# Logical Dynamics and Dynamical Systems

**RASMUS K. RENDSVIG**

**DEP. OF PHILOSOPHY | FACULTY OF HUMANITIES AND THEOLOGY | LUND UNIVERSITY**

Logical Dynamics and Dynamical Systems

# Logical Dynamics and Dynamical Systems

by Rasmus K. Rendsvig

| Organization **LUND UNIVERSITY** | Document name **DOCTORAL DISSERTATION** |
|---|---|
| Department of Philosophy<br>Box 192<br>221 00 Lund<br>Sweden | Date of disputation<br>2018-03-16 |
| | Sponsoring organization<br>The Swedish Research Council |
| Author(s)<br>Rasmus K. Rendsvig | |

Title and subtitle
Logical Dynamics and Dynamical Systems

Abstract

This thesis is on information dynamics modeled using *dynamic epistemic logic* (DEL). It takes the simple perspective of identifying models with maps, which under a suitable topology may be analyzed as *topological dynamical systems*. It is composed of an introduction and six papers. The introduction situates DEL in the field of formal epistemology, exemplifies its use and summarizes the main contributions of the papers.

Paper I models the information dynamics of the *bystander effect* from social psychology. It shows how augmenting the standard machinery of DEL with a decision making framework yields mathematically self-contained models of dynamic processes, a prerequisite for rigid model comparison.

Paper II extrapolates from Paper I's construction, showing how the augmentation and its natural peers may be construed as maps. It argues that under the restriction of dynamics produced by DEL dynamical systems still falls a collection rich enough to be of interest.

Paper III compares the approach of Paper II with *extensional protocols*, the main alternative augmentation to DEL. It concludes that both have benefits, depending on application. In favor of the DEL dynamical systems, it shows that extensional protocols designed to mimic simple, DEL dynamical systems require infinite representations.

Paper IV focuses on *topological dynamical systems*. It argues that the *Stone topology* is a natural topology for investigating logical dynamics as, in it, *logical convergence* coinsides with topological convergence. It investigates the recurrent behavior of the maps of Papers II and III, providing novel insigfhs on their long-term behavior, thus providing a proof of concept for the approach.

Paper V lays the background for Paper IV, starting from the construction of metrics generalizing the Hamming distance to infinite strings, inducing the Stone topology. It shows that the Stone topology is unique in making logical and topological convergens coinside, making it *the* natural topology for logical dynamics. It further includes a metric-based proof that the hitherto analyzed maps are continuous with respect to the Stone topology.

Paper VI presents two characterization theorems for the existence of *reduction laws*, a common tool in obtaining complete dynamic logics. In the compact case, continuity in the Stone topology characterizes existence, while a strengthening is required in the non-compact case. The results allow the recasting of many logical dynamics of contemporary interest as topological dynamical systems.

Key words
formal epistemology, modal logic, dynamic epistemic logic, dynamical systems, general topology

Classification system and/or index terms (if any)

| Supplementary bibliographical information | Language<br>English |
|---|---|
| ISSN and key title | ISBN<br>978-91-88473-70-7 (print)<br>978-91-88473-71-4 (pdf) |
| Recipient's notes | Number of pages<br>226 | Price |
| | Security classification | |

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature _____

Date ____2018-02-09____

# Logical Dynamics and Dynamical Systems

by Rasmus K. Rendsvig

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

# Contents

# List of Papers

This thesis includes the following papers, referred to by their Roman numerals. Authors are listed in alphabetic order. All authors contributed to all parts equally.

I   **Pluralistic Ignorance in the Bystander Effect: Informational Dynamics of Unresponsive Witnesses in Situations calling for Intervention**

    Rasmus K. Rendsvig
    Synthese 191(11): 2471-2498, 2014.

II   **Model Transformers for Dynamical Systems of Dynamic Epistemic Logic**

    Rasmus K. Rendsvig
    Logic, Rationality and Interaction (LORI 2015), W. van der Hoek, W. H. Holliday, and W. Wang (eds.), pp. 316–327, Springer, 2015.

III   **Intensional Protocols for Dynamic Epistemic Logic**

    Hanna S. van Lee, Rasmus K. Rendsvig and Suzanne van Wijk
    Submitted November 7th, 2017.

IV   **Convergence, Continuity and Recurrence in Dynamic Epistemic Logic**

    Dominik Klein and Rasmus K. Rendsvig
    Logic, Rationality and Interaction (LORI 2017), Baltag, A., Seligman, J. and Yamada, T. (eds.), pp. 108–122, Springer, 2017.

V   **Metrics for Formal Structures, with an Application to Kripke Models and their Dynamics**

    Dominik Klein and Rasmus K. Rendsvig
    Submitted December 31, 2017.

VI   **Characterizations of Reduction Law Existence**

    Rasmus K. Rendsvig
    Draft.

Papers I, II and IV are reproduced with kind permission of Springer.

The following papers were written during the doctorant period, but have not been included in the thesis due to lower overlap with the main theme.

**Diffusion, Influence and Best-Response Dynamics in Networks: An Action Model Approach**

Rasmus K. Rendsvig
In Ronald de Haan (ed.) *Proceedings of the ESSLLI 2014 Student Session*.

**Dynamic Epistemic Logic of Diffusion and Prediction in Threshold Models**

Alexandru Baltag, Zoé Christoff, Sonja Smets and Rasmus K. Rendsvig
Accepted for publication in *Studia Logica*.

**Hintikka's *Knowledge and Belief* in Flux**

Vincent F. Hendricks and Rasmus K. Rendsvig
In Hans van Ditmarsch and Gabriel Sandu (eds.) *Jaakko Hintikka on Knowledge and Game Theoretical Semantics*, Outstanding Contributions to Logic Series, p. 317–337, Springer, 2018.

**The Philosophy of Distributed Information**

Vincent F. Hendricks and Rasmus K. Rendsvig
In Luciano Floridi (ed.) *The Routledge Handbook of Philosophy of Information*, Routledge, 2016.

**Youths Exhibit Misperceptions of Norms, False Consensus and Pluralistic Ignorance with respect to Cyberbullying**

**Rasmus K. Rendsvig**, Winnie Alim and Vincent F. Hendricks[*]
Submitted.

# Acknowledgements

# Introduction

This thesis starts with an example.

You have just read a declarative sentence, the first in this section. Perhaps it shaped your idea of what is to come: You might now expect an example, where before you did not. If you've read these lines before, it may have changed nothing. In either case, you partook in a *dynamics*: You began reading this section in some state of mind. You then exposed that state to a piece of information, embodied by the first sentence. Whether affected or not, the result again is a belief state. The pattern is this: Belief state, information, belief state. That is an example of an *information dynamics*.

The topic of this thesis is information dynamics and how to formally model them. To be more precise, the topic is how to model information dynamics using *dynamic epistemic logic*. The perspective investigated is simple, yet not quite community practice: Formalize the model as a *mapping* that allows for the formulation of reduction laws. Then, with a bit of math sprinkled on top, the model will be a *topological dynamical system*.

With that, the main contribution of this thesis is not an analysis of a philosophical paradox or the development of a methodology in the philosophy of science. Nor is it a set of logics for this purpose or that. Rather, the contribution is the formulation of a bond between relatives—from dynamic epistemic logic to the sister field of topology and dynamical systems. With that, the thesis makes a contribution to logic and formal epistemology.

Why the perspective is reasonable, and why the stated consequence follows, is what the thesis argues. This preliminary chapter situates dynamic epistemic logic in the field of formal epistemology in Section 1, exemplifies its use in modeling in prerequisite-free manner in Section 2, and presents the individual papers constituting the thesis, couching them in preceding considerations, in Sections 3–8. The final Section 9 concludes the chapter, summarizing the contributions of the individual papers and listing venues for further research.

# 1   Formal Epistemology and Dynamic Epistemic Logic

'Dynamic epistemic logic' has both a narrow and a broad reading. In the narrow reading, the term refers to a class of *formal logics*, touched later. In the broad sense, it refers to a research area in a broader field labeled *formal epistemology*.

Formal epistemology, in turn, is broadly characterized as a research field by both its subject matter and its methodology. The subject matter is *epistemology*, the study of knowledge, justification, rational beliefs, and related notions, questions and problems, for single agents and for groups. This is where the *epistemic* in dynamics epistemic logic enters: Epistemic is of the Greek *epistēmē*, knowledge. The methodology of formal epistemology is *formal*, employing a suite of tools from formal logic, probability theory, decision and game theory, recursion and complexity theory, various simulation approaches, and more.

Formal epistemology is—relative to other areas of philosophy—a young field of research. In the first anthology on the subject, the 2016 *Readings in Formal Epistemology* [5], the earliest paper included is F. P. Ramsay's paper *Truth and Probability* from 1921, and the earliest of the remaining 36 papers is an excerpt from David Lewis' 1969 *Convention*. The denomination 'formal epistemology' itself stems from the 1990's,[1] but what is to be counted as work falling under the heading of course depends on formal epistemology's subject matter and methodology.

Just as its arsenal of methods vary from those of mainstream epistemology, so has the subject matter of formal epistemology taken on a non-classic tinge: Apart from including formal takes on traditional epistemological problems, the field is further influenced by topics in neighboring disciplines like game theory, computer science, linguistics, and the cognitive and social sciences. As a result, the field concerns not only traditional epistemic notions like knowledge, belief and justification, but also less traditionally epistemological topics like rationality in decision making and strategic interaction, cryptography, preference dynamics, abstract argumentation and formal pragmatics, information pooling and judgment aggregation, learning, and social influence and network effects, to mention but a few [5].

Knowledge, belief and information stand central in formal epistemology, and these concepts are typically discussed in relation to an idealized or hypothetical *agent*—or more than one, in many cases. In these cases, the *rational agency* of the agent is often a matter of scrutiny: It may be discussed whether a given way of making decisions is rational, or whether a given way of incorporating new information is rational. In such cases, it is for most investigations a prerequisite to fix an *epistemic representation* of agents. It is a prerequisite to choose a formal representation of the propositional attitudes agents hold to pieces of information, be these attitudes knowledge, belief, justification, awareness, or other. Taken

---

[1]Cf. the introductions of [5, 55].

at large, two main families of approaches exist, the *probabilistic* and the *logical*.

Among probabilistic approaches, the most widely adopted doctrine is *Bayesianism*. For the Bayesian epistemologist, the primary attitude of interest is *belief*. The doctrine holds that agents' beliefs come in *degrees*: An agent's belief in a proposition is assigned a number, and this number must fall between 0 and 1. This is a consequence of the first of two core normative tenets of Bayesianism: An agent's degrees of beliefs must obey the *laws of probability*. A Bayesian agent's beliefs are thus identified with a *probability distribution*, and the degree of belief in a proposition is identified with the probability assigned to that proposition by the distribution. With $P$ the probability distribution for some agent's beliefs, the agent's belief in the proposition $A$ could then be $P(A) = 0.8$, for example. The second normative tenet of Bayesianism specifies how agents should change their beliefs in the light of new information: When an agent learns of new evidence, it should update its beliefs by *conditioning* in accordance with *Bayes' rule*. Jointly, these two requirements characterize the *Bayesian epistemically rational agent*. So, at least, goes for the rough presentation: Bayesianism is not a position without internal debate, cf. e.g. [149].

If one thinks that it is an unduly high demand that agents must be able to specify an exact probability distribution for their beliefs in order to be rational, the field of *imprecise probabilities* [46, 69, 78, 145] offers an alternative approach. Using imprecise probabilities, agents' beliefs are represented using sets of probability distributions. In this case, the agent above would not have to specify that it's degree of belief in $A$ is *exactly* 0.8. Instead, it could be more imprecise and accept all probability distributions that assign between 0.75 and 0.85 to $A$, for example, as constituting it's beliefs. Alternatively, there are logical approaches.

Within the logical approaches to formal epistemology, the predominant framework is *modal logic*. The starting point of a modal logic is its *language*. A prototypical modal logical language is given by first fixing a set of *propositional variables* $p_1, p_2, \ldots$. For applications, these are given specific readings: For example, $p_1$ could be taken to represent the proposition 'The trashcan is green' and $p_2$ the proposition 'The train is approaching', etc. The propositional variables represent propositions which in the language are represented in no finer detail: They are therefore often referred to as *atomic propositions* or simply *atoms*.[2] Apart from the atoms, modal languages also include formulas with more complex structure. The full set of formulas may be given using a *Backus-Naur form*:

$$\varphi := p \mid \neg\psi \mid \psi \wedge \psi' \mid \Box\psi$$

The Backus-Naur form gives a recursive definition of the formulas of the modal language under construction. It says that $\varphi$ is a formula if it is a propositional variable, or if it is of one of the forms presented where $\psi$ and $\psi'$ are already formulas. The symbol '$\neg$' is a

---

[2]Grammatically and conceptually, the examples contain additional structure. To reflect it, one may move to richer languages, like *first-order languages*, which may also be made modal.

negation and '∧' a conjunction. In most modal logical application, they are given very straightforward readings of 'not' and 'and'. From these, one can define disjunction '$\psi \vee \psi'$' ('$\psi$ or $\psi'$') and implication '$\psi \rightarrow \psi'$' ('if $\psi$, then $\psi'$') and other Boolean connectives.

The tricky bit lies in the the *modal operator* $\square$, often referred to as the *box*. Apart from the interpretation of the atoms, the understanding of this operator details what the modal logic is about. In the majority of interpretations, the box makes an *universal* statement about things seen from a *local perspective*. To exemplify, think of standing at a train station, tracks leading to neighboring stations. Reading $p_1$ as 'The trashcan is green', the formula $\square p_1$ could be read 'At *every* train station one stop away, the trashcan is green', for example. Dual to this universal statement is an *existential* statement: 'There *exists* a train station one stop away where the trashcan is green'. With the current interpretation of the box, this statement is captured by the formula $\neg\square\neg p_1$: '*Not* at every train station one stop away is the trashcan *not* green'. Such statements are used so often that '$\neg\square\neg$' is typically shortened to '$\Diamond$', which is referred to as the *diamond*. The diamond is called the *dual* of the box. More directly, the formula $\Diamond p_1$ is read 'There *exists* a train station one stop away where the trashcan is green'. To further exemplify, the modal language also contains formulas that describe stations more than just *one* stop away, for example $\Diamond\Diamond p_1$ and $\square\Diamond p_1$. The first states that going two stops can bring one to a green trashcan, while the second state at every train station one stop away, there exists a train station one stop away where the trashcan is green. If the trashcan is green at the current station and the trains go both ways along the tracks, $\square\Diamond p_1$ would, in this example, be making a true statement.

The train example serves two purposes: First, thinking of the modal operators as talking about train stations reachable from the current position will hopefully aid intuitions when the type of formal models primary to this thesis, namely *pointed Kripke models*, in Section 2 are introduced as abstract semantics for modal logic. Second, the example is meant to illustrate the interpretational freedom of modal operators: They are not inherently doxastic, epistemic, metaphysical, or other. They are box and diamond, until further notice.

In formal epistemology, the most common interpretation of the box operator is that it represents either *knowledge* or *belief*. As these concepts are often of interest simultaneously, it is not rare that the languages used are *multi-modal*: They include one box operator for knowledge and one for belief. The Backus-Naur form for a prototypical epistemic-doxastic language is then

$$\varphi := p \mid \neg\psi \mid \psi \wedge \psi' \mid K_i\psi \mid B_i\psi$$

with $K_i$ and $B_i$ both boxes. The subscript $i$ refers to an agent: The formula $K_i\psi$ reads 'Agent $i$ knows that $\psi$' and $B_i\psi$ reads 'Agent $i$ believes $\psi$'. The $i$ is a variable, like the $p$, ranging over some specified set of agents, say $\mathcal{A}$. If $\mathcal{A} = \{1, 2, 3\}$, then the language would be build on 6 box operators in total and allow formulas like $B_1K_2\psi$—that agent 1 believes that agent 2 knows $\psi$. When the set $\mathcal{A}$ contains more than one agent term, the language is not only multi-modal, but also *multi-agent*. In any case, the language allows

for the expression of *higher-order information*, i.e., information about information, and so forth. The single-agent formula $K_a K_a \varphi$ illustrates the point: It expresses that agent $a$ knows something about it's own information, namely that it knows $\varphi$.

From the above Backus-Naur form, the sharpest contrast between the probabilistic and logical approaches is apparent: Using a language as the above, beliefs do not come in degrees—they are *all-or-nothing*.[3] That beliefs are all-or-nothing means that for every proposition, either the proposition is believed or it is not: There is nothing in between, no degrees of belief. Following the Backus-Naur form, graded beliefs are simply not *expressible*: The language includes the formula $B_i \psi$ expressing that $i$ believes $\psi$ and it includes the formula $\neg B_i \psi$ expressing that $i$ does not believe $\psi$, but no formula expressing that $\psi$ is believed to some degree. All-or-nothing beliefs are also referred to as *qualitative* beliefs, in contrast with the *quantitative* beliefs of probabilistic approaches.[4]

The first of the two normative tenets of Bayesianism provides an answer to what constitutes rational quantitative beliefs: The beliefs must be probabilistic. In *epistemic logic*, similar foundational questions arise concerning the governing principles for *rational knowledge* and *rational all-or-nothing belief*, considered as modal logical notions. First to provide a systematic answer was the Finnish philosopher and logician Jaakko Hintikka (1929-2015) in his seminal book from 1962, *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, cited here as [91].[5]

Hintikka's analysis starts from the commitment made by a speaker when uttering "I know that $\varphi$". Asking Hintikka, the speaker not only commits to the truth of $\varphi$, but also to being in an evidential situation with conclusive grounds strong enough to warrant the claim.[6] Hintikka's subsequent analysis feeds strongly on a particular consequence of such grounds:

> If somebody says "I know that $p$" in this strong sense of knowledge, he implicitly denies that any further information would have led him to alter his view. He commits himself to the view that he would still persist in saying that he knows that $p$ is true – or at the very least persist in saying that $p$ is in fact true – even if he knew more than he now knows. [91, p. 18]

This consequence of the speaker's commitment lies at the heart of Hintikka's analysis, and

---

[3]A qualification: Logical approaches exist where belief are not all-or-nothing. These include multi-valued logics and probabilistic logics, where varieties of graded beliefs may be represented.

[4]Strong connections exist between the two: Recent works bridge rational quantitative and qualitative beliefs [99, 112], show correspondence between structures modeling higher-order information [64], and design dynamic epistemic logics with probabilities [29, 101].

[5]Hintikka was not the first to consider an epistemic application of modal logic. Seemingly, that was his doctoral adviser, Georg Henrik von Wright (1916-2003) in *An Essay in Modal Logic* (1951), [150], cf. [43, 86].

[6]Hintikka takes a pragmatic view of conclusive grounds: "We must realize, however, that having this right [to claim knowledge] need not mean that one's grounds are so strong that they logically imply that what one claims to know is true. It may merely mean that the grounds one has are such that any further inquiry would be pointless for the normal purposes of the speakers of the language." [91, p. 17-18], footnotes omitted.

throughout *Knowledge and Belief* he routinely returns to re-castings of it. In particular, it is decisive for the nature of an *epistemic alternative*:

> The conditions into which we are trying to catch the logic of knowledge and belief are in terms of certain alternatives to a given state of affairs. Roughly speaking, these alternatives are possible states of affairs in which a certain person knows at least as much as – and usually even more than – he knows in the given state. In short, we are concerned with the different possibilities there are for somebody to gain further information. [91, p. 44][7]

Hence, the commitment on the side of the speaker is one relating his current informational state to other such states—in essence, those in which the speaker has the same or more information. In the train metaphor, the epistemic alternatives are the stations quantified over with the epistemic operators.

In his epistemological program, Hintikka held that the axioms or principles of epistemic logic serves as conditions of a special kind of general (strong) rationality: The statements which may be proved false by application of the epistemic axioms are not inconsistent in the sense that their truth is logically impossible. They are rather rationally 'indefensible'. Indefensibility is annexed as the agent's epistemic laziness, sloppiness or perhaps cognitive incapacity whenever to realize the implications of what he in fact knows:

> In order to see this, suppose that a man says to you, 'I know that $p$ but I don't know whether $q$' and suppose that $p$ can be shown to entail logically $q$ by means of some argument which he would be willing to accept. Then you can point out to him that what he says he does not know is already implicit in what he claims he knows. If your argument is valid, it is irrational for our man to persist in saying that he does not know whether $q$ is the case. [91, p. 31]

Defensibility thus means not falling victim of 'epistemic neglience' [38]. The notion of indefensibility gives away the status of the epistemic axioms and logics embraced by Hintikka in *Knowledge and Belief*. Some epistemic statement for which its negation is indefensible is called 'self-sustaining'. The notion of self-sustenance corresponds to the concept of *validity*. Corresponding to the self-sustaining statement is the logically valid statement. This, in turn, is be a statement which it is rationally indefensible to deny. Hence, to Hintikka, epistemic axioms are descriptions of epistemic rationality.

For an idea of Hintikka's position, and for later reference, a selection of epistemic principles discussed by Hintikka are presented in Table 1. The table also includes principles not directly touched on in [91], but which are standard in epistemic and modal logic at large.

---

[7]The formulation "knows at least as much" should be taken to mean "the same and possibly more", not "more" given some measure.

**Table 1:** Standard modal axiom schemes given for the knowledge operator, $K_a$. In the doxastic reading, replace $K_a$ by $B_a$.

| Axioms | | Rule of inference | |
|---|---|---|---|
| K | $K_a(\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$ | N | From $\vdash \varphi$, infer $\vdash K_a\varphi$ |
| D | $\neg K_a(\varphi \wedge \neg\varphi)$ | | |
| T | $K_a\varphi \rightarrow \varphi$ | | |
| 4 | $K_a\varphi \rightarrow K_aK_a\varphi$ | | |
| 5 | $\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$ | | |

To evaluate the self-sustainability of epistemic and doxastic principles, Hintikka introduces a semantic construct called a *model system*. A model system consists of a set of *model sets*—partial descriptions of the state of affairs using the epistemic language, including the state of the agents knowledge and beliefs—and a relation between them. For knowledge, a model system's relation is interpreted as relating *epistemic alternatives*—formally model sets—where one is *epistemically accessible* from another just in case the agent in the former holds *at least as much information* as in the latter.[8]

Hintikka's model systems are reminiscent of to the relational, possible worlds semantics of *Kripke models* (see Section 2), but Hendricks and I have recently argued (see [89]) that—when taking Hintikka's epistemic program into consideration—identifying model systems with Kripke models mis-represents important aspects of Hintikka's analysis. Rather, we argue, a model system could be seen as a set of Kripke models linked by a relation obtained through *arbitrary public announcements* [9, 50, 42, 108, 153].

Of the principles listed in Table 1, Hintikka explicitly discusses T, 4 and 5, endorsing T and 4 for knowledge and 4 for belief, while explicitly rejecting 5 for both. Endorsement is a result of the self-sustainability in model systems. In case of knowledge, the principle T, capturing the veracity of knowledge, is self-sustaining as one any epistemic alternative is an alternative to itself: In any situation, an agent will have a least as much information as it in fact has. The principle 4 is argued for by appeal to the *transitivity* of the epistemic accessibility relation: In an agent holds at least as much information in $w_1$ as it does in $w_0$ and at least as much information in $w_2$ as in $w_1$, it must hold at least as much information in $w_2$ as in $w_0$. From this, it follows that 4 is self-sustaining—but *not* a product of introspection, as the principle is often taken to be.[9] Contra 5, the argument also uses the accessibility relation: There exists, Hintikka argues, situations where the relation is not *symmetric*, i.e., where an agent may have *strictly more* information in $w_1$ than in $w_0$. In this case, $w_0$ is *not* accessible from $w_1$. If in $w_1$ the agent has attained enough information to obtain knowledge onwards, while not having so in $w_0$, principle 5 would not be satisfied—but again, not due to arguments from introspection.

Of the remaining principles, not explicitly discussed by Hintikka, both K and N seem to fit

---

[8] The interpretation relating to beliefs is omitted here, as they are secondary in the work: See [91] for details.

[9] Hintikka is very explicit on this point and includes a discussion of philosophical position pro introspection.

the spirit of his epistemic program, cf. [89], and so does D for belief: The quote immediately strongly indicates that he would support logical closure of for rational knowledge, and I would suppose the same—together with consistency—for belief.[10]

From Hintikka's work in *Knowledge and Belief*, formally informed discussions have permeated through both mainstream and formal epistemology [88]. Questions raised includes, for example, the acceptability of closure principles, like the axioms K or 4 above, by e.g. Nozick [126], Williamson [154] and Holliday [94, 95], the problem of logical omniscience and suggestions for its resolve by e.g. Hintikka [92] and Rantala [134], and the acceptability of alternative epistemic principles by e.g. Lenzen [113] and Stalnaker [147].

Though the research questions in contemporary epistemic logic may vary substantially from the work of Hintikka, there is a strong sense in which the general methodology of *Knowledge and Belief* is seen in much recent work: The logic of a given epistemic concept is investigated by constructing a suitable formal language and a model theoretic representation of said concept over which one may seek the set of validities and a corresponding axiomatic base. A recent work along these lines in Baltag and Smets' [17], where "weak non-introspective knowledge", introduced by Stalnaker [146, 148] to modally capture Lehrer's defeasibility analysis of knowledge [110, 111] is analyzed.

Terminologically, contemporary research makes claims less bold than that of Hintikka and often they concern knowledge or belief in varying circumstances, ranging from reasoning about finite cases occurring in the now, like uncertainty about the current state of play in games [19, 23, 25, 49, 129, 140, 147], to uncountable cases in infinitely evolving temporal settings, like the learning of scientific theories [67, 68, 87, 98]. With such varying structures the objects of epistemic analyses, the main focus in contemporary research in epistemic logic is not to find an axiomatization of *the* logic of knowledge and belief. Rather, it focuses on logically capturing different types of knowledge, belief, certainty and uncertainty, and their interplay to compare them systematically [17, 90].

With several contemporary research topics requiring an underlying logic of knowledge and belief, two systems very strong in terms of rationality requirements have become obtained a role close to the community standard. These are the systems S5 for knowledge and KD45 for belief, axiomatically given by the systems attributed to Hintikka above, in both cases plus the axiom 5.

Also in terms of models especially one framework has become a standard, namely the relational, possible worlds semantics offered by Kripke models, accredited to Saul Kripke's 1963 [107], cf. [10]. Kripke models also constitute the basic semantic approach pursued in this thesis, and are revisited in greater detail in Section 2 together with their close-to-

---

[10]Jointly, these principles do not quite constitute a *modal logic* in the standard technical reading, cf. [89]. This is due to the partiality of models sets.

standard epistemic interpretation based on *informational indistinguishabilty*. The interpretation holds that an agent know a given proposition if that proposition is true in all the states of the world not ruled out by current evidence, i.e., the states indistinguishable from the actual state of affairs. Evidence thus *partitions* the state space into sets respectively compatible and incompatible with current information. The logic of the resulting knowledge operator is S5, with an accompanying "near-partitioning" of states in the case of belief, accommodating that such may be false, following a KD45 logic.

Historically, the idea of knowledge as given by the partitioning of the set of states based on present information has traces back to Hintikka's 1962, with the idea penetrating his authorship: See for example his 2007 [93]. Mathematically, the partition approach to knowledge goes back at least to Aumann's 1976 paper *Agreeing to Disagree* [8] while the logic S5 as the default for knowledge reached a large audience through Fagin, Halpern, Moses and Vardi's seminal 1995 monograph *Reasoning about Knowledge* [58]. The use of S5 as the logic of knowledge has older roots: It was used for modeling by McCarthy et al. in 1978 [122]. On the side of belief, KD45 was suggested as early as 1984 by Levesque [115, 114], while the semantic plausibility interpretation dates back to Grove's 1988 paper [73], discussed below.

The two tenets of Bayesianism—that rational beliefs are probabilistic and that rational belief change is by conditionalization—reflect two perspectives both of interest also to logical approaches. The first is a *static* perspective, concerning properties of epistemic states. The second is *dynamic*, concerning how epistemic states change under informational influence.

In dynamic terms, knowledge and belief differ in that the former is more conservative than the latter: If a proposition is known, no new information will change this, but if it is merely believed, new information may cause the belief to be dropped. The process of how beliefs should rationally change in the light of new information has been a topic of considerable interest to philosophers and computer scientists alike. The topic and surrounding field of research is commonly referred to as *belief revision.*

Following earlier work by Isaac Levi and William L. Harper, the formal study of belief revision was initiated by Carlos Alchourrón, Peter Gärdenfors and David Makinson in a series of publications through the 1980's [2, 4, 76, 77]. The formal take on belief revision developed through these papers – which have varying configurations of subsets of the three as authors – is often referred to as *AGM belief revision.*

Where the prototypical modal logical language for knowledge and belief allows for both multiple agents and the expression of higher-order information, the subject of study for AGM belief revision is a single agent's beliefs about a set propositions not taken to include statements about the agent's own beliefs. For expositional purposes, the Backus-Naur form

$$\varphi := p \mid \neg\psi \mid \psi \wedge \psi'$$

may be taken to specify the language of AGM belief revision, though the approach does

require the language to be propositional. Denote this language—i.e.,the set of expressions deemed formulas by the Backus-Naur form—by $L$.

Further, fix a classic logic $\Lambda$ in the language given and let $\vdash$ be the *proof relation* of $\Lambda$: Where $\Gamma$ is a set of formulas, $\Gamma \vdash \varphi$ means that the formula $\varphi$ is provable from the assumption of the premises $\Gamma$ in the logic $\Lambda$. Denote by $Cn(\Gamma)$ the set of *logical consequence* of $\Gamma$, i.e., the set of formulas provable in $\Lambda$ from $\Gamma$.

From this basis, AGM belief revision represents an agent's beliefs as a set $K$ of formulas, satisfying two rationality criteria for static belief: The set $K$ must be *consistent* and it must be *logically closed*. I.e., the set $Cn(K)$ must not contain a contradiction and $K$ must equal $Cn(K)$. A set $K$ satisfying these two criteria is called a (nonabsurd) *belief set*.[11]

Belief sets represent the beliefs of the agent in question. Like the modal logical approach, belief sets represent all-or-nothing beliefs. In the terminology of Gärdenfors [76], a belief set may express three different attitudes to a formula $\varphi$:

  *i*) If $\varphi \in K$, then $\varphi$ is *accepted*.
  *ii*) If $\neg\varphi \in K$, then it is *rejected*.
  *iii*) If neither $\varphi \in K$ nor $\neg\varphi \in K$, then $\varphi$ is *indetermined*.

These three attitudes correspond respectively to $B_a\varphi$, $B_a\neg\varphi$ and $\neg B_a\varphi \wedge \neg B_a\neg\varphi$ from the modal logical approach, recalling that the formula $\varphi$ may not contain any modal operators. The AGM approach thus treats a similar and simpler static belief notion.

On the dynamic side, AGM belief revision is concerned with three types of transitions possible between these attitudes towards a proposition $\varphi$.

The first transition concerns moving from a belief state indeterminate about $\varphi$ to one determinate about $\varphi$, i.e., to one where $\varphi$ is either accepted or rejected. This is the simplest form of belief change, referred to as *expansion*. Expansion is simple as it is both *conflict-free* and *unique*: Adding $\varphi$ to $K$ causes no inconsistencies, so no old beliefs need be removed due to conflict, and there is a unique minimal way to ensure that $K$ expanded with $\varphi$ is a belief set: Taking it's logical closure.[12]

The second transition is in the opposite direction, where the agent gives up a belief about $\varphi$. It is the transition moving from a belief state determined about $\varphi$ to one indetermined about $\varphi$, retracting the belief in $\varphi$ (or $\neg\varphi$), but not adding new beliefs. The operation is referred to as *contraction*. Contraction is not as simple as expansion: Though no conflicts may arise when removing $\varphi$ from $K$, the operation of doing so *successfully* is not unique. Obstructing success is that $\varphi$ may be logically implied by other formulas in $K$, some of

---

[11]Cf. [76, p. 24]. The notation $K$ for the belief set is standard notation, used systematically already in the early [76]. It should *not* be confused with the modal operator for knowledge.

[12]The operation may also be given an equivalent basis in terms of rationality postulates: See [76, p. 48-51].

**Table 2:** The Gärdensfors/AGM rationality postulates for contraction functions.

*The basic contraction postulates*

C1     $K - \varphi$ is a belief set (i.e., $K - \varphi = Cn(K - \varphi)$)     (Closure)
C2     $K - \varphi \subseteq K$     (Inclusion)
C3     If $\varphi \notin K$, then $K - \varphi = K$     (Vacuity)
C4     If not $\vdash \varphi$, then $\varphi \notin K - \varphi$     (Success)
C5     If $\varphi \in K$, then $K \subseteq (K - \varphi) + \varphi$     (Recovery)
C6     If $\vdash \varphi \leftrightarrow \psi$, then $K - \varphi = K - \psi$     (Extensionality)

*The supplementary contraction postulates*

C7     $(K - \varphi) \cap (K - \psi) \subseteq K - (\varphi \wedge \psi)$     (Conj. Overlap)
C8     If $\varphi \notin K - (\varphi \wedge \psi)$, then $K - (\varphi \wedge \psi) \subseteq K - \varphi$     (Conj. Inclusion)

which must *also* be removed to ensure that $\varphi$ is not re-added by logical closure. If, e.g., $K$ contains $\varphi, \psi$ and $\psi \to \varphi$, then removing only $\varphi$ will not suffice—either $\psi$ or $\psi \to \varphi$ must be removed as well. Hence the complexity of contraction: What beyond $\varphi$ should be removed?

The third transition is between two belief states, the first where $\varphi$ is accepted, the second where $\varphi$ is rejected (or *vice versa*). This is the case of *revision*: In accepting $\neg\varphi$ (i.e., rejecting $\varphi$), the currently held belief that $\varphi$ must, on pain of contradiction, be given up. Hence, for reasons similar to the case of contraction, a unique revision operation involves a choice not settled on logical grounds alone.

AGM belief revision theory does *not* settle the question of how to uniquely contract or revise a given belief set. Rather, it posses a set of criteria that any *contraction function* or *revision function* that does make a unique choice should satisfy to be deemed rational.

Contraction and revision functions are, respectively, operationalizations of the contraction and revision transitions. Both function types take as input a belief set and a formula and both output a belief set. Hence they are functions $f\colon \mathcal{K} \times L \longrightarrow \mathcal{K}$, where $\mathcal{K}$ is the set of all possible belief sets. One not uncommon notation is to use $-$ for contraction functions and $*$ for revision functions. Additionally, infix notation is used, with $K - \varphi$ the output of contracting $K$ with $\varphi$ using the contraction function $-$. Similarly, $K * \varphi$ is the output of $*$ applied to $K$ and $\varphi$. Lastly, also expansion is considered a function with notation $+$.

An example of a rationality postulate for a contraction function is *Success*, hinted at above. Success requires that if $\varphi$ is not a logical truth (theorem), then the contracted belief set $K - \varphi$ should *not* include $\varphi$. The formal representation of Success and the additional postulates for contraction may be found in Table 2.

In Gärdenfors' 1988 monograph *Knowledge in Flux* [76], a conservative principle motivates the rationality postulates: The heuristic criterion of *informational economy* requires that when changing beliefs, these should be changed in a fashion minimal to obtain the desired

effect.[13] The criterion is often referred to as the principle of *minimal change*, and is now standard [139]. Minimal change, in turn, is motivated by a simple rationality consideration: Information seldom comes for free, so unnecessary information loss is undesired.

The basic principles may be argued for as follows: C1 captures that the result of revision is a rational belief set, which seems reasonable for rational contraction. C2 states that no new beliefs should be *added* when something is to be removed: It may be seen as a consequence of minimal change. So may C3, stating that if $\varphi$ is not believed in the first place, then there is not reason to change anything. C4 captures that contraction should be successful, mentioned above. C5 is more controversial: It states that when $\varphi$ is believed, if it is first retracted, but then re-added, then *all* beliefs possibly lost by the first contraction are recovered. With minimal change applying to both contraction and expansion, the principle may seem valid: It is however the cause of controversy, with an early counterexample due to Sven-Ove Hansson [84]; Rott and Pagnucco investigate contraction without recovery [138]. C6, finally, states that the revision of agents are concerned with the *content* and not the *form* of the information they receive.

The most striking contribution of *the* seminal AGM paper [2]— the unique paper authored by all three authors—is a *representation theorem* for the class of contraction functions that satisfy rationality postulates C1–C6 of Table 2. The class of contraction functions considered rational, on this account, is the explicitly constructable class of *partial meet contraction functions*.

Fixing a class of functions for contraction equally fixes a class of functions for revision and *vice versa*. At least this holds true if one is willing to accept respectively the *Levi* and *Harper identities*. The Levi identity originates from Isaac Levi (1977), who argued that only expansion and contraction are legitimate forms of belief change from which revision should be analyzed as a composite. In particular, revision with $\varphi$ is equivalent with first contracting beliefs in $\neg\varphi$, followed by expansion with $\varphi$. I.e., revision is given by the identity

$$K * \varphi = (K - \neg\varphi) + \varphi.$$

The Levi identity thus allows the definition of revision from contraction and expansion. With the mentioned result of [2], this thus fixes a class of revision functions. Moreover, as shown in [76], the class of revision functions obtained may equally well be motivated on rational grounds: When the contraction function satisfies C1–C6, the corresponding revision function will satisfy R1–R6 of Table 3, which contains the AGM postulates for revision. As with the contraction postulates, these are justified using informational economy considerations in [76].

---

[13]The AGM postulates trace back to earlier suggestions by Gärdenfors [74, 75] and was conveniently re-represented by Alchourrón and Makinson [3] and Makinson [121].

**Table 3:** The AGM postulates for revision functions.

*The basic revision postulates*

| | | |
|---|---|---|
| R1 | $K * \varphi$ is a belief set (i.e., $K * \varphi = Cn(K * \varphi)$) | (Closure) |
| R2 | $\varphi \in K * \varphi$ | (Success) |
| R3 | $K * \varphi \subseteq K + \varphi$ | (Inclusion) |
| R4 | If $\neg\varphi \notin K$, then $K + \varphi = K * \varphi$ | (Vacuity) |
| R5 | $K * \varphi$ inconsistent only if $K$ or $\varphi$ is | (Consistency) |
| R6 | $\varphi \leftrightarrow \psi$ implies $K * \varphi = K * \psi$ | (Extensionality) |

*The supplementary revision postulates*

| | | |
|---|---|---|
| R7 | $K * (\varphi \wedge \psi) \subseteq K * \varphi + \psi$ | (Subexpansion) |
| R8 | If $\neg\psi \notin K * \varphi$, then $K * \varphi + \psi \subseteq K * (\varphi \wedge \psi)$ | (Superexpansion) |

For the opposite direction, William L. Harper (1977) first advanced the suggestion of defining contraction through revision. The Harper identity captures the relationship:

$$K - \varphi = K \cap K * \varphi.$$

As with the Levi identity, also the Harper identity "preserves" rationality: If a revision function satisfies R1–R6, the contraction function induced by the Harper identity satisfies C1–C6, cf. [76].

Beyond the six basic postulates for respectively contraction and revision, Gärdenfors also suggests two supplementary principles for each operation, postulates C7 and C8 of Table 2, and R7 and R8 of Table 3. These postulates are intended to facilitate *iterated* belief revision under the principle of minimal change, cf. [76, p. 55]. The postulates are pairwise preserved by the Levi and Harper identities [76].[14]

In combination with the the six basic postulates for contraction, the supplementary postulates C7 and C8 characterize the sub-class of partial meet contraction functions that are also *transitively relational* [2]. Partial meet contraction functions are not the only candidates for contraction functions, though. In [76], Gärdenfors shows that revision functions based on *epistemic entrenchment relations* are equally characterized by R1–R8. The corresponding contraction functions are, by Hans Rott in [137], shown equivalent to the *safe contractions* of Alchourrón and Makinson [4].

The approaches to belief revision just mentioned all share the fundamental structure of the AGM approach: They concern contraction and revision functions $f \colon \mathcal{K} \times L \longrightarrow \mathcal{K}$ working on belief sets. A different approach was taken by Adam Grove in his 1986 paper [73], which marks a move towards approaching belief revision in an epistemic logical setting.

Recall that underlying the AGM belief sets was a language $L$ and a logic $\Lambda$. Let $L_\Lambda$ be the set of all *maximal* consistent sets of formulas from $L$. Each such set $A$ may be thought of as

---

[14]I.e., if $-$ satisfies C7 (resp. C8), then the induced $*$ will satisfy R7 (resp. R8) and *vice versa*.
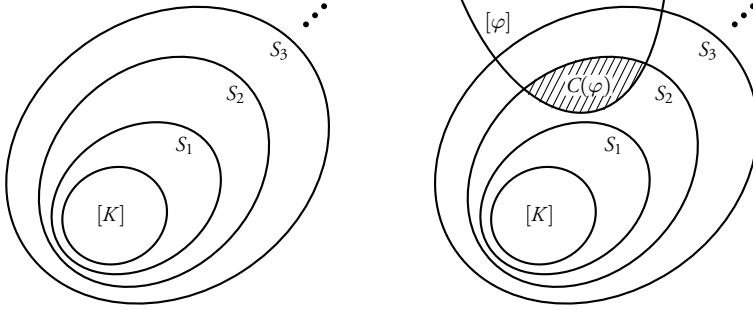
**Figure 1:** A Grove sphere system with center sphere $[K]$ (left) and its revision with $\varphi$ (right).

a *possible world*, assigning a truth value to every formula of the language: If $\varphi \in A$, then $A$ evaluates $\varphi$ as true, and if $\neg\varphi \in A$, then $A$ evaluates $\varphi$ as false. For each maximal consistent set of formulas, exactly one of these two options realizes. To enforce this intuition, denote the elements of $L_\Lambda$ by $w, w'$, etc.

In this construction, the proposition made by a formula $\varphi$ may be identified with the set of possible worlds that makes $\varphi$ true. Formally, this proposition is

$$[\varphi] = \{w \in L_\Lambda : \varphi \in w\}.$$

Likewise, every AGM belief set may be identified with a set of possible worlds. Intuitively, these are the worlds that the agent in question considers the *very most compatible* with the currently held beliefs [73].[15] For the belief set $K$, the set of very most compatible worlds is

$$[K] = \{w \in L_\Lambda : K \subseteq w\}.$$

I.e., a world is very most compatible with the current beliefs if it makes everything currently believed true.

In addition to the worlds very most compatible with current beliefs, Grove builds on the semantics used by David Lewis for counterfactual statements in [117] and adds a *system of spheres* of less and less compatible worlds around the center sphere, $[K]$. Figure 1 illustrates a sphere system while Table 4 contains the formal requirements.

For a sphere system centered on $[K]$, Grove suggests revising the belief set $K$ with $\varphi$ by taking as the new set of very most compatible worlds exactly those among the worlds in $[\varphi]$ that are most compatible with current beliefs. In a sense, the spheres may be thought of as a sequence of *fall-back hypotheses*: When told that the hypothesis the agent holds,

---

[15]Grove does *not* use the terminology "very most compatible". I take responsibility for that. It is to avoid conflation below.

The requirements for $\mathcal{S}$, a Grovian system of spheres centered on a belief set $[K]$.

S1   $\mathcal{S}$ is totally ordered by $\subseteq$: If $S, S' \in \mathcal{S}$, then either $S \subseteq S'$ or $S' \subseteq S$.

S2   $[K]$ is the $\subseteq$-minimum of $\mathcal{S}$: For all $S \in \mathcal{S}$, $[K] \subseteq S$.

S3   $L_\Lambda$ is in $\mathcal{S}$ (and is thus the maximal sphere).

S4   For every formula $\varphi$, if any sphere in $\mathcal{S}$ intersects $[\varphi]$, then there is a smallest sphere that intersects $[\varphi]$.


$[K]$, is wrong—because $\varphi$ is the case—the agent moves to the closest fall-back hypothesis compatible with the new information. The move is presented graphically in Figure 1.

Formally, Grove's suggestion yields a revision function in the following way: Let $C(\varphi)$ be the set $[\varphi] \cap S_\varphi$, where $S_\varphi$ is the $\subseteq$-smallest sphere in $\mathcal{S}$ which has non-empty overlap with $[\varphi]$. By S3 of Table 4, this exists unless $\varphi$ is inconsistent, in which case $C(\varphi)$ is defined to be $L_\Lambda$. Let $K_{C(\varphi)}$ be the belief set consisting of the formulas satisfied at every world in $C(\varphi)$. A revision function may then be defined by, for all $\varphi$ and all $K$,

$$K * \varphi = K_{C(\varphi)}.$$

The main results of Grove's [73] are then that *i)* Any such revision function is AGM rational in the sense of satisfying R1–R8,[16] and *ii)* For any revision function $*$ satisfying R1–R8 and any belief set $K$, there exists a sphere system centered on $[K]$ for which $K * \varphi = K_{C(\varphi)}$ for all $\varphi$. Point *ii)* thus shows that any R1–R8 revision function may be represented by a family of sphere systems containing a system for each belief set $K$.

Two remarks concerning Grove's sphere systems point to aspects outside classic AGM belief revision. The first remark concerns what is given by a Grove revision—and importantly, what is not. What is given is a revised belief set, neatly represented by the set of worlds $C(\varphi)$. What is *not* given is a sphere system centered on $C(\varphi)$. The lack of such a "revised" sphere system means that the formal structure for making a further revision is missing: To $C(\varphi)$, the agent has no fall-back hypotheses. Solving this problem in general—how to capture rational re-revision under the restriction of minimal change—is the study of *iterated belief revision*. The study of this topic was systematically started in the 1990's with important contributions by Boutilier [37], Darwiche and Perle [44] and Nayak [125], cf. [70].

The second remark on Grove's models concerns its relation to epistemic logic, and specifically to the belief operator. For a simple interface, the belief operator $B_a$ may be given a semantics over a Grove sphere system $\mathcal{S}$ centered at $[K]$ representing the beliefs and fall-back hypotheses of the (only) agent $a$ by the following clause:

$$\mathcal{S} \vDash B_a \varphi \text{ iff for all } w \in [K], \varphi \in w.$$

---

[16]Essentially: Grove does not take the logical closure of $C(\varphi)$ to obtain R1 in his Theorem 1; this is done silently by Gärdenfors in his presentation [76, p. 84].

The clause states that the system $\mathcal{S}$ makes $B_a\varphi$ true if, and only if, all the very most compatible worlds makes the formula $\varphi$ true. The clause successfully captures the essence of AGM-style beliefs: $B_a\varphi$ if, and only if, $\varphi$ is in the belief set $K$—at least for any formula $\varphi$ of the *propositional* language of AGM. For more general formulas of the prototypical epistemic language, the right-hand side is not yet well-defined. Omitting the knowledge operator, the situation may be remedied by the clause

$$B_a\varphi \in w \text{ iff for all } w' \in [K], \varphi \in w'$$

stating that agent $a$ believes $\varphi$ *anywhere* in the sphere system $\mathcal{S}$ just in case $a$'s AGM belief set contains $\varphi$.[17]

Baring details, Grove sphere systems with these semantics are a special case of so-called *plausibility models*—a Kripke type models which over the last decade have been of much interest in epistemic logic research. From the epistemic logic community, van Benthem [22] and Baltag and Smets [16] independently introduced *conditional* belief operators on plausibility-like models. A conditional belief formula is of the form

$$B_a^\varphi \psi$$

where both $\varphi$ and $\psi$ are formulas. Like a conditional probability, it is read 'Agent $a$ believes $\psi$, conditional on learning that $\varphi$'.

Over a Grove sphere system $\mathcal{S}$, conditional belief formulas may be evaluated in the style of [22] by the clause

$$B_a^\varphi \psi \in w \text{ iff for all } w' \in \min[\varphi], w' \in \psi$$

where $\min[\varphi]$ is the intersection of $[\varphi]$ with the most compatible sphere $S$ in $\mathcal{S}$ for which this intersection is non-empty. Stated using the AGM notation from above, $\min[\varphi]$ is thus the set $C(\varphi)$. Hence, by the first result of Grove, it follows that the belief revision function $*$ defined by for all $K$, for all formulas $\varphi, \psi$ from the propositional language $L$, $K * \varphi = \{\psi \in L : B_a^\varphi \psi\}$ satisfies the rationality postulates R1–R8.

Using conditional beliefs, one may thus describe elements of AGM belief revision. In *dynamic epistemic logic*, this form of belief revision is often called *static*, as it involves no changes to the model representing beliefs: The model is merely described from a different perspective. Static belief revision stands in contrast with *dynamic belief revision*, which involves a change in the model. In the context of Grove sphere systems, a dynamic belief revision would involve a change to the system of sphere itself. A different update could be a strong knowledge gain, where some of the possible worlds were removed.

In the terminology of Katsuno and Mendelzon [97], static and dynamic belief revision corresponds to the difference between *belief revision* and *belief update*: Belief revision is a

---

[17] The properties of belief under these semantics are captured by the KD45 axioms.

process where both the beliefs and the revised beliefs refer to the same situation, whereas in belief update, beliefs are brought up to date with a changing world. In case of a single agent having beliefs only about the propositional nature of an unchanging world—as in the AGM case—belief revision suffices. However, when multiple agents with higher-order information are in play, the change of one agent's beliefs mitigates a change in the model, which the remaining agents—and the agent itself—must take into account. Hence the case for belief update.

Even a single-agent with higher-order information may serve to illustrate: Consider an agent $a$ that does not explicitly believe some $p$ and who is informed of this fact by a credible source. I.e., the agent is told $p \wedge \neg B_a p$: '$p$ is the case, but you do not believe it.' For this type of information, the *Success* postulate fails: After the update, the agent should rationally believe $p$, but should not believe that it does not believe it. Propositions of this form are called *Moore sentences*, attributed to G.E. Moore. Such *unsuccessful* updates is one topic of interest in dynamic epistemic logic [15, 52, 54, 96, 142].

Belief revision and belief update were not *per se* the topics first studied under the heading of dynamic epistemic logic, though several works have since the mid-2000's dealt explicitly with the topic [7, 16, 17, 18, 32, 22, 48, 54, 67]. Topically, the field comes from a close neighborhood: The paper typically accredited with initiating the field, the 1989 *Logics of Public Communication* by Jan A. Plaza [133], concerned knowledge change in multi-agent systems as the result of so-called *truthful public announcements*.

Truthful public announcements ensure that their content is afterwards *common knowledge* [8, 116]. Continuing in the Grove sphere system framework with a single agent, this would amount to the agent simply knowing the content stated. To illustrate, evaluate the knowledge operator $K_a$ over a Grove sphere system $\mathcal{S}$ by the clause

$$\mathcal{S} \vDash K_a \varphi \text{ iff for all } w \in \mathcal{S}, \varphi \in w.$$

I.e., the agent knows $\varphi$ if its negation is simply not on the radar: It is not considered possible at all. Allowing for a generalization of sphere models where not *all* possible worlds need be present in $\mathcal{S}$, the public announcement of $\varphi$ in $\mathcal{S}$ is the updated sphere model $\mathcal{S}_{|\varphi}$ where all worlds containing $\neg\varphi$ have been deleted, but the sphere system among the remaining worlds stay the same. No matter whether the original model satisfies $K_a \varphi$ or not, then the updated model will make the formula true.

Public announcements serve to illustrate one of two main foci in dynamic epistemic logic, namely *model change*. Semantically, it is the hall-mark topic of study: Given a static model, operations are defined to systematically change it. Such updates—or, more generally, such *model transformations*—are then defined for a variety of purposes. The most well-known examples are of communicative acts, broadly construed, in epistemic and doxastic multi-agent systems. Model change is a main focus in this thesis.

The second focus of dynamic epistemic logic concerns the *logics* of model change. The *modus operandi* involves adding so-called *dynamic operators* to a static language. Back to public announcement as an illustration: Construct a dynamic language from the static prototypical epistemic language by allowing formulas of the form

$$[!\varphi]\psi$$

whenever $\varphi$ and $\psi$ are dynamic formulas. The formula $[!\varphi]\psi$ reads 'It is *now* the case that if it is announced that $\varphi$, then *after* the announcement, it will be the case that $\psi$.' Over Grove sphere systems, the formula could be evaluated by the clause[18]

$$\mathcal{S} \vDash [!\varphi]\psi \text{ iff } \mathcal{S}' \vDash \psi.$$

The *logic* of dynamic epistemic logic is then the logic involving such dynamic operators. In the present thesis, dynamic modalities are used in Papers I and II, while the standard method for obtaining logics proper is touched on in Paper VI. The method, also introduced by Plaza in [133], is by translation: Essentially, it is shown that every formula with a dynamic modality may be translated to an equivalent static formula. In a slogan, talk about dynamics may be *reduced* to talk about statics. The method involves adding so-called *reduction laws* as axioms to a suitable static logic, a method returned to in Section 8 on Paper VI.

Since the publication of Plaza's [133], dynamic epistemic logic has grown into a small research field. Early contributions are the operations for truthful public announcements, *untruthful* announcements and *non-public* announcements, i.e., announcements to subgroups of agents, defined, independently of Plaza cf. [54], by Gerbrandy [65] and Gerbrandy & Groeneveld [66] in the late 1990's, though with heavy influences from the general 'dynamic turn' in logic, for more on which [24, 54] are sources.

The dominant approach to model transformation representation, however, came with Baltag, Moss and Solecki's 1998 paper *The Logic of Public Announcements, Common Knowledge, and Private Suspicions* [13]. The authors introduced a general methodology for constructing complex transformations. The key insight, both technically and intuitively, was that information events may be represented as *models*. An example of such a model, in general called *action models* or *event models*, is illustrated in Figure 2. The effect of the action model cannot be illustrated in the hitherto introduced Grove sphere models: For that, the situation must be represented by a Kripke model. The details involve taking the product of a Kripke model and an action model using *product update*, in loose terms a type of graph theoretic "multiplication", exemplified in the next section.

Action models and product update jointly constitute the main type of model transformations used in this thesis. Exceptions are Paper I, which uses a similar construct, but tailored

---

[18]The clause does not cover the case where $\mathcal{S}$ contains only $\neg\varphi$ worlds, in which case $[\varphi!]\psi$ is standardly assumed satisfied.
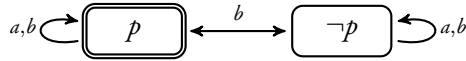
Figure 2: An example of an action model with two agents, *a* and *b*. The boxes represent *information events*: In the left, it is truthfully announced that *p*, in the right that ¬*p*. The left event is what in fact happens, indicated by the double line. The labeled arrows represent what the agents *cannot distinguish*: Agent *b* cannot distinguish whether *p* or ¬*p* is being announced, but *a* can. Agent *b* will therefore learn nothing about *p*, but *a* will learn that *p* is the case. Agent *b* will learn something, though, namely that *a* learn *whether p or not*.

for beliefs (the *action-priority update* of Baltag and Smets [17]), and Paper vi, which abstracts away from the details of the transformations. The versions of action models and product update used in Papers ii-v are not precisely those of Baltag, Moss and Solecki, but generalizations thereof. In particular, the papers use *multi-pointed* action models [20] with *postconditions* [27, 53], allowing for complex program constructions and factual change. The reason for this focus is the generality and wide applicability of action models: The framework is not perfect for every occasion, but it is simple, rich, and of broad interest, as witnessed by the vast literature that employ, investigate or build on it. In seeking to apply topological and dynamical systems concepts to logical dynamics in the dynamic epistemic logic tradition, it is a natural starting point.

However, action models and product update are not a natural stopping point for what falls in the field of dynamic epistemic logic. Quite to the contrary, the literature sports a wide variety of model transformation types, designed for belief revision [7, 16, 17, 18, 22, 32, 48, 54, 67], as general alternatives or supplements to action models [27, 35, 53, 56, 71, 102], for hybrid [81] and probabilistic extensions [29, 101], for topological variations [14, 51, 128], and for dynamics of awareness [151], evidence [11, 31], preferences [30, 119], and diffusion in social networks [12, 39, 40, 41, 120, 136, 144], to illustrate.

## 2    Elements of Dynamic Epistemic Logic

For all its variety, there is one perspective on dynamics shared by most works in dynamic epistemic logic: A dynamics is a sequence of self-contained models $m_1, m_2, m_3, \ldots$ where all but the initial is obtain from the former by some transformation. This perspective stands in contrast to models representing dynamics internally, from the outset offering a full, unfolded view of time. This is the contrast between the *local* and *global*—or *Grand Stage*—views on dynamics, in the terms of van Benthem [24]. Grand Stage models are touched on in Section 5 on Paper iii.

The local view on dynamics is not unique to dynamic epistemic logic. The the perspective is, for example, shared by the works on belief revision from the previous section: An agent's belief state is a self-contained model, a snap-shot of the agent's information here-and-now, independent of past or future changes. When this belief state is changed, a new, again

self-contained belief state is the result. Equally, the local view is shared by discrete-time dynamical systems, with one restriction: In the case of a dynamical system, the sequence is the result of iterating a map on the initial model.

In the classic and most common dynamic epistemic logical setting—that which this thesis is on—the self-contained snap-shot models are so-called *pointed Kripke models*.[19] A pointed Kripke model is a simple beast. Described graphically, it consists of some *dots*, each marked with zero or more *colors*. One of these dots is bigger than the others, and is called the *point*. Between some of the dots are drawn *arrows*. Here is an example, using the colors red and green:
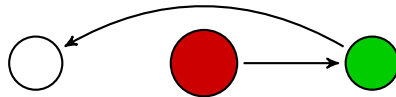


In this drawing, the left dot has no color, the center dot is red, and the right dot is green. There could be a dot with two colors, but there is not.

A transformation of a pointed Kripke model is then just changing the model to another one. Typically, transformations are given in some systematic, rule-based manner. An example could be "If a dot points at a green dot, then it must point at itself, too." This rule would transform the above to this model:
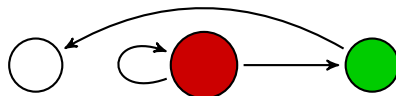


That's it—a dynamics in the spirit of dynamic epistemic logic. The sequences of interest typically have a deeper interpretation, a longer time-span and well-motivated rules for transformations, but in the abstract, the idea is the same.

One possible interpretation of a pointed Kripke model could be as a network of train stations with the point marking the commuter's current position, the colors the colors of trashcans, the arrows the running lines. Transformations could then involve changing the commuter's position, painting trashcans, adding or removing stations, closing redundant lines or opening new ones—maybe even some running both directions.

The standard interpretation in dynamic epistemic logic, however, has little to do with trains. In the standard interpretation, each pointed Kripke model represents a state of affairs and one or more agents' information and higher-order information about this state. Under an informational interpretation, a pointed Kripke model is consequently often called an *epistemic state*. A sequence of epistemic states then represents an information dynamics,

---

[19] *See* e.g. [34, 72] for textbook references

where the state of affairs may change, the agents' information about the state of affairs may change, or the agents' information about others' information may change—in any combination.

Such sequences constitute a main object of study in the field of dynamic epistemic logic, and has a number of applications, mainly in theoretical philosophy and theoretical computer science. In philosophy, one may be interested in which types of knowledge or belief are robust when agents learn new information; in computer science, one may be interested in whether some communication protocol is secure even when messages are overheard.

To model a dynamics using dynamic epistemic logic, there are three essential elements to consider: *i*) what aspects to include, both in the state of affairs, but also in terms of the agents' information, *ii*) how to represent these aspects, and *iii*) how to update the representations to obtain dynamics.

For *i*), what to include depends, of course, on application: In one case, certain features may be abstracted away that may be considered essential in another. To model an agent undergoing the information dynamics of the example in the introduction, it might be deemed relevant to include propositions concerning the content of this thesis, but less so to include propositions about the position of the moon. Likewise, knowledge and beliefs may seem germane, while the agents' ethical considerations might be safe to omit.

For a simple exposition which still illustrates the key elements of dynamics as modeled using dynamic epistemic logic, take the following view on the initial example: Two agents are sitting in a room, one them holding this thesis, closed and unread. Call this agent *a*, the other *b*. Of interest is only whether the first section contains an example or not, together with *a* and *b*'s (lack of) *knowledge* about this fact.[20]

With the relevant features selected, describing them using a formal language has the advantage of conciseness. First, use *p* to denote the proposition 'The first section contains an example'. Then *p* corresponds to a color in the illustration above, and represents a proposition considered "atomic" in the sense that no attempt to represent its internal linguistic structure is made: The *p* in the language cannot be further broken down into subformulas. For knowledge, let $K_i\varphi$ read that agent *i* knows that $\varphi$, with *i* either *a* or *b*. Include negation ($\neg$) and conjunction ($\wedge$), so the formal language of the model is given by the Backus-Naur form

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_i\varphi$$

read as described on page 3.

For *ii*), the representation of the relevant aspects, choose pointed Kripke models, to be interpreted as epistemic states. The standard interpretation is as follows, starting with the

---

[20]To include beliefs, too, would be more natural, but the constructions are slightly more involved, cf. Paper 1.

dots, interpreted as states of affairs. Each such state is colored by a selection of our chosen atomic propositions: When an atomic proposition is present at a state, it is considered *true* at that state. Otherwise, it's false, indicated here by marking the state with its negation. In the example with $p$ the only atomic propositions of interest, there are only two possible states of affairs:



Here, the double line indicates the point, representing the actual state of affairs. Given the introduction, it seems the most reasonable representation of the real world. The second state represents an *alternative* state of affairs, one where there is no example.

Last, the arrows. Enter the representation of information and knowledge. When two states are connected by an arrow labeled by an agent name, this represents that the agent cannot tell the two states apart—to the agent, the dots are *informationally indistinguishable*. Informational indistinguishability is the central epistemic notion in the standard interpretation. Per the semantic definition of truth for knowledge formulas given below, it is on this notion the standard interpretation's understanding of knowledge rests.

To exemplify, consider the following off-scale black-and-white illustration of one of the dots from the colored Kripke model above:



Is it the white, the red, or the green dot? We can eliminate white: The brightness is off. But whether it is red or green, we lack color and size information to determine. Stated differently, if it does illustrate the green dot, then *for all we know*, it might as well be the red dot. *Vice versa*, if it is the red dot, then *given our information*, it might as well be the green. Conversely, if two states are not connected, then it means the agent *does* possess the information to distinguish one from the other: If asked whether the dot above is white, we have sufficient information to answer *no*.

When states are indistinguishable in this informational sense, they are called *epistemic alternatives* to each other. Arrows, then, connect epistemic alternatives.

Returning to the two agent example, to determine which states are epistemic alternatives to each other, we should ask whether the agents possess any information that allow them to distinguish the states. If they do not, arrows are drawn, not otherwise. With the thesis held by $a$, closed an unread, neither agent has any information about whether the introduction contains an example or not. I.e., neither agent can distinguish the left from the right state, nor *vice versa*:
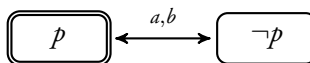
A final question and the epistemic state is complete: Are there any of the states the agents can distinguish *from themselves*? It's not a trick question—the answer is just a trivial *no*. No information will ever allow them to distinguish something from itself, seeing as there are no differences between "the two". Hence, the epistemic state:
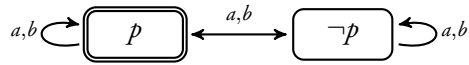


This epistemic state captures the situation where the agents have no information about the state of affairs. In the literature, such an epistemic state, where all the possible states are connected by bi-directional arrows, are sometimes referred to as one of *blissful ignorance*.

The ignorance in the epistemic state may be expressed using the formal language, in particular the knowledge operators $K_a$ and $K_b$. Their interpretation over an epistemic state provides the link between the standard interpretation's view on knowledge and the notion of indistinguishability: A state in an epistemic state is makes the formula $K_i\varphi$ true if, and only if, the formula $\varphi$ is true in all of agent $i$'s epistemic alternative to that state. Hence, for an agent to know $\varphi$, it must have information that eliminates *all* the alternatives that falsify $\varphi$.

The agents' ignorance is exposed by checking which knowledge formulas are true at the actual state of the epistemic state. The formula $K_a p$, for example, is false: Agent $a$ has an epistemic alternative where $\neg p$ is true. Likewise $K_a \neg p$ is false because of the $p$ alternative— the actual state itself. Hence, both $\neg K_a p$ and $\neg K_a \neg p$ are true: Jointly, they state that, in the epistemic state, agent $a$ knows nothing about the example in the first section. The same goes for agent $b$. No information, no knowledge, only blissful ignorance.

The denotation 'blissful ignorance' comes with a caveat: It refers only to ignorance with respect to the *atomic propositions*. Agents $a$ and $b$ possess quite a lot of information about each others' information: For example, the actual state satisfies $K_b(\neg K_a p \wedge \neg K_a \neg p)$—that agent $b$ knows that agent $a$ has no information. To check this requires checking epistemic alternatives for both agents, but the modal logical semantics ensures that the formula has a definitive truth value. Likewise, any formula of the language is evaluated as either true or false. An epistemic state thus represents all the agents' *higher-order* information—their information about other's information (about others' information, etc.)—in quite a small construct.

There is much more to be said about epistemic states, both formally and on interpretation, than it is feasible to recant here. In passing, it may be mentioned that the arrows represent binary relations, under the standard interpretation taken to constitute *equivalence relation*s, resulting in the logic of the knowledge operators being S5, mentioned in Section 2, and that operators for the information of groups—like *common* and *distributed* knowledge—are often included in the language. In-depth introductions may be found in e.g. [24, 54, 58, 88].

For *iii*), how to update an epistemic state with new information to obtain dynamics, the example initiated first requires a continuation. Assume, for narrative purposes, that agent *a* opens the thesis and reads the first sentence of the introduction. Assume further that agent *b* witnesses this, but does not know what agent *a* reads. Finally, assume that both agents take whatever is read as absolutely guaranteed to be irrevocably *true* (and that all this is common knowledge).

This continuation of the example encapsulates more than the agents' individually receiving a new piece of information. Rather, the continuation is a complex informational event, where one agent receives new information directly, while the other's information changes as a consequence. The ability to represent events of this complexity is one of the strong suits of dynamic epistemic logic, and in particular of the action model approach touched in Section 2. As mentioned there, a key insight of Baltag, Moss and Solecki's [13] is that informational events may be represented as semantic structures, so-called pointed action models.

The standard interpretation of a pointed action model is very close to that of an epistemic state, but where the individual 'dots' do not represent states, but instead *events*, with an event colored not with the atomic propositions it makes true, but with *the information that the event truthfully conveys*. The point, then, represents the event that actually occurs. The arrows between events still represent indistinguishability.

One possible modeling of the continuation is the following action model:
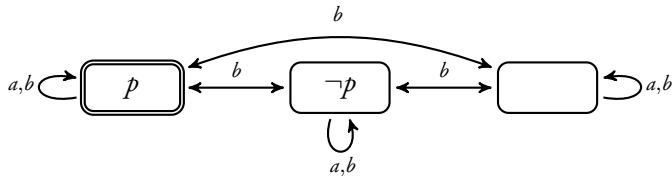


This pointed action model represents that, in fact, the information that *p* is being truthfully conveyed (somewhat unnaturally, to both agents simultaneously), but that agent *b* cannot tell whether the information truthfully conveyed is *p*, ¬*p* or no information at all. Agent *a*, however, can tell all three events apart (but cannot distinguish any event from itself).

To obtain a new epistemic state from the old and the action model, the two are "multiplied". The process takes two steps.

An intuition for the first step is that each event in the action model works like an informational "filter" which allows some states through, but not others. Take the left most event, for example. It *truthfully* conveys the information that *p*. Hence it *rules out* the state where ¬*p* is the case: This state does not satisfy the *precondition* that *p* is the case, and so does not pass through that filter. The ¬*p* state is however compatible with both the middle and right most events: Nothing in the information conveyed there rules out that ¬*p*.

This "filtering" through preconditions determines the set of states in the new epistemic state. In this example, there will be four states:

$p$     $\neg p$     $p$

$\neg p$
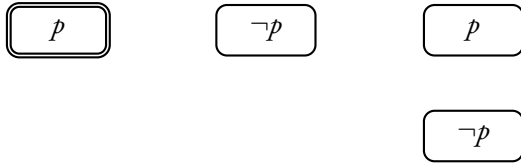
The left most state is the *child* of the original $p$ state and the $p$ event: It is the actual state because is is the result of what was actually the case and what actually happened; the middle is the child of the $\neg p$ state and event; and the two right most states are the children of the $p$ and $\neg p$ states with the null information event. Though some a identical in the propositions they evaluate, none are redundant.

In the second step, the arrows are calculated. The device here is that two children are indistinguishable if, and only if, all of their parents are indistinguishable. For an example, think of a person of whom you do not know whether they prefer $A$ or $B$—you cannot distinguish one case from the other. In answering your query, they mumble to the degree that you miss the answer: You cannot tell whether they answered $A$ or $B$. You are no wiser, and still do not know their preferences. Had you been able to distinguish either set of parents—either the preferences or the answers—there would have been no uncertainty.

In a more sterile formulation, the arrows are calculated as follows: If states s1 is the child of state s and event 1, and t2 the child of state t and event 2, then an $i$-arrow must be drawn from s1 to t2 if, and only if, there is an $i$-arrow from state s to state t and an $i$-arrow from event 1 to event 2. For fully detailed introductions, see e.g. [15, 54, 123].

Applying the second step complete the construction of the new epistemic state, which becomes a bit of a mouthful. Perhaps it helps to notice that for agent $b$ *all* the states are linked together, while for agent $a$, the epistemic state can be chopped into three disconnected components (left, middle and right).
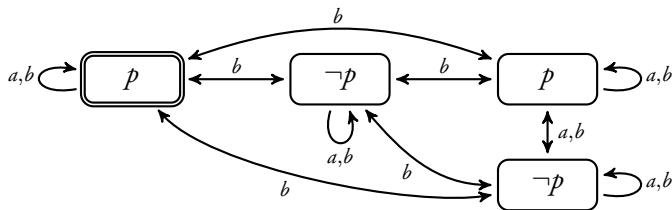
In this updated epistemic state, both agents have new information, but it is information of very different kinds. Agent $a$ has learned something, namely that $p$: The actual state now satisfies $K_a p$. Agent $b$, however, has not so much gained knowledge as gained confusion: The agent now considers *more* things possible than previously—for example that agent $a$

may know that $p$. Specifically, where agent $b$ before knew that $a$ had no information—as $K_b(\neg K_a p \wedge \neg K_a \neg p)$ was true—agent $b$ now knows that they new *do not* know this: $K_b \neg K_b(\neg K_a p \wedge \neg K_a \neg p)$. Whether to call this new information for agent $b$ may be contested, but that it is a change in information is certain.

To finalize the example, consider a last informational event: Let agent $a$ tell $b$—without mumbling—of the newfound knowledge that $p$. This event may be represented by the pointed action model

$$a,b \circlearrowleft \boxed{\boxed{K_a p}}$$
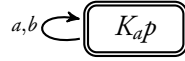
which in effect is equivalent to the truthful public announcement of $K_a p$ in the style of Plaza [133], mentioned in Section 2.

This time around, the calculation of the product is quite a lot simpler: Only the actual state of the epistemic state satisfies the precondition of the single event in the action model, so the resulting epistemic state has only one state, indistinguishable from itself:

$$a,b \circlearrowleft \boxed{\boxed{p}}$$
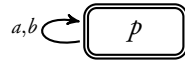
In this final epistemic state, both agents know $p$, know that the other knows $p$, and so forth, for all levels of higher-order information.

The now modeled example illustrates one of two aspects of what dynamic epistemic logic does well, namely the semantic aspect of constructing information dynamics. The second aspect concerns how to obtain *logics* for such dynamics, a topic postponed to Section 42 on Paper VI. The example also implicitly illustrates an aspect of information dynamics which the standard dynamic epistemic logic framework does not handle, as the framework only represents a part of the story of how the dynamics come about. Specifically, the framework includes the machinery to formally represent epistemic states, informational events, and how to update one of the former with one of the latter. The framework, however, is silent on what informational event a given epistemic state should be updated with. This is perfectly fine for epistemic analyses where the interest is on how a given epistemic state will evolve under some sequence of informational events. As with Bayesianism, it is a full-fledged theory of rational epistemic agency. Bayesianism, however, is easily coupled with *decision theoretical* frameworks, such as expected utility maximization. For dynamic epistemic logic, no such standard decision framework exists. When a dynamics partially develops as the result of agents' decisions, dynamic epistemic logic thus requires an augmentation. Such an augmentation is a topic in Paper I.

# 3 Paper I: Decision Making in the Bystander Effect

The path this thesis has taken was not foreseen: The papers included are each the product of ideas that sprung from the writing of those before. They progress from attempting to solve a practical modeling task in Paper I to abstract characterization theorems in Paper IV.

Paper I is an attempt to use formal information dynamics—like those of the previous section—to model an empirical phenomenon known as *the bystander effect*. The bystander effect is a well-documented phenomenon in social psychology, where the term covers the tendency that witnesses are less prone to offer assistance when others are present.

In the social psychological literature (see e.g. [109, 124]) the phenomenon is explained by a composite of three features: First, that the presence of others may cause us to overlook an accident; second, that the presence of others may offer misleading information causing us to mis-classify the accident as inconsequential; and third, that the presence of others may cause us to not take responsibility for intervening. Paper I focuses on the second of the three sub-explanations.

The second sub-explanation consists of an information dynamics in which the observation of an event, the observation of others, and informed choices are key components. In a nutshell, the story goes as follows:

A group of agents all witness an accident. Privately, they conclude that there is need of help, but they are uncertain. To not rush to unwanted aid, they choose to observe their co-witnesses: Do they think help is needed? As each agent observes the others covertly, their act of observation is observed by the others as inaction—they are not rushing to help. From this observation of others, each agent concludes that they other witnesses think that help is unnecessary: Else they would be rushing to help.

This belief state, where everybody believes the same one thing, while believing that the others believe the opposite, is in the literature called a state of *pluralistic ignorance*.

In the state of pluralistic ignorance, the agents re-evaluate the situation at hand. Based on their new information—that no-one else believes that help is needed—they revise their beliefs: They treat the perceived inaction of the others as *social proof* that help is unnecessary. From this, each agent privately concludes that help indeed *is* unnecessary: They just misinterpreted the original situation. In the story's final act, each agent therefore chooses to not offer aid.

This pluralistic ignorance explanation of the bystander effect is contrasted by another, shorter and for a time also popular explanation: *People just don't care about each other*. This apathy explanation also explain the unresponsiveness of witnesses, but it does not explain why people intervene when alone. It is therefore a poor explanation, but returned to below.

Several reasons made it compelling to attempt a modeling of the bystander effect.

First—very generally—there is the interest in having a formal model in the first place, cf. e.g. the recent thesis by Dominik Klein [100]: A formal representation of a phenomenon— or its various explanations, as in this case—typically offers a *clarification* of the involved concepts by a rigid, formal representation. This in turn allows the *verification* of said explanation—a test of whether involves explanatory gaps more easily missed in the natural language presentation. Finally, a formalization allows *exploration* of the phenomenon, e.g. by the varying of parameters to find the boundaries of a particular outcome—like the unresponsiveness of witnesses. In Paper i, these three aspects are all present: The dynamic transfer of information between agents is clarified compared to the informal social psychological model, elements required to make the explanation complete are identified and boundary parameters are estimated (roughly: the data is scarce).

Second, modeling the bystander effect presented an opportunity to attempt a modeling of an interesting empirical phenomenon. Based on a paper co-authored with P. G. Hansen and V. F. Hendricks [83], the hope was, through modeling a sequence of information dynamical phenomena, to obtain overarching structural insights into why reasonable individuals as a group gets it *wrong*. Given the freedoms that dynamic epistemic logic offers in terms of information updates, modeling the bystander effect using specifically this tool seemed a good way to put the tool to the test.

Third, the pluralistic ignorance explanation in it self is an interesting information dynamics, posing several open questions concerning modeling choices. The explanation involves informational social influence by revision based on aggregated social proof obtained through action interpretation, all elements that offered interesting challenges. In addition, the explanation involves agents' *decisions*, an aspect of dynamic phenomena not standardly treated in dynamic epistemic logic, cf. the above.

In particular the aspect of agent decisions presented a conundrum: If one chose to impart a game theoretic payoff structure [127], it would be over the terminal outcomes of the bystander effect story. But for the agents to reason about these terminal outcomes, the model would have to included them from the outset. Such a model would then no longer be just the representation of a current epistemic state: It would include all the possible choices, the full temporal structure—the Grand Stage. But then the model would no longer be cast in dynamic epistemic logic, but something else entirely, somewhat defeating the second point of motivation. Hence, I tried to come up with an mechanism for invoking agent choices suited to dynamic epistemic logic. The resulting "transition rules" are the thesis' first step towards identifying models with maps.

Paper i, then, is in summary as follows: After having presented the dynamics of the bystander effect in a step-by-step fashion, it includes an introduction to a dynamic epistemic logic framework, namely the *plausibility models* and *plausibility action models* of Baltag and

Smets [17], augmented with postconditions. It is shown how the occurrence of an accident may be modeled using a plausibility action model capturing that each agent perceives the event as an accident, while being in the blind about other agents' perception.

At this stage, agents' decisions enter the picture. Agents are modeled as having three possible actions: To intervene, to evade, or to observe. Each action is modeled as a pointed action model where the acting agent know her action, while others cannot distinguish between evasion and observation, but find the former more plausible. As the agents are assumed to act simultaneously, the paper defines an operation for composing individual actions into a joint (pointed) action model. The accident state is then updated with one of these composite models—which one depends on the agents' individual choices.

To formally model how the agents decide their actions, the paper introduces *transition rules*. Informally, a transition rule is a requirement to the next action of an agent. It is of the form "*If $\varphi$ is the case, make a choice that makes Y the case.*" The choice of the agent in a situation making $\varphi$ the case would then be a solution to this "equation", with the available solution candidates the actions above.

Using transition rules, the paper defines four different types of agents. One of them is apathetic and just doesn't care about others. That agent type's rule colloquially reads "*No matter what, make a choice that makes you not be in the middle of this.*" This apathetic agent type always chooses to evade. The other three types are first responders, observers, observers that are affected by social proof. The latter represents the agents hypothesized by the pluralistic ignorance explanation of the bystander effect.

The individual actions, the transition rules and the composition operation jointly allow for the definitions of variety of models for the bystander effect, depending on the composition of types in the population. The paper investigates only the four pure populations, the models of which are compared to data from classic studies in social psychology. It is concluded that only the observer influenced by social proof fits the bill.

For the thesis at large, the key point of Paper 1, at best implicit in the paper itself, is methodological: In adding transition rules as a decision making framework on top of the standard dynamic epistemic logic, the models of the various explanations become *mathematically self-contained*: Given an epistemic state, the augmented framework fully specifies how the next epistemic state is to be calculated from the former.[21] To underline the point, then irrespective of whether decision making is at stake or not, such self-containedness is *not* a default in dynamic epistemic logic modeling practices. As it is a prerequisite for a rigid comparison of models that they explicit in their assumptions and constructions on all points, it is in such cases a *desideratum* that the models are fully—rather than only partially—formal. In such

---

[21]In retrospect, this self-containedness essentially amounts to identifying models with maps: From an epistemic state $m$, a set of transition rules $t$ dictate a joint action model $t(m)$ subsequently applied using a product, $\otimes$, all in all specifying the next epistemic state $f(m) = m \otimes t(m)$.

cases, the standard dynamic epistemic logic framework requires additional functionality. Whether transition rules specifically are a good choice for obtaining this functionality is a topic in Paper II.

# 4  Paper II: On the Plurality of Choice Mechanisms

In evaluating explanations for the bystander effect, Paper I follows the lead of the social psychological literature: A small collection of models are compared to the empirical findings and the best fitting is concluded the best explanation of the phenomenon. Best, that is, among the small collection of conceived of models. To this small collection of models, there might, however, exists a large collection of alternatives. One thing is that the pluralistic ignorance model contains a free parameter that could possibly be estimated for a better fit, but there may also be completely different dynamic epistemic logic models that—in general—would do a better job of explaining the bystander effect. Hence, having gotten some distance to Paper I, the following question arose: Given that the paper already used non-standard functionality, what tricks could these hypothetical, alternative models possibly invoke? Asking the same differently, when seeking the best model of the bystander effect, then the best *among what*?

To answer, another question presented itself: Mathematically, set-theoretically, what *are* these thing referred to as "the models"? Abstract away from the concrete form of the transition rules and look at their function: They pick actions. Specifically, given a pointed Kripke model as input, a transition rule picks an individual action which, when composed with the individual actions picked by other transition rules on the same input, produce an action model. Zooming out a bit, a set of transition rules thus simply specify an action model, given a pointed Kripke model as input. In other words, a set of transition rules induce a mapping from a set of pointed Kripke models to a set of action models.

As an action model and a product jointly define mappings taking pointed Kripke models to pointed Kripke models, the big picture structure of the type of models is that they are maps from a set of pointed Kripke models to the set of maps on that set of Kripke models. With $Q$ the set of pointed Kripke models in question and $Q^Q$ the set of maps from $Q$ to $Q$, "the models" are thus of the form $f \colon Q \longrightarrow Q^Q$.

However, in answering the *among what?* question, not all maps $f$ of this form seemed relevant. First, the set of maps $Q^Q$ seemed too unrestricted: Reasonably, it seemed, the subset $A \subseteq Q^Q$ of maps representable using action models should suffice. Second, not every map $f \colon Q \longrightarrow A$ would be of interest either: At least, the map should be reasonably representable.

At this point, a lesson from a Grand Stage model type—*interpreted systems*—enters the

picture. In their seminal *Reasoning about Knowledge*, Fagin, Halpern, Moses and Vardi dedicate a number of chapters to the design of protocols for interpreted systems, a widely used Grand Stage framework for multi-agent systems. In particular, they spend a chapter on *knowledge-based programs*. In essence, a knowledge-based program is a set of instructions for an agent: Each instruction takes the form "*If $\varphi$, then do $a$*", where $a$ is some individual action. In the context of dynamic epistemic logic, it is natural to think of $a$ as an individual action model. As with transition rules, a set of knowledge-based programs may thus be used to define a map $f : Q \longrightarrow A$. Hence, they offer an alternative way of representing models.

This realization caused doubt: The approach using transition rules is somewhat cumbersome compared to knowledge-based programs, but it also seemed to required that the agents solve a problem to decide on an action—akin to solving a utility maximization problem[22] in game theory—instead of having an action just dictated to them whenever a particular condition arose. This doubt motivated Paper II: It is more pleasant to check whether there is a difference than to toss and turn under uncertainty.

The particular question asked in Paper II is inspired by dynamical systems theory: With a map $f : Q \longrightarrow Q^Q$, the map given by $g(x) = f(x)(x)$ maps $Q$ to $Q$ and is hence an instantiation of the broadest understanding of a *discrete-time dynamical system*: A set together with a map acting on it (cf. e.g. [57]). Paper II then asks which types of *orbits*, i.e. sequences of pointed Kripke models from the set $Q$, that may be produced by different types of update mechanisms.

The three update mechanisms are transition rules (referred to in the paper as *problems*), knowledge-based programs and and *deterministic multi-pointed* action models. The latter shares the functionality of the other two: Given a pointed Kripke model as input, it supplies an action model as output with which the input model is updated.

The main result shows that knowledge-based programs and multi-pointed action models are equivalent in the orbits produceable, that these are a subset of the orbits produceable by problems,[23] and that under two finiteness conditions, the three are equivalent. The two finiteness conditions are quite liberal, and it is on a conceptual note concluded that the orbits produceable are adequate for modeling empirical phenomena finite in nature while befitting of the modeling style of dynamic epistemic logic.

For the thesis at large, the main result contributes that there exists natural ways of fully formalizing dynamic epistemic logical models as maps, and that this approach is sufficient for most empirical purposes.

On a different conceptual note, Paper II initially makes an argument akin to that concluding Section 3 on Paper I: The standard dynamic epistemic logic framework leaves something

---

[22]Or other choice principle problem: See e.g. Paolo Galeazzi's [63].

[23]Whether it is a strict subset is not checked.

to be desired when seeking full formalization: Its models are not sufficiently described to be identified with any one type of mathematical object. The dynamical systems framework is suggested as a way to achieve precision, settling the *among what?* question. The paper shortly compares the approach with a different framework playing a similar role: The *extensional protocols* framework of van Benthem, Gerbrandy, Hoshi and Pacuit's [28]. Paper III contributes with a deeper comparison.

## 5  Paper III: Intensional and Extensional Protocols

When supplying procedures for action—when specifying protocols—to specification forms stand out. One dictates a sequence of action, not unlike a cooking recipe: First crack the eggs into a bowl. Beat them, and then stir in the vanilla. The second is conditional, not unlike a survival guide: In case of hypothermia, raise the body temperature; in case of hyperthermia, lower the body temperature.

In the terminology of Parikh and Ramanujam [130], the two instruction forms are referred to as *extensional* and *intensional* protocols, respectively.

The transition rules, the knowledge-based programs and the multi-pointed action models of Paper II are all—when applied iteratively as per the dynamical systems approach—intensional protocols.

Paper III compares the intensional protocols of dynamic epistemic logical dynamical systems with the extensional protocols for dynamic epistemic logic of van Benthem, Gerbrandy, Hoshi and Pacuit's [28]. The comparison is mediated by the models of *epistemic temporal logic*, following the methodology of [28].

The semantic structures of epistemic temporal logic are Grand Stage models—one is illustrated in Figure 3. Grand Stage models were hinted at in the consideration surrounding Paper I: Instead of a step-by-step modeling in the style of dynamic epistemic logic, the model of the bystander effect could also have been constructed using a single, all-including model, representing both information, time, interplay of the agents' possible choices, and the resulting outcomes.

Such epistemic temporal models are based on *histories*: Branching strings of the states of Kripke models, linked by a relation representing the passing of time. Each time step, in turn, represents actions taken by agents—this is the temporal dimension. The epistemic dimension is as in epistemic states: States of the world are linked to the effect that agents may be uncertain about the current state of the world and may in addition have detailed knowledge, beliefs or expectations about what past actions were taken, what futures are possible, or what futures they or others can force about.
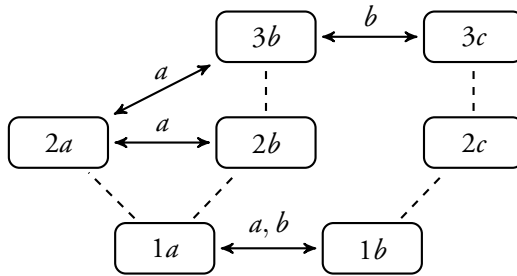
**Figure 3:** A simple epistemic temporal model for two agents, *a* and *b*, with states named by time and position. Arrows represent epistemic indistinguishability—reflexive and transitive links are implicit. The dashed line represents the passing of time. In the model, neither agent knows which state is the case at time 1, but at state 2*c*, they are both certain. Agent *a* does not always know the time, cf. the links between 2*a* and 3*b*. Agent *b* does not have perfect recall: At time 3, it has forgotten that it knew whether it was at a *b* or a *c* state at time 2.

Epistemic temporal models are common-place in computer science and game theory, exemplified by interpreted systems or extensive-form games, where they allow the modeling of agents that undertake long-term reasoning. They thus stand in contrast with the approach of dynamic epistemic logic, where only the current time is represented in any one pointed Kripke model.

As both time and information are portrayed in epistemic temporal models, interplay between these may be represented. For example, two dots different numbers of times-steps away from some time *zero* may be connected by an arrow: Then the agent doesn't know what time it is. Another example is memory: If a history is epistemically ruled out at one time-point—i.e., if there are no arrows connecting that time-point to any of the history's dots—then it will remain ruled out at later time-points. This property is called *perfect recall* and epitomizes the opposite of Douglas Quayle's problems in Philip K. Dick's *We Can Remember It for You Wholesale*, adapted to film as *Total Recall*, twice.

Properties such as these may be used as design requirements for rational agents. Agents may be required to, e.g., have perfect recall, to always know the time, to learn systematically from events, among others.

As dynamic epistemic logic models do not have explicit temporal structure, it is not readily possible to determine whether their agents satisfy such temporal rationality requirements. Enter van Benthem, Gerbrandy, Hoshi and Pacuit: In the paper *Merging Framework for Interaction* [28], they show how dynamic epistemic logic models may be used to build epistemic temporal models. Moreover, as the resulting all-in-one models always exhibit certain properties—perfect recall being one of them—they obtain a characterization of the types of agents implicitly assumed by the dynamic epistemic logical approach.

In obtaining the characterization, the authors must show that *any* dynamic epistemic logic model will result in an epistemic temporal model that possesses the properties in question.

For this, it is a prerequisite to fix what is meant by a dynamic epistemic logic model—as in Paper II.

The framework used in [28] is based on a formal representation of extensional protocols. Simplified, an extensional protocol is a finite *string of actions*: $a, a, b, a, ..., c, b, d$. Such a string together with an initial pointed Kripke model then constitute a dynamics: For the string here, the initial pointed Kripke model is first updated by $a$, then by $a$ again, then by $b$, and so forth. The result is a sequence of pointed Kripke models, just as in the example of Section 2 and as with the dynamical systems approach.

To build an epistemic temporal model from an extensional protocols model, then, one simply stacks the pointed Kripke models in the order they are produced, connecting their dots with temporal links whenever one is results from another.

The dynamical systems approach to building epistemic temporal models is in many ways the same, but in an essential way different. It is the same in starting from an initial model which is sequentially updated by actions, leading to a string of stackable Kripke models, but it is different in how the actions that update the Kripke models are specified; in how—and when—they are selected. In the extensional protocol approach, the full string of actions is pre-specified by the modeler, given only the initial Kripke model. In the dynamical systems approach, the modeler instead pre-specifies a map, which at the initial and subsequent Kripke models pick an action to be executed.

Both approaches supply instructions for action, for detailing protocols, but one is extensional, the other is intensional. Like cooking recipes and survival instructions, both are useful, but serve different purposes.

In Paper III, Hanna van Lee, Suzanne van Wijk and I compare these two approaches to protocols for dynamic epistemic logic. To do this, we copy the methodology of van Benthem, Gerbrandy, Hoshi and Pacuit, and characterize which types of epistemic temporal models that are constructable when taking the dynamical systems approach to model building. Given the results in Paper II, we choose, for simplicity, to use multi-pointed action models for the update mechanism. The two frameworks are then compared by contrasting the properties—like perfect recall—they induce in the epistemic temporal models they produce.

Following a presentation of the intensional and extensional protocol frameworks, these are initially compared in a simple modeling task, *the muddy children* [58]. Modeling this puzzle, it is concluded that the intensional protocols approach has an advantages in being finitely representable.

Ensuingly, the epistemic temporal structures are introduced, along with 8 structural properties. Having defined how to generate a big model from an intensional protocol, a first result

is shown: 7 of the 8 properties are necessary consequences, but the 8th is not. Conversely, it is shown that if a epistemic temporal model possesses all 8 properties, then there exists an intensional protocol that will generate it. These two results *almost* constitute a characterization of the epistemic temporal models that can be build using intensional protocols, but not quite. When attention is restricted to situations where both the epistemic temporal model and the intensional protocol satisfy two additional finiteness criteria, then a proper characterization is obtained: The eight properties are both necessary and sufficient. These technical results are the main results of the paper.

Subsequently, the paper focuses on a difference between the two protocol frameworks not mentioned above: Originating from mappings, dynamical systems are *deterministic*: Given an input, they produce one output only. Extensional protocols are more general, and allow for *non-deterministic protocols*. To facilitate comparison, Paper III therefore turns to epistemic temporal models generated by sets of dynamical systems, imitating non-deterministic evolution, and identifies necessary properties inherited by the resulting epistemic temporal models.

Finally, the paper compares the two protocol frameworks. It would be wonderful to be able to shortly list the exact difference between the two, but a comparison is not trivial: There are subtle details in the two frameworks that make some properties of the generated models natural for the one, but not for the other, and *vice versa*. In particular, working with *pointed* Kripke and action models is a prerequisite in the dynamical systems approach and, as a consequence, some of the characterizing properties are relative to structural features of models' points. Working with pointed models is not a prerequisite in the extensional protocols framework, so the results of van Benthem, Gerbrandy, Hoshi and Pacuit concern properties without reference to such. This makes a direct comparison difficult: In attempting to abstract away from the points of pointed models, structural features important to the dynamical systems approach are ignored, yielding a sub-optimal comparison.[24]

In a rough summary, the comparison of the two frameworks shows that *i*) every intensional protocol may be mimicked by an extensional protocol but not *vice versa*, albeit even for finitely represented intensional protocols, the mimicking extensional protocol may require a countably infinite representation; *ii*) as a consequence of the possibility of mimicking, the epistemic-temporal agent types implicitly encoded using intensional protocols inherit the rationality properties shown in [28] to characterize extensional protocol agents; *iii*) in addition, intensional protocol driven agents (and models) satisfy the stronger uniformity property that identical circumstances produce identical reactions, reflecting that intensional protocols are mappings. This property is not satisfied by extensional protocols, as these dictate actions upon consulting an "external clock".

---

[24]Should one seek an exact comparison between the two frameworks, my suggestion would be to re-undertake the original study of van Benthem, Gerbrandy, Hoshi and Pacuit, but using pointed models.

For the thesis at large, the results of Paper III contribute that the dynamical systems approach to modeling information dynamics does not stand back to the extensional protocols approach: Each has benefits over the other and framework choice would advisably be made based on application.

# 6 Paper IV: Proof of Concept

Like Paper III, Papers IV and V also concern two frameworks, though this time around, the relationship of interest is not a comparison, but an embedding. In these two papers, Dominik Klein and I approach the problem of how logical dynamics of dynamic epistemic logic may be recast in a manner that allows for analysis through *topology* and *dynamical systems theory*.

When discussing dynamical systems up to this point, this has been done with reference to the broad conception from Paper II: A set together with a map acting on it. In most treatments, however, a more advanced starting point is taken by assuming that the set has some sort of *structure* and that the map somehow behaves *reasonably* with respect to this structure. Specifically, a *topological dynamical system* is often defined as a *metrizable compact topological space* together with a *continuous map* acting on it (see e.g. [57, 152]).

The two questions of what that means and why one would want it are *not* equally hard to answer. The latter is easier: The properties are desirable as they impart structure, structure facilitates proving theorems, and theorems aids in understanding the often complex beast under investigation. As topological dynamical systems have been investigated *extensively* in both pure and applied mathematics since the early twentieth century, the hope Dominik and I share is that some of the theoretical developments will aid our understanding of logical dynamics. Paper IV is a first proof of concept.

On to the harder question: First, what is relevant here is first the slightly simpler notion of a *compact metric space*. An intuition is an ordinary room with objects in it, including a ruler. The ruler may be used to measure the distances between the objects: It provides a *metric* on the set of objects. The walls of the room provide boundaries: They ensure that all the objects in the room are within finite distance of each other, that there are no infinite horizons. The room is then called *compact*. Jointly, the objects, the ruler and the walls constitute a compact metric space.

To illustrate continuity, put a person in the room and watch them behave, watch them move about. At a later time, it becomes of interest to duplicate the first behavior, up to some given margin of error. If we are then allowed also a very small margin of error in putting them back into their original position while their behavior still stays within the desired range, then they are continuous. In short, If close outputs are desired, then close enough inputs

suffice. On the other hand, if the person sometimes immediately does something erratic just because the initial position was off the slightest, then they are discontinuous. Erratic people are not the easiest to work with—and the same goes for maps.

A topological dynamical system is, then, a room with a non-erratic person, and the study of a topological dynamical system concerns how the person moves through the room as time progresses. Questions asked may include where the person moves given some start state, whether there are points in the space with respect to which the always stays close, or perhaps whether there are points that the person always returns to after some time.

Paper IV contains a set of results on the long-term dynamics of multi-pointed action models, recast as topological dynamical systems. The paper relies heavily on the work on metrics and topologies developed in Paper V, detailed in the next section.[25] Paper IV is included before Paper V in this thesis as it includes a self-contained overview of the approach as well as a concrete motivating problem of study, two features which in combination hopefully eases the reading of the theoretical Paper V.

The results concern a conjecture by Johan van Benthem, concerning the long-term behavior of a particular type of maps: Those based on finite action models and without postconditions. The conjecture was that whenever a such was iteratively applied to a finite pointed Kripke model, it would have a periodic orbit. I.e., after finite time, it would return to to a point previously visited, from where it would loop. This was the content of van Benthem's *Finite Evolution Conjecture* [21]. The conjecture was refuted using a counterexample by Thomas Sadzik in his 2006 paper, [141].

In the Finite Evolution Conjecture, the requirement is that the map returns *exactly* to a point it has previously visited. However, if working in a metric (or just topological) space, there are weaker concepts of "returning to". In particular, the concept of a *recurrent point* is of interest: Roughly, a point is recurrent if, when the map moves on from it, then no matter where the map goes, it will again return to being arbitrarily close to the point. This may not happen in finite time, but the map will—in the very long run—return.

Paper IV, then, in a sense, investigates a weakening of the Finite Evolution Conjecture: Will every finite action model when iteratively applied to a finite pointed Kripke model give rise to a finite set of recurrent points?

To use the topological notion of recurrence, a topology is required. However, the topology will taint the results: What is recurrent in one topology may not be recurrent in another. In general, the topological results are only interesting to the degree that the topology fits with the intuitions of the subject matter. Paper IV therefore first seeks a topology that fits

---

[25]The difference between the settings in Paper IV and Paper V is small, with the setting of Paper V slightly simpler: In defining *modal spaces*, the logic is not necessary in defining the quotient, and is consequently omitted in Paper V.

with natural intuitions about logical dynamics.

Though our approach to topology was originally driven by metrics—as is visible from Paper v—Paper iv takes a different starting point, namely *convergence*. Informally, a sequence of points converge to some final destination when for no matter how small a distance, there is some element in the sequence that is within that distance from the destination, and so is every element after it. Roughly, the sequence forever grows closer and closer to the destination.

In Paper iv, we suggest a logical analogue of this geometric intuition: A sequence of models *logically converge* to some destination when for every formula of the chosen language which is satisfied at the destination, there is some element in the sequence that satisfies that formula, and so does every element after it. Roughly, the sequence forever grows more and more like the destination.

Logical convergence is then used to discriminate between topologies: We seek one satisfying the intuitive demand that a sequence topologically converges to a point if, and only if, the sequence logically converges to the point. We find that the *Stone topology* satisfies the desideratum where others fail. The Stone topology is further natural from a logical point of view: It is induced by a basis of clopen sets with each basis element the truth set of a formula. Based on these reasons, we feel that the Stone topology is sufficiently natural to make results concerning recurrence in it interesting.

Paper iv could have contained a stronger argument for the choice of topology: In Paper v, it is shown that the *desideratum* in fact *characterizes* the Stone topology.

Relative to the Stone topology, Paper iv then shows three main proposition concerning update with finite action models. The first result is immediate when combining the topological approach with a result shown by Sadzik: Every finite action model with only Boolean preconditions and without postconditions have exactly one recurrent point. Somewhat ironically, an application of the proposition shows that Sadzik's counterexample to the Finite Evolution Conjecture in the topological setting converges: It has a single recurrent point. Where in the non-topological setting, the example constitutes a complex dynamics never returning to the same point, in the topological setting, it constitutes a very simple sequence.

The analysis of Sadzik's example shows the usefulness of the topological approach to logical dynamics. Analyzing the example in the non-topological setting is somewhat akin to analyzing the sequence $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ without a metric for measuring distance between numbers: That sequence, too, never returns to the same point. Treating numbers as—well, as *numbers*, the sequence may never return to the same place, put it's clearly *going* somewhere: To 0. In this light, the sequence is not infinitely complex, but nicely structured.

Paper ɪᴠ's last two propositions on recurrence show that once updates are allowed to be more complex, they are no-longer guaranteed to produce so nicely structured dynamics: Both finite action models with Boolean preconditions and with postconditions, and finite action models with non-Boolean preconditions, but no postconditions, may produce orbits with uncountably many recurrent points. Hence, a conjecture concerning the finiteness of the sets of recurrent points would be incorrect.

The proofs of two latter propositions rely on the *Turing completeness* of the respective classes of maps, the first class shown such by Thomas Bolander and Mikkel Birkegaard Andersen in [36], the second by Dominik Klein and I in the note *Turing Completeness of Finite Epistemic Programs*, included as an appendix to Paper ɪᴠ.

For the thesis at large, the results of Paper ɪᴠ provide a proof of concept for the applications of the topological concepts from dynamical systems theory to logical dynamics: In light of the concepts of convergence and recurrence, insights into the structure of the unfolding dynamics are discernible where without they are not. To iterate the point, then these results are only interesting to the degree that the Stone topology is found natural, a point which recurs in Paper ᴠ.

## 7 Paper ᴠ: Metrics and a Topological Basis

Where Paper ɪᴠ concerns recurrence in the Stone topology, the joint work with Dominik initiated with a search for metric between pointed Kripke models: As a metric induces a topology while representing distance in an intuitive manner, finding a natural metric is a possibly more intuition-rich approach to obtain a natural topology.

Indeed, the first metric we designed was heavily based on our intuitions of what makes pointed Kripke models more or less alike: It is included in Paper ᴠ in the example *Close to Home, Close to Heart*. It is included as one among several examples as we have since its design severely generalized the approach. Thus, the starting point of Paper ᴠ is not a metric applicable to sets of pointed Kripke models, but a family of metrics applicable to sets of countably infinite strings. By extension, these metrics are applicable in general to formal structures serving as semantics for countable languages. Among such are pointed Kripke models, on which the paper focuses.

The intuition behind the original metric illustrates the general approach: Think of going out the door, taking a walk past a sequence of features. You pass by an open shop, a broken tile, a filled bike-rack, and so forth. The next day, take the same walk: The shop is still open, the tile has been fixed, the bike-rack is still filled. Throughout the walk, you witness similarities and differences. How alike are these walks? If all the possible walks were objects in a room, how far would these two be apart?

If the all the walks are finite, then one solution is straightforward: Simply count the number of differences between any two walks. If there are three differences, then the distance is 3. If the walks pass by some number $n$ features, then the most different walks will be $n$ apart. This way of measuring distance, this metric, is known as the *Hamming distance* [80].

If the walks are infinite, however, then the Hamming distance doesn't work: Two walks may then be different on an infinity of features, but as "infinitely far apart" is not an entry on any regular ruler, it is not an acceptable answer.

Working with formal models, infinite walks occur more often than not, so the Hamming distance is not applicable. This is where the intuition behind the *Close to Home, Close to Heart* metric enters the picture. The idea is that *all differences should weigh in, but not equally much*. Specifically, the further from the front door a difference is seen, the less it counts. If the weights of the different features are then picked suitably, adding all the weighed differences between any two walks will be a real number, present on the ruler.

One way to pick the weights in a suitable manner is to pick them so they give rise to a *convergent series*: An example is $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ Thought the sum of these numbers, $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$, contains infinitely many terms, the result is still finite: In this case, it's 2. That, then, is the maximum distance between two walks, using these weights. If only the two walks differ on features 2 and 4, the difference is $\frac{5}{8}$.

Interpreting the features met on the walk as formulas being either true or false, this weighted Hamming distance for infinite sequences[26] defines a metric on (quotients of) sets of formal structures, in particular on sets of pointed Kripke models.

This first metric induces a well-behaved topology, with respect to which maps induced by multi-pointed action models with postconditions applied using product update are continuous. Under this metric, then, sets of pointed Kripke models under the action of such induced maps constitute topological dynamical systems. Success.

However, as in the case of the choice of topology in Paper IV, results obtained using a metric are tainted by the choice of metric: If the metric is not right for the subject matter, then neither will be the results.[27]

There are many cases where this first metric is not right, as when one cares more about other properties of features than when on the walk they are met. Very reasonably, one might consider that it makes a bigger difference between walks if a favorite café on the other side of town is one day a bookstore, than if a tile just outside the door has been fixed.

---

[26]To the best of our knowledge, the generalization is new—at least we have failed to find it in the comprehensive *Encyclopedia of Distances* [47].

[27]We thank Johan van Benthem for this critical point, without which we might have stayed satisfied with the single metric. The point was in private conversation: Any possible misconstrual is fully my responsibility.

Moreover, one may no even care about certain features, like the tile.

Based on these considerations, Paper v presents a general approach to defining metrics for formal structures described by countable languages. The approach combines weighing with the Hamming distance, but allows adjustments of the weights as one sees fit together with a few additional generalizations. Paper v makes the the approach formally precise and investigates it.

Having introduced the approach to metrics and their application to formal structures in general, the paper turns specifically to pointed Kripke models, initially illustrating the general approach applied. Ensuingly, the paper shows a number of examples, defining various metrics natural from a modal logical point. One of the examples show that the syntactic approach adopted strictly generalizes several other metrics from the very recent literature.

The paper then turns to topology. Initially, the paper shows that two metrics are topologically equivalent whenever they agree on which formulas of the modal language should receive strictly positive weight. The resulting topologies are slight generalizations of the Stone topology, which we refer to as as *Stone-like topologies*.

Having established this initial result, Stone-like topologies become the focal point. They are shown to be always *Hausdorff* and *totally disconnected*. Moreover, if one is so lucky that certain assumptions apply, they are also compact.

The same section also paints a picture of Stone-like topologies by relating their open, closed and clopen sets to formulas of the modal language. In this respect, Stone-like topologies are well-behaved, in particular in the compact case. There, a set is clopen if, and only if, it is the truth set of some formula.

The paper then turns to convergence and limit points. Here, the result mentioned in Section 6 on Paper IV is stated and shown: The Stone-like topology for a set of formulas $D$ is the unique topology for which logical convergence with respect to $D$ is equivalent with topological convergence.

It is not in the style of Paper v to dwell on its results, but for the sake of this exposition, the result should be highlighted: Conceptually, the results implies that if one agrees that logical convergence captures the intuition of convergence when working with logical dynamics, then the unique topology that will be satisfactory is the corresponding Stone-like topology. Essentially, then, if one want to work with logical convergence, then working with the Stone-like topology is *the* choice of topology that does not taint results.

Beyond this key conceptual result, Paper v presents a simple characterization of isolated points and exemplifies perfect, imperfect and discrete spaces. In the compact case, the former of these types are homeomorphic to the Cantor set.
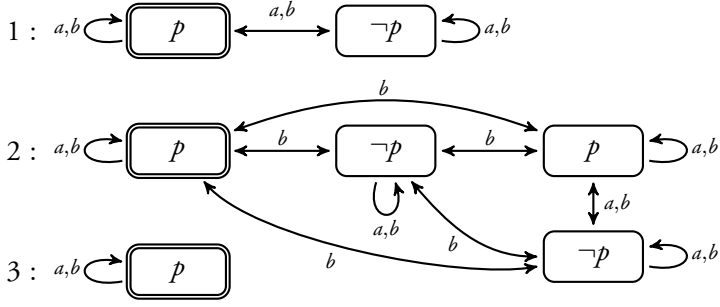
**Figure 4:** The three epistemic states of the dynamics from Section 2, labeled on the left by time.

Finally, the paper turns to mappings, where the continuity result mentioned above is generalized: Maps induced by multi-pointed action models with postconditions applied using product update are continuous with respect to the Stone-like topology. With this result, the paper establishes the desired connection between dynamic epistemic logic and topological dynamical systems, essential for the thesis at large.

# 8 Paper VI: Reduction Laws and Continuity

Papers I-V have focused on one aspect of the perspective mentioned in the, namely to formalize models as *mappings*, a not quite common community practice. That the updates used in model building allow for *reduction laws*, though, is so much common practice as to almost characterize the community.

Reduction laws have hitherto been mentioned but in parsing, as they make no entries in the thesis but in the final Paper VI. Equally, the *logic* which is dynamic epistemic logic has been glanced over: Up to now, only the semantic side of dynamics has been considered.

To sketch the approach, consider again the sequence of the three epistemic states produced by the dynamics in Section 2, as shown in Figure 4.

The formal language introduced included knowledge formulas to talk about the agents' information. The first epistemic state satisfies $\neg K_a p \wedge \neg K_b p$, the second satisfies $K_a p \wedge \neg K_b p$ and the third $K_a p \wedge K_b p$.

The first step in creating a logic of these dynamics involve extending the formal language in a way so it can also *talk about dynamics*. This is done by augmenting the language with so-called *dynamic operators*. To describe the above dynamics, two such are needed, one for each informational event. Let $\Sigma_1$ denote the model transformation defined by the pointed action model that applied using product update produced epistemic state 2 from epistemic

state 1, and let $\Sigma_2$ denote that produced 3 from 2. The "dynamic diamonds" $\langle\Sigma_1\rangle$ and $\langle\Sigma_2\rangle$ are then our two dynamic operators.

To include the dynamic operators in the language, augment the grammar so that both $\langle\Sigma_1\rangle\varphi$ and $\langle\Sigma_2\rangle\varphi$ are formulas whenever $\varphi$ is. The formula $\langle\Sigma_1\rangle\varphi$ states, roughly, that applying $\Sigma_1$ to the current epistemic state produces an epistemic state in which $\varphi$ is true. For a precise semantics, see e.g. [20].

In epistemic state 1, the dynamic formula $\langle\Sigma_1\rangle K_a p$ is then true, as verified by checking that $K_a p$ is satisfied in epistemic state 2. On the contrary, both $\langle\Sigma_1\rangle K_b p$ and $\langle\Sigma_2\rangle p$ are false, but for different reasons: The first fails as $K_b p$ is false in epistemic state 2, the second because $\Sigma_2$ cannot be applied successfully to epistemic state 1—its execution requires that the precondition $K_a p$ of the actual event is true in the epistemic state, but it is not. However, the more complex $\langle\Sigma_1\rangle\langle\Sigma_2\rangle p$ again is true: In epistemic state 1, applying $\Sigma_1$ takes us to epistemic state 2 where applying $\Sigma_2$ produces epistemic state 3, satisfying $p$.

With dynamic operators, it thus possible to talk about the semantic dynamics using the formal language. As a result, it is possible to construct *logics* for the dynamics: Formal theories proving properties about the dynamics, cast in the formal language. Now, logics for Kripke models cast in languages without dynamic operators are common place: If we ignore the dynamic part of the story, then it is often a simple task to find a logic that is *sound* and *complete* with respect to the class of models use to represent the static states.[28]

Enter reduction laws: A reduction law is a formula which claims the equivalence of a formula with a dynamic operator and a formula without it.[29] If such a reduction law is valid on the class of models of interest, then it allows one to *reformulate* the dynamic statement to a static statement—it allows the reduction of dynamics to statics.

Here's the kicker: If one can find a reduction for every dynamic formula, then adding the corresponding (valid) reduction laws to the static logic as axioms results in a dynamic logic that is sound and complete for the dynamic semantics of interest. As logicians revel in sound and complete axiomatizations for semantics, and dynamic semantics have been immensely popular over recent years, the contemporary literature is brimming with new semantic dynamics, dynamic languages, and dynamic logics based on reduction laws.

This is point of entry of Paper VI. It proves two characterization theorems linking the topological approach of Papers IV and V with the reduction law approach to completeness. The theorems are of similar form and apply to respectively compact and non-compact logics. Roughly, the theorem for the compact case states the following:

---

[28] A logic sound with respect to a class $C$ proves *only* things that are true (valid) in $C$. A logic complete with respect to $C$ proves *all* the true (valid) things. Thus, a sound and complete logic complete with respect to $C$ captures *all* and *only* truths (validities). In short, it's a very, very nice theory.

[29] Or at least with a lower *dynamic complexity*, but for details see e.g. [54].

Let $f\colon S \to S$ be a map on a set of structures $S$ described by a static language $\mathcal{L}$. Let $\langle f \rangle$ be the dynamic operator of $f$ and $\mathcal{L}_f$ the dynamic extension of $\mathcal{L}$. Then $\mathcal{L}_f$ is reducible to $\mathcal{L}$ if, and only if, $f$ is continuous with respect to the Stone topology on $S$.

The left-to-right direction also holds in the non-compact case, but the right-to-left direction needs a stronger notion than continuity. Paper VI shows that requiring that the preimage of any basis element is a basis element does the job. Maps of this type are dubbed *confined*.

In one reading, the characterization theorems of Paper VI thus show that to recast reduction law-friendly logical dynamics as topological dynamical systems, one only needs to sprinkle on a bit of math: The Stone topology.


# 9 Concluding Remarks

As stated in the introduction, the topic of this thesis is information dynamics and how to formally model them using dynamic epistemic logic. The perspective investigated is simple: Formalize the model as a map that allows for the formulation of reduction laws. Then, with a bit of math sprinkled on top, the model will be a topological dynamical system.

Though Papers I-VI were not written with the forethought of composing a coherent argument concerning why this perspective is reasonable nor with the forethought of establishing the stated consequence, I hope that the above exposition lends credit to the case. In summary, a tentative argument may be put together as follows.

Paper I on the bystander effect shows how augmenting the standard machinery of dynamic epistemic logic with a decision making framework yields mathematically self-contained models of dynamic processes, a prerequisite for rigid model comparison.

Paper II extrapolates from Paper I's construction, showing how the added decision making framework and its natural peers may be construed as maps. It makes the link to dynamical systems explicit, and shows that under the restriction of dynamics produced by the iteration of the identified maps still falls a collection rich enough to be of interest.

Paper III compares the dynamical systems approach of Paper II with the main existing augmentation to dynamic epistemic logic for obtaining self-contained models, namely extensional protocols. It concludes that each framework has its benefits, depending on application. In favor of the dynamical systems approach, it shows how extensional protocols designed to mimic simple, finite models of the former kind require infinite representations.

With Paper IV, the focus shifts to topological dynamical systems. The paper argues that the Stone topology is a natural topology under which to investigate logical dynamics by showing that it satisfies the intuitive *desideratum* that an introduced notion of 'logical con-

vergence' coincides with topological convergence. Ensuingly, the paper investigates the recurrent behavior of the map types from Papers II and III, where it is shown that the topological perspective adds novel, structural insights to the analysis of their long-term behavior.

Paper V provides the theoretical backbone of Paper IV, but also shows an important strengthening of the argument underlying the choice of analyzing logical dynamics under specifically the Stone topology: It shows that the Stone topology is the unique topology for which the logical and topological convergences are equivalent. It further includes a metric-based proof that the hitherto analyzed maps are continuous with respect to the Stone topology.

Paper VI, finally, shows a tight connection between continuity in the Stone topology and the existence of reduction laws for dynamic modalities, in the compact case even yielding a characterization. With showing that reducibility always implies continuity, the results of Paper VI makes it straightforward to recast many types of logical dynamics of contemporary interest as topological dynamical systems.

In sum, then, taking a dynamical systems perspective on logical dynamics increases rigidity in model construction while allowing large freedoms in the dynamics representable, on top of which there is a unique natural candidate topology allowing for novel structural insights of the behavior of maps that may in addition be shown continuous by the well-established logical method of reduction laws. With that, the thesis formulates a bond between relatives, with the hope that it will be seen as a reasonable contribution to logic and formal epistemology.

Of course, nothing shows a contribution reasonable as the further research it can help facilitate. The following three venues stand salient:

A first venue is to explore the promised lands of dynamical systems theory and its region of logical dynamics. The dynamical systems theory is rich with concepts, methodologies and results which may be of aid and interest in analyzing logical dynamics. To understand which tools are applicable and how to wield them in the logical setting is a broad endeavor. With Paper IV, the thesis takes a first step, showing that the maps induced by multi-pointed action models and product update show so-called *nontrivial recurrence*, a term adopted from Hasselblatt and Katok's survey of the dynamical systems field, [85]. The authors remark that it is the first indication of complicated asymptotic behavior, but that it in come low-dimensional systems it is possible to obtain comprehensive description of it. As the Stone topology is 0-dimensional, a concrete place to start a survey is on recurrence, in the references provided by Hasselblatt and Katok on low-dimensional spaces. Similarly starting from 0-dimensionality is in seeking insights from of well-known setting sharing this feature, such as *Cantor spaces*, and *shifts* and *symbolic dynamics* [118, 152]. Hopefully, studies on these topics will provide general tools to illuminate structural features of logical dynamics.

A second venue is applications in formal epistemology. Paper IV remarks on a puzzling aspect of common knowledge: Where it cannot be attained by asynchronous communication in finite time [58, 59, 79], it may be obtained semantically in the limit—but only in languages without the formal operator. Hence, you can attain common knowledge in the limit iff you cannot talk about it. There is an interplay here between expressibility, non-compactness, topology and formal epistemology which I hope to address with Dominik in later work. Of interest here is also whether the alternative notion of *relativized common knowledge* [27]—which in contrast to the standard notion plays nice with reduction axioms—will facilitate agents' getting it right. But formal epistemology is not only about getting it right—it's also about understanding the structures involved in getting it wrong. Paper I is an example, and so is the classic model of *informational cascades* [33]. Logical models have captured the higher-order reasoning by which individually rational agents as a group go wrong when cascading [1, 11, 135]. With dynamics highly dependent only initial conditions, it is possible that these—and other social influence dynamics—are *chaotic*, but might be nudged towards socially desirable *attractors* [152]. How specifically to study such *information control problems* is an open, but engaging problem [82, 83]. As a final topic pertaining to formal epistemology, then the relations between logical dynamics cast in dynamical systems and *formal learning theory* [67, 68, 87, 98] deserve clarification. Formal learning theory is well-established in formal epistemology and deals with topics and concepts also arising with the dynamical systems approach, such a long-term and limit behavior, topology and convergence, applied in relation to rational learners. With the link between dynamic epistemic logic and formal learning theory already established [67], the hope is that the addition of the topological, dynamical systems perspective could be of mutual interest.

A third venue concerns the logic of the dynamical systems approach to dynamic epistemic logic. As mentioned in Section 8 on Paper VI, sound and complete logics may be obtained through reduction laws for certain model transformations. However, formulas with dynamic operators are still finite, and can therefore only talk of finite dynamics from any given pointed Kripke model. With the Finite Evolution Conjecture mentioned in Section 6 on Paper IV is refuted and relevant properties of dynamic epistemic logical dynamics occurring in the limit, it thus seems that the finite horizon language of dynamic modalities stay silent on dynamically relevant properties. To capture these aspects of the dynamic epistemic logical semantics, it seems that a different language is required. To find a suitable language, it may serve to consult a body of logical literature which takes perspective on the interplay of logic and dynamical systems opposite of this thesis: Applying logics to dynamical systems. In this line, Sarenac takes a modal approach to describing *iterated function systems*[143] and Platzer takes a dynamic logic approach to *hybrid systems* [131, 132], while van Benthem outlines possible approaches to fixed points and limit cycles of dynamical systems by applying fixed-point and oscillation operators galvanized by *modal μ-calculus* [26]. Directly focused on logics of topological dynamical systems is the literature on *dynamic*

*topological logic*, springing from early work by combinations of Artemov, Davoren, Kremer, Mints, Narode and Rybakov [6, 45, 103, 104, 106] in the late 1990's. Both Kremer and Mints' handbook chapter [105] and the series of papers by Fernandéz-Duque [60, 61, 62] seem of especial interest, with results on, respectively, continuous maps on Cantor spaces and metric spaces. Whether either of these approaches may capture the dynamics of dynamic epistemic logic to satisfaction is an open question—looking at logical dynamics as dynamical systems allows it to be posed.

# References

[1] Andreea Achimescu, Alexandru Baltag, and Joshua Sack. The Probabilistic Logic of Communication and Change. *Journal of Logic and Computation*, 2016.

[2] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[3] Carlos E. Alchourrón and David Makinson. Hierachies of Regulations and their Logic. In R. Hilinenn, editor, *New Studies in Deontic Logic*, volume 152 of *Synthese Library*, pages 125–148, 1982.

[4] Carlos E. Alchourrón and David Makinson. On the Logic of Theory Change: Safe Contraction. *Studia Logica*, 44:405–422, 1985.

[5] Horacio Arló-Costa, Vincent F. Hendricks, and Johan van Benthem, editors. *Readings in Formal Epistemology*. Springer Graduate Texts in Philosophy, Vol. 1. Springer, 2016.

[6] Sergei Artemov, Jennifer M. Davoren, and Anil Nerode. Modal Logics and Topological Semantics for Hybrid Systems. Technical report, Cornell University, 1997.

[7] Guillaume Aucher. A Combined System for Update Logic and Belief Revision. In M. W. Barley and N. Kasabov, editors, *Intelligent Agents and Multi-Agent Systems*, pages 1–17. Springer, 2005.

[8] Robert J. Aumann. Agreeing to Disagree. *Annals of Statistics*, 4:1236–1239, 1976. Republished in [5], 2016.

[9] Philippe Balbiani, Alexandru Baltag, Hans van Ditmarsch, Andreas Herzig, Tomohiro Hoshi, and Tiago de Lima. What Can We Achieve by Arbitrary Announcements?: A Dynamic Take on Fitch's Knowability. In *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '07, pages 42–51. ACM, 2007.

[10] Roberta Ballarin. Modern Origins of Modal Logic. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2017 edition, 2017.

[11] Alexandru Baltag, Zoé Christoff, Jens Ulrik Hansen, and Sonja Smets. Logical Models of Informational Cascades. In J. van Benthem and F. Liu, editors, *Logic Across the University: Foundations and Applications—Proceedings of the Tsinghua Logic Conference, Beijing, 14-16 October 2013*, Studies in Logic, Volume 47, pages 405–432. College Publications, 2013.

[12] Alexandru Baltag, Zoé Christoff, Rasmus K. Rendsvig, and Sonja Smets. Dynamic Epistemic Logic of Diffusion and Prediction in Threshold Models. *Studia Logica*, to appear.

[13] Alexandru Baltag, Lawrence S. Moss, and Sławomir Solecki. The Logic of Public Announcements, Common Knowledge, and Private Suspicions. In *TARK 1998*, pages 43–56. Morgan Kaufmann Publishers, 1998.

[14] Alexandru Baltag, Aybüke Özgün, and Ana Lucia Vargas Sandoval. Topo-Logic as a Dynamic-Epistemic Logic. In A. Baltag, J. Seligman, and T. Yamada, editors, *Logic, Rationality, and Interaction (LORI 2017, Sapporo)*, pages 330–346. Springer, 2017.

[15] Alexandru Baltag and Bryan Renne. Dynamic Epistemic Logic. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2016 edition, 2016.

[16] Alexandru Baltag and Sonja Smets. Conditional Doxastic Models: A Qualitative Approach to Dynamic Belief Revision. *Electronic Notes in Theoretical Computer Science*, 165:5–21, 2006.

[17] Alexandru Baltag and Sonja Smets. A Qualitative Theory of Dynamic Interactive Belief Revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 9–58. Amsterdam University Press, 2008.

[18] Alexandru Baltag and Sonja Smets. Group Belief Dynamics under Iterated Revision: Fixed Points and Cycles of Joint Upgrades. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '09, pages 41–50. ACM, 2009.

[19] Alexandru Baltag, Sonja Smets, and Jonathan A. Zvesper. Keep 'Hoping' for Rationality: a Solution to the Backward Induction Paradox. *Synthese*, 169(2):301–333, Jul 2009.

[20] Alexandu Baltag and Lawrence S. Moss. Logics for Epistemic Programs. *Synthese*, 139(2):165–224, 2004.

[21] Johan van Benthem. "One is a Lonely Number": Logic and Communication. In Z. Chatzidakis, P. Koepke, and W. Pohlers, editors, *Logic Colloquium '02*, Lecture Notes in Logic, 27, pages 95–128. Association for Symbolic Logic, 2002.

[22] Johan van Benthem. Dynamic Logic for Belief Revision. *Journal of Applied Non-Classical Logics*, 17(2):129–155, jan 2007.

[23] Johan van Benthem. Rational Dynamics and Epistemic Logic in Games. *International Game Theory Review*, 09(1):13–45, 2007.

[24] Johan van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011.

[25] Johan van Benthem. *Logic in Games*. The MIT Press, 2014.

[26] Johan van Benthem. Oscillations, Logic, and Dynamical Systems. In S. Ghosh and J. Szymanik, editors, *The Facts Matter. Essays on Logic and Cognition in Honour of Rineke Verbrugge*, pages 9–22. College Publications, 2016.

[27] Johan van Benthem, Jan van Eijck, and Barteld Kooi. Logics of Communication and Change. *Information and Computation*, 204(11):1620–1662, 2006.

[28] Johan van Benthem, Jelle Gerbrandy, Tomohiro Hoshi, and Eric Pacuit. Merging Frameworks for Interaction. *Journal of Philosophical Logic*, 38(5):491–526, 2009.

[29] Johan van Benthem, Jelle Gerbrandy, and Barteld P. Kooi. Dynamic Update with Probabilities. *Studia Logica*, 93:67–96, 2009.

[30] Johan van Benthem and Fenrong Liu. Dynamic Logic of Preference Upgrade. *Journal of Applied Non-classical Logics*, 17:157–182, 2007.

[31] Johan van Benthem and Eric Pacuit. Dynamic Logics of Evidence-Based Beliefs. *Studia Logica*, 99:61–92, 2011.

[32] Johan van Benthem and Sonja Smets. Dynamic Logics of Belief Change. In H. van Ditmarsch, J. Y. Halpern, W. van der Hoek, and B. Kooi, editors, *Handbook of Logics for Knowledge and Belief*, pages 299–368. College Publications, 2015.

[33] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992–1026, jan 1992.

[34] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, 2001.

[35] Thomas Bolander. Seeing is Believing: Formalising False-Belief Tasks in Dynamic Epistemic Logic. In A. Herzig and E. Lorini, editors, *European Conference on Social Intelligence (ECSI 2014)*, pages 87–107. CEUR Workshop Proceedings, vol. 1283, 2014.

[36] Thomas Bolander and Mikkel Birkegaard Andersen. Epistemic Planning for Single- and Multi-agent Systems. *Journal of Applied Non-Classical Logics*, 21(1):9–34, 2011.

[37] Craig Boutilier. Iterated Revision and Minimal Change of Conditional Beliefs. *Journal of Philosophical Logic*, 25:263–305, 1996.

[38] Roderick M. Chisholm. The Logic of Knowing. *Journal of Philosophy*, (60):773–795, 1963.

[39] Zoé Christoff. *Dynamic Logics of Networks*. PhD thesis, Institute for Logic, Language and Computaiton, University of Amsterdam, 2016.

[40] Zoé Christoff and Jens Ulrik Hansen. A two-tiered formalization of social influence. In D. Grossi, O. Roy, and H. Huang, editors, *Logic, Rationality, and Interaction (LORI 2013, Hangzhou)*, volume 8196 of *Lecture Notes in Computer Science*, pages 68–81. Springer, 2013.

[41] Zoé Christoff and Jens Ulrik Hansen. A Logic for Diffusion in Social Networks. *Journal of Applied Logic*, 13:48–77, 2015.

[42] Michael Cohen. A Dynamic Epistemic Logic with a Knowability Principle. In W. van der Hoek, W. H. Holliday, and W. Wang, editors, *Logic, Rationality, and Interaction*, volume 9394 of *Lecture Notes in Computer Science*, pages 406–410. Springer, 2015.

[43] Kenneth Collier. Hintikka's Epistemic Logic. In R. J. Bogdan, editor, *Jaakko Hintikka*, pages 181–198. Springer, 1987.

[44] Adnan Darwiche and Judea Pearl. On the Logic of Iterated Belief Revision. *Artificial Intelligence*, 89(1-2):1–29, 1997.

[45] Jennifer M. Davoren. *Modal Logics for Continuous Dynamics*. PhD thesis, Cornell University, 1998.

[46] Arthur P. Dempster. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Statist.*, 38(2):325–339, 04 1967.

[47] Michel Marie Deza and Elena Deza. *Encyclopedia of Distances*. Springer, fourth edition, 2016.

[48] Hans van Ditmarsch. Prolegomena to Dynamic Logic for Belief Revision. *Synthese*, 147(2):229–275, Nov 2005.

[49] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. Playing Cards with Hintikka. *Australasian Journal of Logic*, (3):108–134, 2005.

[50] Hans van Ditmarsch, Sophia Knight, and Aybüke Özgün. Arbitrary Announcements on Topological Subset Spaces. In N. Bulling, editor, *Multi-Agent Systems*, volume 8953 of *Lecture Notes in Computer Science*, pages 252–266. Springer, 2015.

[51] Hans van Ditmarsch, Sophia Knight, and Aybüke Özgün. Announcement as Effort on Topological Spaces. *Synthese*, 2017.

[52] Hans van Ditmarsch and Barteld P. Kooi. The Secret of My Success. *Synthese*, 151:201–232.

[53] Hans van Ditmarsch and Barteld P. Kooi. Semantic Results for Ontic and Epistemic Change. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 87–117. Amsterdam University Press, 2008.

[54] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer, 2008.

[55] Igor Douven and Jonah N. Schupbach. Formal Epistemology. In *Oxford Handbooks Online*, 2017.

[56] Jan van Eijck, Floor Sietsma, and Yanjing Wang. Composing models. *Journal of Applied Non-Classical Logics*, 21(3-4):397–425, 2011.

[57] Tanja Eisner, Bálint Farkas, Markus Haase, and Rainer Nagel. *Operator Theoretic Aspects of Ergodic Theory*. Springer, 2015.

[58] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. The MIT Press, 1995.

[59] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. Common Knowledge Revisited. *Annals of Pure and Applied Logic*, 96(1):89–105, 1999.

[60] David Fernández-Duque. A Sound and Complete Axiomatization for Dynamic Topological Logic. *Journal of Symbolic Logic*, 77(3):1–26, 2012.

[61] David Fernández-Duque. Dynamic Topological Logic of Metric Spaces. *Journal of Symbolic Logic*, 77(1):308–328, 2012.

[62] David Fernández-Duque. Non-Finite Axiomatizability of Dynamic Topological Logic. *ACM Transactions on Computational Logic*, 15(1):1–18, 2014.

[63] Paolo Galeazzi. *Play without Regret*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2017.

[64] Paolo Galeazzi and Emiliano Lorini. Epistemic Logic meets Epistemic Game Theory: a Comparison Between Multi-agent Kripke Models and Type Apaces. *Synthese*, 193(7):2097–2127, Jul 2016.

[65] Jelle Gerbrandy. *Bisimulations on Planet Kripke*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1998.

[66] Jelle Gerbrandy and Willem Groeneveld. Reasoning about Information Change. *Journal of Logic, Language and Information*, 6(2):147–169, 1997.

[67] Nina Gierasimczuk. *Knowing One's Limits*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2010.

[68] Nina Gierasimczuk, Vincent F. Hendricks, and Dick de Jongh. Logic and Learning. In A. Baltag and S. Smets, editors, *Johan van Benthem on Logic and Information Dynamics*, volume 5 of *Outstanding Contributions to Logic*, pages 267–288. Springer, 2014.

[69] Itzhak Gilboa and David Schmeidler. Maxmin Expected Utility with Non-unique Prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[70] Patrick Girard and Hans Rott. Belief Revision and Dynamic Logic. In A. Baltag and S. Smets, editors, *Johan van Benthem on Logic and Information Dynamics*, volume 5 of *Outstanding Contributions to Logic*. Springer, 2014.

[71] Patrick Girard, Jeremy Seligman, and Fenrong Liu. General Dynamic Dynamic Logic. In T. Bolander and T. Braüner, editors, *Advances in Modal Logic, vol 9.*, pages 239–260. College Publications, 2012.

[72] Valentin Goranko and Martin Otto. Model Theory of Modal Logic. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier, 2008.

[73] Adam Grove. Two Modellings for Theory Change. *Journal of Philosophical Logic*, 17:157–170, 1988.

[74] Peter Gärdenfors. Conditionals and Changes of Belief. In R. Niiniluoto and R. Tuomela, editors, *Acta Philosophica Fennica*, volume 30, pages 381–404, 1982.

[75] Peter Gärdenfors. Rules for Rational Changes of Beliefs. In T. Pauli, editor, *320311: Philosophical essays dedicated to Lennart Åqvist on his fiftieth birthday*, Philosophical Studies no. 34, pages 88–101, 1982. Uppsala: Department of Philosophy, University of Uppsala.

[76] Peter Gärdenfors. *Knowledge in Flux*. The MIT Press, 1988.

[77] Peter Gärdenfors and David Makinson. Revisions of Knowledge Systems using Epistemic Entrenchment. In M. Vardi, editor, *Proceedings of the second conference on Theoretical Aspects of Reasoning about Knowledge (TARK'88)*, pages 83–95. Morgan Kaufmann, 1988.

[78] Joseph Y. Halpern. *Reasonong about Uncertainty*. MIT Press, 2003.

[79] Joseph Y. Halpern and Yoram Moses. Knowledge and Common Knowledge in a Distributed Environment. *Journal of the ACM*, 37(3):549–587, 1990.

[80] Richard W. Hamming. Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 29(2):147–160, 1950.

[81] Jens Ulrik Hansen. *A logic toolbox for modeling knowledge and information in multi-agent systems and social epistemology*. PhD thesis, Roskilde University, 2011.

[82] Pelle G. Hansen and Vincent F. Hendricks. *Infostorms*. Copernicus, 2nd edition, 2016.

[83] Pelle G. Hansen, Vincent F. Hendricks, and Rasmus K. Rendsvig. Infostorms. *Metaphilosophy*, 44(3):301–326, 2013.

[84] Sven-Ove Hansson. Belief Contraction without Recovery. *Studia Logica*, 50:251–260, 1991.

[85] Boris Hasselblatt and Anatole Katok. Principal Structures. In *Handbook of Dynamical Systems, Vol. 1A*. Elsevier, 2002.

[86] Vincent Hendricks and John Symons. Epistemic Logic. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2015 edition, 2015.

[87] Vincent F. Hendricks. *The Convergence of Scientific Knowledge*. Studia Logical Library – Trends in Logic 9. Kluwer, 2001.

[88] Vincent F. Hendricks. *Mainstream and Formal Epistemology*. Cambridge University Press, 2006.

[89] Vincent F. Hendricks and Rasmus K. Rendsvig. Hintikka's *Knowledge and Belief* in Flux. In Hans van Ditmarsch and Gabriel Sandu, editors, *Jaakko Hintikka on Knowledge and Game Theoretical Semantics*, Outstanding Contributions to Logic, pages 317–337. Springer, 2018.

[90] Vincent F. Hendricks and Olivier Roy, editors. *Epistemic Logic: 5 Questions*. Automatic Press / VIP, 2010.

[91] Jaakko Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. College Publications, 2nd (2005) edition, 1962.

[92] Jaakko Hintikka. Impossible Possible Worlds Vindicated. *Journal of Philosophical Logic*, 4:475–484, 1975.

[93] Jaakko Hintikka. Epistemology without Knowledge and without Belief. In *Socratic Epistemology*. Cambridge University Press, 2007.

[94] Wesley H. Holliday. Epistemic Closure and Epistemic Logic I: Relevant Alternatives and Subjunctivism. *Journal of Philosophical Logic*, 44:1–62.

[95] Wesley H. Holliday. *Knowing what Follows: Epistemic Closure and Epistemic Logic*. PhD thesis, Stanford University, 2012.

[96] Wesley H. Holliday and Thomas F. Icard, III. Moorean Phenomena in Epistemic Logic. In L. Beklemishev, V. Goranko, and V. Shehtman, editors, *Advances in Modal Logic (AiML), vol. 8*, pages 178–199. College Publications, 2010.

[97] Hirofumi Katsuno and Alberto O. Mendelzon. On the Difference Between Updating a Knowledge Base and Revising It. In *Proceedings of the Second International Conference on the Principles of Knowledge Representation and Reasoning (KR'91)*, pages 387–394. Morgan Kaufmann, 1991.

[98] Kevin T. Kelly. *The Logic of Reliable Inquiry*. Oxford University Press, 1996.

[99] Kevin T. Kelly and Hanti Lin. Propositional Reasoning that Tracks Probabilisic Reasoning. *Journal of Philosophical Logic*, 41:957–981, 2012.

[100] Dominik Klein. *Social interaction: A formal exploration*. PhD thesis, Tilburg University, 2015.

[101] Barteld P. Kooi. Probabilistic Dynamic Epistemic Logic. *Journal of Logic, Language and Information*, 12:381–408, 2003.

[102] Barteld P. Kooi and Bryan Renne. Generalized Arrow Update Logic. In Krzysztof R. Apt, editor, *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Thirteenth Conference (TARK 2011)*, pages 205–211. ACM, 2011.

[103] Philip Kremer. Temporal Logic over S4: an Axiomatizable Fragment of Dynamic Topological Logic. *Bulleting of Symbolic Logic*, 3:375–376, 1997.

[104] Philip Kremer and Grigori Mints. Dynamical Topological Logic. *Bulleting of Symbolic Logic*, 3:371–372, 1997.

[105] Philip Kremer and Grigori Mints. Dynamic Topological Logic. In M. Aiello, I. Pratt-Hartmann, and J. van Benthem, editors, *Handbook of Spatial Logics*, chapter 10, pages 565–606. 2007.

[106] Philip Kremer, Grigori Mints, and Vladimir Rybakov. Axiomatizing the Next-Interior Fragment of Dynamic Topological Logic. *Bulleting of Symbolic Logic*, 3:376–377, 1997.

[107] Saul A. Kripke. Semantic Analysis of Modal Logic I. Normal Modal Propositional Calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9:67–96, 1963.

[108] Louwe B. Kuijer. How Arbitrary Are Arbitrary Public Announcements? In M. Colinet, S. Katrenko, and R. K. Rendsvig, editors, *Pristine Perspectives on Logic, Language, and Computation*, volume 8607 of *Lecture Notes in Computer Science*, pages 109–123. Springer, 2014.

[109] Bibb Latané and John M. Darley. *The Unresponsive Bystander: Why Doesn't He Help?* Appleten-Century-Crofts, 1970.

[110] Keith Lehrer. *Theory of Knowledge*. Routledge, 1990.

[111] Keith Lehrer and Thomas Paxton Jr. Knowledge: Undefeated Justified True Belief. *Journal of Philosophy*, 66:225–237, 1969.

[112] Hannes Leitgeb. *The Stability of Belief: How Rational Beliefs Coheres with Probability*. Oxford University Press, 2017.

[113] Wolfgang Lenzen. Recent Work in Epistemic Logic. *Acta Philosophica Fennica*, 30:1–219, 1978.

[114] Hector J. Levesque. A logic of Implicit and Explicit Belief. *Proc. National Conference on Artificial Intelligence (AAAI '84)*, pages 198–202, 1984.

[115] Hector J. Levesque. Foundations of a functional approach to knowledge representation. *Artificial Intelligence*, 23:155–212, 1984.

[116] David Lewis. *Convention: A philosophical study*. Harvard University Press, 1969.

[117] David Lewis. *Counterfactuals*. Cambridge University Press, 1973.

[118] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995.

[119] Fenrong Liu. *Changing for the Better: Preference Dynamics and Agent Diversity*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2007.

[120] Fenrong Liu, Jeremy Seligman, and Patrick Girard. Logical Dynamics of Belief Change in the Community. *Synthese*, 191(11):2403–2431, 2014.

[121] David Makinson. How to Give Up: a Survey of Some Formal Aspects of the Logic of Theory Change. *Synthese*, 62:347–363, 1985.

[122] John McCarthy, Masahiko Sato, Takeshi Hayashi, and Shigeru Igarashi. On the model theory of knowledge. Technical report, No. STAN-CS-78-657, Stanford Artificial Intelligence Laboratory, Computer Science Department, Standford University, 1978.

[123] Lawrence S. Moss. Dynamic Epistemic Logic. In H. van Ditmarsch, J. Y. Halpern, W. van der Hoek, and B. P. Kooi, editors, *Handbook of Epistemic Logic*. College Publications, 2015.

[124] David G. Myers. *Social Psychology*. McGraw-Hill, 11 edition, 2012.

[125] Abhaya C. Nayak. Iterated Belief Change Based on Epistemic Entrenchment. *Erkenntnis*, 41(3):353–390, Nov 1994.

[126] Robert Nozick. *Philosophical Investigations*. Harvard University Press, 1981.

[127] Martin J. Osbourne and Ariel Rubinstein. *A Course in Game theory*. The MIT Press, 1994.

[128] Aybüke Özgün. *Evidence in Epistemic Logic: A Topological Perspective*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2017.

[129] Eric Pacuit and Olivier Roy. A Dynamic Analysis of Interactive Rationality. In H. van Ditmarsch, J. Lang, and S. Ju, editors, *Logic, Rationality, and Interaction (LORI 2011, Guangzhou)*, Lecture Notes in Computer Science, Vol. 6953, pages 244–257. Springer, 2011.

[130] Rohit Parikh and Ramaswamy Ramanujam. A Knowledge Based Semantics of Messages. *Journal of Logic, Language and Information*, 12(4):453–467, Sep 2003.

[131] André Platzer. *Logical Analysis of Hybrid Systems: Proving Theorems for Complex Dynamics*. Springer, 2010.

[132] André Platzer. Logics of Dynamical Systems. In *Proceedings of the 27th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2012)*, pages 13–24. IEEE, 2012.

[133] Jan A. Plaza. Logics of Public Communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.

[134] Veikko Rantala. Urn Models: A New Kind of Non-Standard Model for First-Order Logic. *Journal of Symbolic Logic*, 4:455–474, 1975.

[135] Rasmus K. Rendsvig. Aggregated Beliefs and Informational Cascades. In D. Grossi and O. Roy, editors, *Logic, Rationality, and Interaction (LORI 2013, Hangzhou)*, Lecture Notes in Computer Science, pages 337–341. Springer, 2013.

[136] Rasmus K. Rendsvig. Diffusion, Influence and Best-Response Dynamics in Networks: An Action Model Approach. In Ronald de Haan, editor, *Proceedings of the ESSLLI 2014 Student Session*, pages 63–75. arXiv:1708.01477, 2014.

[137] Hans Rott. On the Logic of Theory Change: More Maps between Different Kinds of Contraction Functions. In P. Gärdenfors, editor, *Belief Revision*, Cambridge Tracts in Theoretical Computer Science, pages 122–141. Cambridge University Press, 1992.

[138] Hans Rott and Maurice Pagnucco. Severe Withdrawal (and Recovery). *Journal of Philosophical Logic*, pages 501–547, 1999.

[139] Hans Rott and Mary-Anne Williams. Introduction. In H. Rott and M. Williams, editors, *Frontiers in Belief Revision*. Springer, 2001.

[140] Olivier Roy. Epistemic Logic and the Foundations of Decision and Game Theory. In J. van Benthem and A. Gupta, editors, *Logic and Philosophy Today (Vol.2)*, number 2, pages 301–332. College Publications, 2011.

[141] Tomasz Sadzik. Exploring the Iterated Update Universe. *ILLC Report PP-2006-263*, 2006.

[142] Sanchit Saraf and Sumit Sourabh. Characterizing Successful Formulas: The Multi-Agent Case. arXiv:1209.0935, 2012.

[143] Darko Sarenac. Modal Logic for Qualitative Dynamics. In O. Roy, P. Girard, and M. Marion, editors, *Dynamic Formal Epistemology*, volume 351 of *Synthese Library*, pages 75–101. Springer, 2011.

[144] Jeremy Seligman, Fenrong Liu, and Patrick Girard. Logic in the Community. In M. Banerjee and A. Seth, editors, *Logic and Its Applications*, pages 178–188. Springer, 2011.

[145] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[146] Robert Stalnaker. A Theory of Conditionals. *Studies in Logical Theory, American Philosophical Quarterly*, Monograph: 2:98–112, 1968.

[147] Robert Stalnaker. Knowledge, Belief and Counterfactual Reasoning in Games. *Economics and Philosophy*, 12:133–163, 1996.

[148] Robert Stalnaker. On Logics of Knowledge and Belief. *Philosophical Studies*, 128:169–199, 2006.

[149] William Talbott. Bayesian Epistemology. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

[150] Georg Henrik von Wright. *An Essay in Modal Logic*. North-Holland Publishing Company, 1951.

[151] Fernando Raymundo Velazquez-Quesada. *Small steps in dynamics of information*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2011.

[152] Jan de Vries. *Topological Dynamical Systems*. de Gruyter, 2014.

[153] Xuefeng Wen, Hu Liu, and Fan Huang. An Alternative Logic for Knowability. In H. van Ditmarsch, J. Lang, and S. Ju, editors, *Logic, Rationality, and Interaction*, volume 6953 of *Lecture Notes in Computer Science*, pages 342–355. Springer, 2011.

[154] Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, 2010.

# Papers

# Paper 1

# Pluralistic ignorance in the bystander effect: informational dynamics of unresponsive witnesses in situations calling for intervention

**Rasmus K. Rendsvig**

**Abstract** The goal of the present paper is to construct a formal explication of the pluralistic ignorance explanation of the bystander effect. The social dynamics leading to inaction is presented, decomposed, and modeled using dynamic epistemic logic augmented with 'transition rules' able to characterize agent behavior. Three agent types are defined: First Responders who intervene given belief of accident; City Dwellers, capturing 'apathetic urban residents' and Hesitators, who observe others when in doubt, basing subsequent decision on social proof. It is shown how groups of the latter may end in a state of pluralistic ignorance leading to inaction. Sequential models for each agent type are specified, and their results compared to empirical studies. It is concluded that only the Hesitator model produces reasonable results.

## 1 Introduction: the bystander effect and pluralistic ignorance

On March 13, 1964, in Queens, New York, Catherine Susan "Kitty" Genovese was raped and stabbed, the assailant fleeing multiple times during the ongoing assault that resulted in Genovese's death. Multiple residents witnessed parts of the nearly one hour long attack, without successfully intervening.

The foremost explanation put forth in the ensuing media coverage was *apathy* among urban citizens. Through the pressures of city life, "*homo urbanis*" has lost his sense of empathy for fellow man, all but grown indifferent to their quarrels (Latané and

R. K. Rendsvig (✉)
Department of Media, Cognition and Communication, University of Copenhagen,
Copenhagen, Denmark
e-mail: rendsvig@gmail.com

Darley 1968, 1970). This explanation struck social psychologists John M. Darley and Bibb Latané as incorrect, whereupon they set out to provide an alternate explanation. The first publication of experimental results is the classic (Latané and Darley 1968), which was soon followed by a vast amount of studies (see Latané and Nida 1981 for a review). In these studies, focus has been moved from why *urban citizens* do not help, to why *people in groups* are less prone to help. The collected data shows a robust tendency, namely that the chance of help being offered diminishes as the number of witnesses increases. This tendency will be referred to as *the bystander effect*.

## 1.1 An explanation of the intervention process

One currently used text-book explanation (see e.g. Myers 2012) of the bystander effect stems from Latané and Darley (1970), and involves three steps which each bystander must go through before he or she will intervene.[1] First, a bystander must *notice the event* in question. With busy street life involving traffic and pedestrians, the risk of overlooking a seizure is higher than on desolate streets. Where the problematic situation goes unnoticed, help will not be offered. Second, if noticed, the bystander must *interpret the event* in order to decide whether an emergency is occurring or not. In many cases, this will not prove to be a problem: situations involving car accidents or bleeding victims are seldom epistemically ambiguous. However, a man slumped on a bench may provide an epistemic conundrum, as he may be merely mumbling curses against the general youth, marooned following a too *Saké* intense business lunch, or moaning in pain from the onset of a seizure. Such ambiguities may be sought resolved by the acquisition of further information, readily present in the form of *social proof*: if other bystanders are not showing signs of distress, the event will be perceived as less critical and therefore ultimately bypassed. Third, in case the event is interpreted as requiring help, the bystander has to gauge *whether to take responsibility*: when alone, there is no question as to who should intervene, but when gathered in groups, *diffusion of responsibility* may arise. Such diffusion may be caused by uncertainty as to whether we are among the best qualified to handle the situation, whether others have already called for paramedics or are just about to act. When alone, the responsibility to intervene rests on one individual, but when in a group, the same pressure is apparently distributed among all, thereby diminishing the chances that anyone will act.

## 1.2 Pluralistic ignorance

The goal of this paper is to model the social informational dynamics and decision procedures running the second of these steps, specifying conditions under which a group of agents in an ambiguous situation may choose to seek social proof in order to individually determine a correct course of action and the associated consequences thereof, hereby providing a detailed explanation for (this part of) the bystander effect. This narrower focus is taken as the second step of the bystander effect explanation

---

[1] For a comprehensive walk-through of this explanation and supportive data, refer to Latané and Darley (1970).

constitutes an interesting informational dynamics in its own right, useful to the analysis of social situations in which neither distractions nor diffusion play important roles.

The second step of the dynamics revolve around a *belief state* often referred to as one of *pluralistic ignorance*: a situation in which everybody believes that everybody else believes a given proposition/endorses a given norm, while no-one in fact believes it/endorses it.[2] Pluralistic ignorance has been put forth as a decisive factor in a plethora of social situations, including the introduction of various unpopular norms such as college binge drinking (Prentice and Miller 1993) and violent gang behavior (Bicchieri and Fukui 1999), the persistence of poor strategies in light of poor firm performance (Westphal and Bednar 2005) and lack of help seeking in class rooms (Miller and McFarland 1987); the allowance of ongoing mortgage deed merry-go-rounds in Denmark during the financial 'upswing' of 2007 (Hendricks and Rasmussen 2012; Hansen et al. 2013).

Pluralistic ignorance may cause individuals not to act for more than one reason. One may be due to social inhibition—you may not wish to be the only one raising your hand to ask a question. Here, inaction results from vanity and social identity. There may be *doubt*—you may not wish to call the police if there is no cause for alarm. In case of doubt, inaction follows from incorrect information processing. Both causes may individually lead to inaction and may further co-occur. In the following, 'pluralistic ignorance' will be used to refer only to processes of the latter kind.

## 1.3 Structure of the paper

In Sect. 2, an example of a social dynamics involving pluralistic ignorance which lead to unfortunate inaction is presented, and an informal sketch of the information processing involved is outlined, the structure of which is used as a guideline for the formal representation. In Sect. 3, elements from dynamic epistemic logic (DEL) are presented. DEL is the primary modeling tool, used to represent static epistemic states and belief revision in light of new information. Section 3 also presents the modeling of the occurrence of 'the accident'. In Sect. 4, it is shown how *agent types* may be defined for DEL, hereby augmenting the framework with a notion of *choice agency*. Three agent types relevant to the bystander effect are defined: First Responders, a 'good samaritan' type agent, who will choose to intervene in emergencies if she believes one such is occurring; City Dwellers, capturing the 'apathetic urban resident' and Hesitators, who observes others when in doubt. In Sect. 5, it is shown how Hesitators' *misinterpretation* of other Hesitators' choice to observe may lead to a state of pluralistic ignorance. Coupled with Hesitators basing subsequent action on social proof obtained through observation, it is further shown that this agent type will choose to evade the scene of the accident, though they privately believe there is cause for intervention. In Sect. 6, models run with each of the three agent types are compared to empirical studies. It is concluded that the Hesitator model is the only of the three producing reasonable results. In Sect. 7, we conclude.

---

[2] Halbesleben and Buckley (2004) provides an illuminating overview of the history and development of the term.

## 2 From ambiguity to inaction

Consider the following example, inspired by Halbesleben and Buckley (2004). A firm is performing poorly and this is caused by the currently implemented business strategy. Every member of the board of directors has equal access to the relevant information pertaining to strategic choice and firm performance, which to each member strongly indicates that the prudent action is to change strategy. The available information is, however, not conclusive: the possibility that a strategy change will leave the firm performance-wise worse off cannot be ruled out. Hence, every board member is in an epistemically ambiguous situation, where they privately believe intervention will be fruitful, but all would prefer additional information before settling for a vote against the status quo. As the voting situation arises, all seek such further information from their peers, hoping that the votes of others will illuminate them. As all look to each other, nobody initially raises an objection, which is interpreted by others as the choice that no objection need be raised. Hereby, all conclude that their intelligent peers believe that the status quo should not be changed. This perceived consensus among peers is then seen as providing evidence to the conclusion that the currently implemented strategy is in fact desirable. Having thus reflected, board members choose not to intervene in the status quo, and the low firm performance continues another term or two. Though a fictitious example, pluralistic ignorance does occur on corporate boards, *often resulting in poor strategic decisions* (see Westphal and Bednar 2005).

While the example differs in topic from a street accident, the two share common information dynamics. The street incident revolves around the commonly unwanted case of physical pain, the board meeting example focuses on the commonly unwanted event of a sub-optimal status quo strategy. Both involve uncertainty about whether an unwanted situation is the case or not, both require intervention in case the unwanted situation indeed is occurring, and both allow for the gathering of further information by observing peers.

Notice how the example implicitly utilizes *two* instances of pluralistic ignorance. First, there is one instance of what may be called *norm-based* pluralistic ignorance: though every member uses the decision rule "if in doubt, seek further information", they assume that others will follow a different rule, namely "if in doubt, raise an objection". This is an instance of pluralistic ignorance as everybody believes that everybody else follows a given norm (here, a decision rule), while in fact no-one follows it. The second instance of pluralistic ignorance is *proposition-based*: everybody ultimately believes that everybody else believes that status quo is fine, while everybody privately believes the strategy should be changed.

The dynamics involved in the example may be decomposed into eleven elements (see Fig. 1): six static states and five state-altering transitions. To start from scratch, in the first state, nothing has happened (1). This is followed by the occurrence of an event, epistemically ambiguous between being an accident (stabbing, onset of poor performance) or nothing of consequence (dispute, reasonable performance) (2). This event results in a second state, where everybody privately believes that an accident occurred, while remaining ignorant about the beliefs of others (3). Based on this state, one may choose to intervene (rush to help, object to strategic choice), may choose actively to evade (ignore the stabbing, withhold objection), or may choose to seek
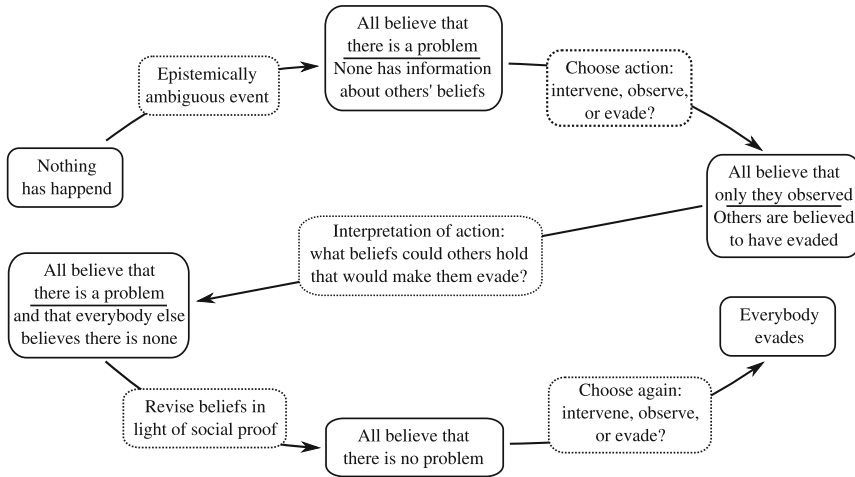
**Fig. 1** Flowchart of the dynamics leading to the bystander effect. *Boxes with solid lines* represent epistemic states—*boxes with dotted lines* indicate events

further information by observing the choice made by others. Crucially, the performed actions of evasion and observation are here considered to be epistemically ambiguous: in seeking further information, we do not want to flaunt our ignorance, so further observation is made discreetly. Given the ambiguity of the accident, observation is chosen and executed (4). It is claimed here that a crucial further element for the dynamics is the resulting *mis-perception* of this choice when made by others: though we ourselves may choose to observe, when we see others do the same, we consider it plausible that they in fact chose to evade. Given this norm-based pluralistic ignorance, in the ensuing third state (5), all still believe there was an accident, while believing that all others evaded. To obtain information about the beliefs of others, their perceived actions must now be *interpreted* (6): given that you evaded, what may I conclude about your beliefs pertaining to the accident? Under the assumption that you are a reasonably decent person, only that you believed there was none. Such interpretations conducted by all then results in a fourth state (7) of proposition-based pluralistic ignorance: though we all believe there was an accident, we also believe that no-one else believes so. Revising our beliefs in the light of the obtained social proof (8), all conclude that no accident occurred (9). Given a further chance to act (10), evasion will be the natural choice, leading to the final state (11), where the accident in fact occurred, everybody believed so, but nevertheless chose to evade it, due to the social information dynamics.

## 3 Plausibility models for states and actions

The sketch presented above suggests several ingredients required for a suitable model, including propositional- and higher-order beliefs (beliefs about beliefs), belief change
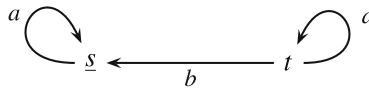
**Fig. 2** A simple, two agent plausibility frame, denoted $\mathsf{F}_1$

in light of new information and agent action. To that end, dynamic epistemic logic[3] with updating by action models with postconditions is a suitable framework: All higher-order beliefs are specified in relatively small models, factual change may be modeled using postconditions, and the dynamics may be built step-by-step, allowing for a detailed overview of each step. Step-by-step construction allows for easy replacement of single 'modules', whereby alternative runs may be investigated. For each such run, the dynamics terminate when either someone agent intervenes, or all agents choose to evade. The dynamics presented concern a three agent case, with group size variations presented in Sect. 6. Throughout, the same complete graph 'all see all' social network structure is assumed.

## 3.1 Statics

*Multi-agent plausibility frames* Where $\mathcal{A}$ is a finite set of *agents*, a *multi-agent plausibility frame* (*MPF*) is a structure $\mathsf{S} = (S, \leq_i)_{i \in \mathcal{A}}$ where $S$ is a finite set of *states* and each $\leq_i$ is a well-preorder.[4]

The idea behind plausibility frames is that such encode the knowledge and beliefs of a group of agents, $\mathcal{A}$, capturing which states each agent may tell apart, and how plausible these states are relative to one another. If two states $s, t$ are connected by $\leq_i$, then $i$ cannot tell these states apart, but if $s <_i t$ (i.e. $s \leq_i t$ and $t \not\leq_i s$), then $i$ considers $s$ *more plausible* than $t$.[5] Figure 2 illustrates a simple plausibility frame $\mathsf{F}_1$ with two states, $s$ and $t$, and two agents, $a$ and $b$. The arrow from $t$ to $s$ captures that $s <_b t$, i.e. that $b$ cannot distinguish between $s$ and $t$, but finds $s$ strictly more plausible.[6] Reflexive arrows will only be drawn if they are the only arrows for a given agent. In Fig. 2, $a$ cannot tell $s$ from $s$ nor $t$ from $t$.

### 3.1.1 Indistinguishability relation: information and plausibility cells

Given an MPF $\mathsf{S} = (S, \leq_i)_{i \in \mathcal{A}}$, the *indistinguishability relation* for agent $i$ is the equivalence relation $\sim_i := \leq_i \cup \geq_i$. Further, the *information cell* of agent $i$ at state $s$ is $\mathcal{K}_i[s] = \{t : s \sim_i t\}$ and the *plausibility cell* of agent $i$ at state $s$ is $\mathcal{B}_i[s] = Min_{\leq_i} \mathcal{K}_i[s] = \{t \in \mathcal{K}_i[s] : t \leq_i s', \text{ for all } s' \in \mathcal{K}_i[s]\}$. The plausibility cell $\mathcal{B}_i[s]$

---

[3] Technically, no *logic* is introduced; the dynamics are investigated using only model theory.

[4] Reflexive and transitive binary relation where every non-empty subset has a minimal element, cf. (Baltag and Smets 2008).

[5] The relation $\leq_i$ may therefore more appropriately be thought of as an *implausibility* relation, where $s \leq_i t$ is read '$t$ is as implausible as $s$, or more so'.

[6] A heuristic to aid recall is that $s < t$ in form is similar to $s \leftarrow t$ when looking at the arrowhead. When $s \leq_i t$ and $t \leq_i s$, arrowheads are omitted altogether.

contains the worlds the agent find *most plausible* from the information cell $\mathcal{K}_i[s]$ and represent the "doxastic appearance" (Baltag and Smets 2008, p. 25) of $s$ to $i$.[7] Notice that $s \leq_i t$ means that $s$ is at *least as plausible* as $t$ for $i$.

In Fig. 2, it is the case that $s \sim_a s$ and $t \sim_a t$, but not $s \sim_a t$. Hence $\mathcal{K}_a[s] = \{s\}$ and $\mathcal{K}_a[t] = \{t\}$. For agent $b$, $s \sim_b s$, $t \sim_b t$ and $s \sim_b t$, entailing that $\mathcal{K}_b[s] = \mathcal{K}_b[t] = \{s, t\}$. However, even if the actual state is $t$, agent $b$ finds $s$ more (most) plausible, so it is the sole element in $b$'s plausibility cell at $t$: $\mathcal{B}_b[t] = \{s\}$.

### 3.1.2 Doxastic propositions

When considering plausibility frames, sets of states may be identified with propositions; e.g. the set $\{s\}$ could be identified with the proposition that intervention is desirable, and $\{t\}$ with the same proposition's negation. In this case, Fig. 1 would represent a situation in which 2 finds it more plausible that intervention is desirable than that it is not, while 1 would know whether or not this was the case.

More specifically, let a *doxastic proposition* (henceforth just *proposition*) be a map $P$ that assigns to every MPF $\mathsf{S}$ with state-space $S$ a subset $(P)_\mathsf{S} \subseteq S$.[8] Denote *true*, $\top$, *false*, $\bot$, and Boolean operations for arbitrary propositions $P$ and $Q$ by

$$
\begin{array}{l|l}
(\top)_\mathsf{S} := S & (P \wedge Q)_\mathsf{S} := P_\mathsf{S} \cap Q_\mathsf{S} \\
(\bot)_\mathsf{S} := \emptyset & (P \vee Q)_\mathsf{S} := P_\mathsf{S} \cup Q_\mathsf{S} \\
(\neg P)_\mathsf{S} := S \backslash P_\mathsf{S} & (P \rightarrow Q)_\mathsf{S} := (S \backslash P_\mathsf{S}) \cup Q_\mathsf{S}
\end{array}
$$

Propositions with epistemic and doxastic modalities are given by

$$(K_i P)_\mathsf{S} := \{s \in S : \mathcal{K}_i[s] \subseteq P_\mathsf{S}\}$$
$$(B_i P)_\mathsf{S} := \{s \in S : \mathcal{B}_i[s] \subseteq P_\mathsf{S}\}$$

The Boolean case simply follows the immediate set-theoretic interpretation. Propositions with epistemic and doxastic operators represent statements of knowledge and belief: $K_i P$ / $B_i P$ reads 'agent $i$ knows / believes that $P$', and $(K_i P)_\mathsf{S}$ / $(B_i P)_\mathsf{S}$ are the sets of states of these propositions given an MPF $\mathsf{S}$. For knowledge, the definition entails that $s \in (K_i P)_\mathsf{S}$ iff for all $t$ in $i$'s information cell relative to $s$, $t$ is a $P$-state. Belief has the same reading, but restricted to the plausibility cell for $i$ at $s$.[9]

### 3.1.3 Example: King or Queen?

Assume a situation with two players, $a$ and $b$, where $a$ has one card in hand, being either a King or a Queen. Let $Q$ be the doxastic proposition "$a$ has a Queen on hand"

---

[7] The notation for information and plausibility cells are adopted from Dégremont (2010).

[8] Parentheses will be omitted where no confusion should arise.

[9] The reason for using this atypical definition of propositions is that it allows us to speak about the *same* proposition across multiple models. This is practical as model transformations will play a large role in the latter. A Kripke model valuation may easily be extracted from set of doxastic atomic propositions; see (Baltag and Smets 2008) for details.

and $K$ ditto for King. Over $\mathsf{F}_1$, Fig. 1, set $(Q)_{\mathsf{F}_1} = \{s\}$ and $(K)_{\mathsf{F}_1} = \{t\}$ (hence $(\neg Q)_{\mathsf{F}_1} = (K)_{\mathsf{F}_1}$). In this case, $(K_a Q)_{\mathsf{F}_1} = \{s\}$, $(K_b Q) = \emptyset$ and $(B_b Q)_{\mathsf{F}_1} = \{s, t\}$. That is, in state $s$, agent $a$ knows $Q$, in *no* state does $b$ know $Q$, but $b$ believes $Q$ in both states. The structure thus models a situation in which $a$ knows whether she holds a Queen, while $b$ only has a belief this regarding. Whether this belief is correct depends on which of the two states is the *actual*. Combined with such a valuation of propositions, a plausibility frame is called a *model*. Denote $\mathsf{F}_1$ with the described valuation $\mathbf{S}_{\mathbf{Q/K}}$.

### 3.1.4 Epistemic plausibility models

Let a *valuation set* be a set $\Phi$ of doxastic propositions, considered the *atomic* propositions. An *epistemic plausibility model* (*EPM*) is an MP frame together with a valuation set $\Phi$, denoted $\mathbf{S} = (S, \leq_i, \Phi)_{i \in \mathcal{A}}$. For $s \in P_\mathbf{S}$, write $\mathbf{S}, s \models P$, and say that $P$ is *true* or *satisfied* at state $s$ in model $\mathbf{S}$. A *pointed EPM* $\mathbf{S} = (S, \leq_i, \Phi, s_0)_{i \in \mathcal{A}}$ is an EPM with a designated state $s_0 \in S$, called the *actual state*. Where $s_0 \in P_\mathbf{S}$, write $\mathbf{S} \models P$.

### 3.1.5 EPMs and Kripke models

Where epistemic plausibility frames are special instances of *Kripke frames* (see e.g. Blackburn et al. 2001), epistemic plausibility models are not special instances of *Kripke models*.[10] However, every EPM $\mathbf{S}$ gives rise to a Kripke model $M_K$. First, let $\Phi'$ be $\Phi$ where the functional nature of each doxastic proposition is ignored. $\Phi'$ may then be treated as a set of *atomic proposition symbols*. Second, define a *valuation map* $\|\cdot\| : \Phi' \longrightarrow \mathcal{P}(S)$, assigning to the elements of $\Phi'$ a set of states from the state-space of the underlying frame. Simply let $M_K$ and $\mathbf{S}$ be based on the same frame, and let the valuation map $\|\cdot\|$ for $M_K$ be given by $\|P\| := P_\mathbf{S}$, for all $P \in \Phi$. The alternative definition of EPMs is used as it is natural when dealing with doxastic propositions rather than a syntactically specified language.

### 3.1.6 Relevant propositions and the initial state

To model the second step of the bystander effect dynamics for three agents, ten atomic propositions are required. First, use $A$ to denote that an accident has occurred. This is the basic fact about which the agents must establish a belief. Second, each agent $i \in \mathcal{A} = \{a, b, c\}$ must choose one of three actions: either to intervene, $I_i$, to observe, $O_i$, or to evade the scene, $E_i$. The set of these atoms is denoted $\Phi$. As the model constructed is temporally simple, the propositions are best read as "agent $i$ has intervened / observed / evaded". It is assumed that no agent can perform two actions simultaneously, i.e. that $I_i \cap O_i = I_i \cap E_i = O_i \cap E_i = \emptyset$. Denote the set of all doxastic propositions obtainable from $\Phi$ and the above construction rules by $\mathsf{Prop}_\Phi$.

The initial state (where nothing has happened) may now be represented by the EPM $\mathbf{S_0}$, Fig. 3.

---

[10] As opposed to the terminology of Baltag and Smets (2008), Benthem (2007), Demey (2011) where epistemic plausibility models are Kripke models.

**Fig. 3** The initial state where nothing has happened; all atoms are set to false: $(A)_{\mathbf{S_0}} = (I_i)_{\mathbf{S_0}} = (O_i)_{\mathbf{S_0}} = (E_i)_{\mathbf{S_0}} = \emptyset$

In this simple state, every agents knows *exactly* what has transpired so far: nothing. All propositions are false, everybody knows this, which is again known by all, etc. Among others, it is the case that $\mathbf{S_0}, s_0 \models \neg(A \vee I_1 \vee O_2 \vee E_3) \wedge \bigwedge_{i \in \mathcal{A}}(K_i(\neg A \wedge \bigwedge_{j \in \mathcal{A}} K_j \neg A))$.

The end conditions of runs mentioned in the beginning of Sect. 3 may now formally be specified: identify the *end of a run* with any EPM satisfying at its actual state either $\bigvee_{i \in \mathcal{A}} I_i$ or $\bigwedge_{i \in \mathcal{A}} E_i$, capturing respectively that at least one agent intervenes, or all evade.

### 3.2 Changing models: action models and action-priority update.

To capture factual and informational changes that occur due to events, a static epistemic plausibility model may be transformed using an *action model*, capturing the factual and epistemic representation of the event, and the *action-priority update product*. The guiding idea is that an action model encodes the belief and knowledge agents have about an *ongoing* event, the information from which is combined with the static model by taking the two models' *product*: the result is a new static model in which the agents' new information takes priority over that of the previous static model. The present formulation rests on Baltag and Smets (2008)), with the addition of *postconditions*, as used in Ditmarsch and Kooi (2008), Bolander and Birkegaard (2011). The latter allows action models to not only change the knowledge and belief of agents, but also effectuate *ontic*[11] *changes*, needed when the environment or agents perform actions.

#### 3.2.1 Action plausibility models

A *(pointed) action plausibility model (APM)*

$$\mathbf{E} = (\Sigma, \leq_i, pre, post, \sigma_0)_{i \in \mathcal{A}}$$

is an MP frame $(\Sigma, \preceq_i)_{i \in \mathcal{A}}$ augmented with a *precondition map*, $pre : \Sigma \longrightarrow \mathsf{Prop}_\Phi$ and a *postcondition map* $post : \Sigma \longrightarrow \mathsf{Prop}_\Phi$ such that $post(\sigma) = \psi$ where $\psi \in \{\top, \bot\}$ or $\psi = \bigwedge_{i=1}^{n} \varphi_i$ with $\varphi_i \in \{P, \neg P : P \in \Phi\}$. Finally, $\sigma_0 \in \Sigma$ is *the actual event*.

Just as every world in an EPM represents a possible state of affairs, specified by the world's true propositions, so every action in an APM represents a possible *change*. *What* change is specified by the pre- and postconditions; preconditions determine what is required for the given action to take place, i.e. what conditions a world must satisfy for an action to be executable in that world, and postconditions what *factual* change the action brings about.

---

[11] Ontic facts are all non-doxastic facts, i.e. propositions that do not contain belief or knowledge operators.

### 3.2.2 Example: King or Queen?, cont.

Continuing the example above, let now $a$ play her card face down. Assuming that $a$ knows which cards she is playing, this situation may again be presented by the MPF $\mathsf{F_1}$ in Fig. 1, with $s$ representing the action 'a plays a Queen' and $t$ 'a plays a King'. The precondition for $s$, that $a$ is playing a Queen, is $Q$, that $a$ has a Queen on hand, and *vice versa* for $t$ and $K$. Hence $pre(s) = Q$ and $pre(t) = K$. Following the play, $a$ will either no longer have a Queen on hand, or no longer have a King. Hence, $post(s) = \neg Q$ and $post(t) = \neg K$. With these pre- and postconditions, $\mathsf{F_1}$ is an action model, call it $\mathbf{E_{Q/K}}$, representing the event where $a$ is certain that she is playing a Queen, while $b$ is uninformed about which of the two plays is the actual, finding it more plausible that $a$ plays $Q$.

### 3.2.3 Doxastic programs

Where $\mathbf{E_{Q/K}}$ represents a situation where $a$ plays her card face down, a show-and-tell play by $a$ is captured by the two strict subsets of the model, i.e. by the *doxastic programs* $\Gamma_Q = \{s\}$ and $\Gamma_K = \{t\}$. A doxastic program is the action model equivalent of a proposition, i.e. a subset of all actions in the models' event space: $\Gamma \subseteq \Sigma$. Over $\mathbf{E_{Q/K}}$, the program $\Gamma_Q$ captures the event where $a$ plays Queen and $b$ sees this, and $\Gamma_K$ the same for $a$ playing King. In the ensuing, it will be assumed that doxastic programs contain the actual action.

### 3.2.4 An accident occurs

What is a suitable APM capturing both the factual change that the accident occurs, as well as $a$, $b$ and $c$'s information about this? Given that the accident in fact occurs, it is clear that the actual event $\sigma_0$ of the model must change the truth value of $A$ from false in $\mathbf{S_0}$ to true in the ensuing EPM $\mathbf{S_1}$. Further, no agents perform actions during the event, so $post(\sigma_0) = A$.

Focusing on $a$, then how does she perceive the occurrence of accident? As "[m]ost emergencies are, or at least begin as, ambiguous events" (Latané and Darley 1968, p. 216) $a$ will at least be uncertain regarding whether it occurs or not, and therefore considers an alternative event $\tau_0$ with $post(\tau_0) = \top$ possible.[12] Moreover, *ex hypothesi*, her perception of the event indicates that in fact $A$, so $\sigma_0 \prec_a \tau_0$. How does $a$ perceive that $b$ and $c$ perceive the accident? Not being telepathic, $a$ cannot tell, and she considers it possible that both, neither, or either of $b$ and $c$ perceive the event as she does. It is assumed, though, that $a$ perceives $b$ and $c$ during the event *as forming an opinion about whether or not $A$*. This assumption is made for two reasons: (1) it produces a smaller model, and (2) pertaining to social proof, only agents perceived as informed are interesting from $a$'s point of view. Variations to this assumption would be interesting, but are not dealt with here. Finally, $a$ must consider it possible that $b$

---

[12] The postcondition $\top$ leaves all atomic propositions as they were in the previous model. This is specified by the action priority update product below.
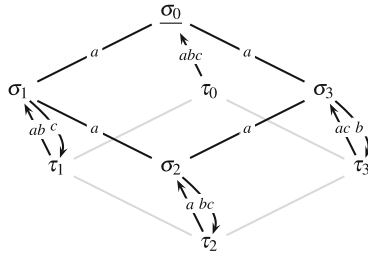
**Fig. 4** Agent $a$'s perception of the accident, $\sigma_0$. All $\sigma_k$ actions have $post(\sigma_k) = A$; for all $\tau_k$, $post(\tau_k) = \top$. Some links are *gray* only for presentation; they mirror the labels above. Diagonal links implied by transitivity are omitted; so are many links for $b$ and $c$ (see Fig. 5)

and $c$ are wrong about the way $a$ perceives the event. Taking this into consideration, Fig. 4 illustrates $a$'s doxastic perception of the accident.

Agent $a$ considers it more plausible that an accident is occuring, but is (at this point) agnostic about the beliefs of her peers; she finds it possible they all 'agree' ($\sigma_0, \tau_0$), that she agrees with only $b$ ($\sigma_1, \tau_1$) or with only $c$ ($\sigma_3, \tau_3$), or that both $b$ and $c$ perceive the ongoing event as non-hazardous ($\sigma_2, \tau_2$). Finally, she cannot rule out that $b$ and $c$ both find the event unproblematic and that they perceive $a$ as doing the same ($\sigma_7, \tau_7$).

Assuming that $b$ and $c$ perceive the accident in an identical manner, the model in Fig. 3 may be suitably duplicated and combined, resulting in the joint model $\mathbf{E_0}$, Fig. 5. Notice that $b$ and $c$'s perceptions are identical to $a$'s. The only states not obtained from a duplication of Fig. 3 are $\sigma_7$ and $\tau_7$. In fact, no one considers these possible, but neither can anyone rule out that others entertain them.

To incorporate the (new) information from an action model or a doxastic program in an EPM, the two must be combined. A natural procedure for doing so is the *action-priority update product* (Baltag and Smets 2008).

### 3.2.5 Action-priority update

The *action-priority update* is a binary operation $\otimes$ with first argument an EPM $\mathbf{S}$ with relations $\leq_i$ and second argument a doxastic program $\Gamma \subseteq \Sigma$ over some APM $\mathbf{E}$ with action space $\Sigma$ and relations $\preceq_i$. The *APU product* is an EPM

$$\mathbf{S} \otimes \Gamma = (S \otimes \Gamma, \leq_i^{\uparrow}, \Phi^{\uparrow}, (s_0, \sigma_0))$$

where the updated state space is $S \otimes \Gamma = \{(s, \sigma) \in S \times \Gamma : \mathbf{S}, s \models pre(\sigma)\}$; each updated pre-order $\leq_i^{\uparrow}$ is given by $(s, \sigma) \leq_i^{\uparrow} (t, \tau)$ iff either $\sigma \prec_i \tau$ and $s \sim_i t$, or else $\sigma \simeq_i \tau$ and $s \leq_i t$;[13] the valuation set $\Phi^{\uparrow}$ is identical to $\Phi$, with the requirement that for every atom $P \in \Phi$,

$$P_{\mathbf{S} \otimes \Gamma} = \{(s, \sigma) : s \in P_{\mathbf{S}} \text{ and } post(\sigma) \not\models \neg P\} \cup \{(s, \sigma) : post(\sigma) \models P\}$$

for states $(s, \sigma) \in S \otimes \Sigma$. Finally, $(s_0, \sigma_0)$ is the new actual world.

---

[13] $\preceq_i$ is from $\mathbf{E}$ and $\leq_i$ from $\mathbf{S}$. $\sigma \prec_i \tau$ denotes ($\sigma \preceq_i \tau$ and not $\sigma \succeq_i \tau$), $\sigma \simeq_i \tau$ denotes ($\sigma \preceq_i \tau$ and $\sigma \succeq_i \tau$).
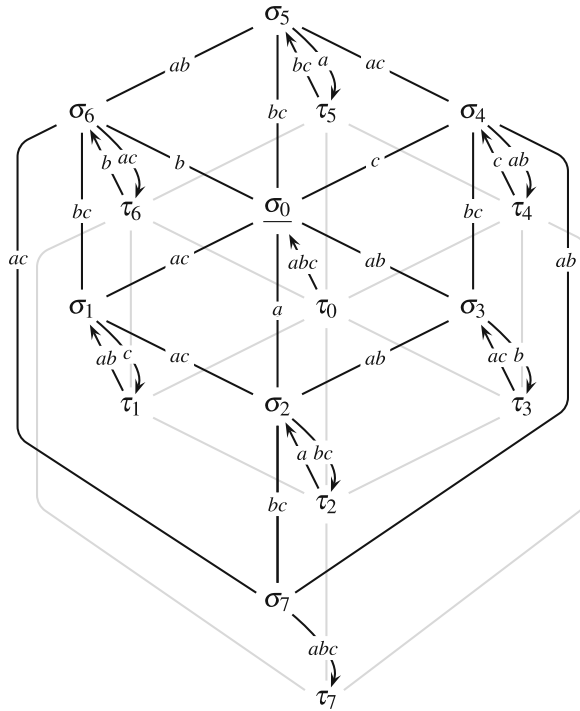
**Fig. 5** The EPM denoted $\mathbf{E_0}$, representing the joint perception of the occurrence of an accident for three agents $a$, $b$ and $c$. All $\sigma_k$ actions have $post(\sigma_k) = A$; for all $\tau_k$, $post(\tau_k) = \top$. *Grey* links are only for presentation; they mirror the labels above. Links implied by transitivity are omitted, e.g. between $\sigma_6$ and $\sigma_4$ for $a$. In essence, $\mathbf{E_0}$ captures (1) that the accident in fact occurs, (2) that this is ambiguous for every agent, who all find it more plausible that it does occur and (3) that no agents learns anything about others' perception of the event, except that *all have an opinion as to whether or not the accident occurred*

The APU product gives priority to new information encoded in $\Gamma$ over the old beliefs from $\mathbf{S}$ by the anti-lexicographic specification of $\leq_i^{\uparrow}$ that gives priority to the APM plausibility relation $\preceq_i$. The definition further clarifies the role of pre- and postconditions; if a world does not satisfy the preconditions of an action, then the given state-action pair does not survive the update, and if postconditions are specified, these override earlier ontic facts, else leave all as was.[14]

*3.2.6 Example, concl.: King or Queen?*

The factual and doxastic consequences of $a$'s play of the Queen is calculated by finding the APU product of $\mathbf{S_{Q/K}}$ and $\mathbf{E_{Q/K}}$. The result is (again) an EPM $\mathbf{S_{Q/K}} \otimes \mathbf{E_{Q/K}}$ with underlying frame $\mathbf{F_1}$, with state space $\{(s, s), (t, t)\}$. The state $(s, s)$ 'survives' as $s$ from $\mathbf{S_{Q/K}}$ satisfies the preconditions of $s$ from $\mathbf{E_{Q/K}}$, while $(s, t)$ does not, as

---

$pre(t) = K$ while $(K)_{\mathbf{S_{Q/K}}} = \{t\}$. Further, in the actual state, $Q$ changes truth value from true in $s \in \mathbf{S_{Q/K}}$ to false in $(s, s) \in \mathbf{S_{Q/K}} \otimes \mathbf{E_{Q/K}}$ as a result of the postconditions of $s \in \mathbf{E_{Q/K}}$. Finally, the information of the agents have changed: e.g., $b$ now *knows* that $a$ does not hold a Queen $((s, s) \models K_b \neg Q)$.

### 3.2.7 Updating with the accident

Updating the simple initial state model $\mathbf{S_0}$ with accident event model $\mathbf{E_0}$ results in a state of great uncertainty. In summary, every agent believes that $A$: an accident has occurred (though no-one *know*s), all know that their peers either believe or disbelieve $A$ (none are indifferent between $A$ and $\neg A$ states), and all consider it possible that *the others* consider it possible that all believe there is no accident.

Formally, updating the simple structure of $\mathbf{S_0}$ with $\mathbf{E_0}$ produces the EPM $\mathbf{S_1} := \mathbf{S_0} \otimes \mathbf{E_0}$ which shares frame with $\mathbf{E_0}$ (Fig. 5) and has $(s_0, \sigma_0)$ as actual state. In $\mathbf{S_1}$, for all $i \in \{0, ..., 7\}$, $(s_0, \sigma_i) \in A_{\mathbf{S_1}}$, and $(s, \tau_i) \in (\neg A)_{\mathbf{S_1}}$. Among others, the following doxastic propositions are true at $(s_0, \sigma_0)$: $A$, $\bigwedge_{i \in \mathcal{A}} B_i A$, $K_a(B_b A \vee B_b \neg A)$, $\neg K_a B_b \neg A$.

Based on this second static state, the agents must make their first decision, as specified by their *decision rules*, which jointly determine the *type* of agent they are.

## 4 Decisions: agent behavior characterized by transition rules

Though agent action may be represented in the introduced DEL framework using suitable atoms and postconditions, the notion of agency in DEL is purely *doxastic*. To move from only *believing* agents to *acting* agents, a richer framework is called for. One possibility would be to introduce a game- or pay-off structure in parallel to the DEL framework or embed the entire dynamics modeled in a temporally extended game tree, whereby actions could be made 'rationally', based on utility maximization at end nodes. A drawback to this method is the large models required: every branch must be fully specified before decisions may follow. Further, considering all possible branches is a cognitively complex task, making the approach empirically unrealistic.

Instead, an alternative approach involves utilizing 'rule of thumb' decisions, brute-forced by the current beliefs of agents. This method, detailed below, forfeits "rational" decisions, but overcomes the two drawbacks of the game theoretic approach by letting choice be dictated locally by current beliefs. Given an EPM, a set of doxastic programs provides a multitude of possible updates. In modeling a dynamic process, the modeler must choose which model is suitable for the next update, based on no strict directions from the to-be-updated EPM. However, environment or agent behavior will often be seen as dictated to some degree by facts or beliefs from the current EPM, thus used as a guideline. To incorporate the next action model choice in a formal manner, *transition rules* are introduced, locally specifying the next update as a function of the current EPM.

Transition rules are used to characterize agent behavior. Each behavior is specified by a set of transition rules, each with a *trigger condition* and a *goal formula*. If an EPM satisfies some trigger conditions, the ensuing EPM must satisfy the matching goal formulas. An APM that ensures that the goals are obtained then satisfies, or *solves*,

the rules, and is seen as a possible choice for the agent in question. Hereby an EPM, a set of behavior-governing rules and a set of APMs jointly specify the transition to the next EPM.

## 4.1 Transition rules

A *transition rule* $\mathcal{T}$ is an expression $\varphi \rightsquigarrow [X]\psi$ where $\varphi, \psi \in \mathsf{Prop}_\Phi$. Call $\varphi$ the *trigger* and $\psi$ the *goal*. If EPM (**S**, $s_0$) satisfies the trigger of a transition rule $\mathcal{T}$, $\mathcal{T}$ is said to be *active* in **S** (else *inactive*).

Specified below, transition rules may be used to choose the next update based on local conditions of the current EPM. E.g., updates by the 'environment' may be specified using atoms in the trigger. To exemplify, let $R$ and $W$ be atoms with resp. readings 'it rains' and 'the street is wet', then the transition rule $\mathcal{T}_1 = R \rightsquigarrow [X]W$ reads 'if it rains, then the next update must be such that after it, the street is wet'. Transition rules may also be used as agent decision rules for factual change, using $B_i\varphi/K_i\varphi$-formulas as triggers and suitable formulas as effects. E.g., the set of transition rules $\{B_i R \rightsquigarrow [X]U_i, B_i\neg R \rightsquigarrow [X]\neg U_i\}$ may be used to specify agent behavior relative to rain: if $i$ believes it rains, then next $i$ will have an umbrella, and if $i$ believes it does not rain, then next $i$ will not have an umbrella. Used thus, transition rules are akin to the *programs* and *knowledge-based programs* of Fagin et al. (1995), here tailored to the DEL framework. They further instantiate one-step *epistemic planning problems*, in the terminology of Bolander and Birkegaard (2011).

## 4.2 Dynamic modalities

Note that *transition rules are not doxastic propositions*: the "modality" $[X]$ has no interpretation, and construed as a formula, $\mathcal{T}_1$ has no truth conditions. Instead, transition rules are *prescriptions* for choosing the next action model. The choice of model is made by implementing a transition rule over an EPM **S** and a set **G** of doxastic programs over one or more APMs using *dynamic modalities*.

For any program $\Gamma$ over APM **E**, $[\Gamma]$ is a dynamic modality, and the doxastic proposition $[\Gamma]\varphi$ is given by

$$([\Gamma]\varphi)_{\mathbf{S}} := \{s \in S : \forall \sigma \in \Gamma, \text{ if } (s, \sigma) \in S \otimes \Gamma \text{ then } (s, \sigma) \in \varphi_{\mathbf{S} \otimes \Gamma}\}.$$

That is, a state $s$ from **S** is a $[\Gamma]\varphi$-state iff every resolution of $\Gamma$ over $s$ is a $\varphi$-world in $\mathbf{S} \otimes \Gamma$. $[\Gamma]$-modalities are natural when $\varphi$ is desired common knowledge among $\mathcal{A}$.

Further, where a $\Gamma$ is doxastic program, let $[\Gamma]_i\varphi$ be given by

$$([\Gamma]_i\varphi)_{\mathbf{S}} := \{s \in S : \forall \sigma \in \Gamma, \text{ if } (s, \sigma) \in S \otimes \Gamma \cap \mathcal{K}_i[(s_0, \sigma_0)] \text{ then } (s, \sigma) \in \varphi_{\mathbf{S} \otimes \Gamma}\}.$$

That is, a state $s$ from **S** is a $[\Gamma]_i\varphi$-state iff every resolution of $\Gamma$ over $s$ *that is included in $i$'s information cell relative to the actual world* $(s_0, \sigma_0)$ *in* $\mathbf{S} \otimes \Gamma$ is a $\varphi$-world in $\mathbf{S} \otimes \Gamma$. Hence $\mathbf{S}, s \models [\Gamma]_i\varphi$ iff $\mathbf{S} \otimes \Gamma, (s_0, \sigma_0) \models K_i\varphi$.

The $[\Gamma]_i$-modalities are natural when transition rules prescribe prescribe agent choices, as they ensure that the performing agent *knows her choice* following the action, while allowing others to be unaware of the choice made.

### 4.3 Solutions and next APM choice

A set of transition rules dictates the choice for the next APM by finding the transition rule(s)'s *solution*. A *solution* to $\mathcal{T} = \varphi \rightsquigarrow [X]\psi$ over pointed EPM $(\mathbf{S}, s)$ is a doxastic program $\Gamma$ such that $\mathbf{S}, s \models \varphi \rightarrow [\Gamma]\psi$. $\Gamma$ is a solution to the set $\mathbf{T} = \{\mathcal{T}_1, ..., \mathcal{T}_n\}$ with $\mathcal{T}_k = \varphi_k \rightsquigarrow [X]\psi_k$ over $(\mathbf{S}, s)$ if $\mathbf{S}, s \models \bigwedge_1^n (\varphi_k \rightarrow [\Gamma]\psi_k)$, i.e. if $\Gamma$ is a solution to all $\mathcal{T}_i$ over $(\mathbf{S}, s)$ *simultaneously*.[15] Finally, a set of doxastic programs $\mathbf{G}$ is a solution to $\mathbf{T}$ over $\mathbf{S}$ iff for every $t$ of $\mathbf{S}$, there is a $\Gamma \in \mathbf{G}$ such that $\Gamma$ is a solution to $\mathbf{T}$ over $(\mathbf{S}, t)$.[16]

If $\mathbf{G}$ is a solution to $\mathbf{T}$ over $\mathbf{S}$, then given a state from $\mathbf{S}$, the transition rules in $\mathbf{T}$ will specify one (or more) programs from $\mathbf{G}$ as the next choice. A deterministic choice will be made if $\mathbf{G}$ is selected suitably, in the sense that it contains a *unique* $\Gamma$ for each $s$. In the ensuing, solution sets will be chosen thus.

### 4.4 Example: looping system

Consider the very simple 'system', consisting of an EPM $\mathbf{S}$ with $s_0 \in P_{\mathbf{S}}$, and APM $\mathbf{E}$ with $pre(\sigma_0) = P$, $post(\sigma_0) = \neg P$, and the set $\mathbf{T} = \{\mathcal{T}_0, \mathcal{T}_1\}$ of transition rules:

$$\mathcal{T}_0 = P \rightsquigarrow [X]\neg P \qquad \mathbf{S} : \begin{array}{|c|} \hline s_0 \\ \hline P \\ \hline \end{array} \quad \mathbf{E} : \begin{array}{|cc|} \hline \sigma_0 & \sigma_1 \\ \langle P; \neg P \rangle & \langle \neg P; P \rangle \\ \hline \end{array}$$
$$\mathcal{T}_1 = \neg P \rightsquigarrow [X]P$$

With $\Gamma_0 = \{\sigma_0\}$ and $\Gamma_1 = \{\sigma_1\}$, $\mathbf{G} = \{\Gamma_0, \Gamma_1\}$ is a solution to $\mathbf{T}$ over $\mathbf{S}$. For $\mathcal{T}_1$, $\mathbf{S}, s_0 \models \neg P \rightarrow [\Gamma_0]P$ as $s_0 \notin (\neg P)_{\mathbf{S}}$. For $\mathcal{T}_0$, it is easy to check that $\mathbf{S} \otimes \Gamma, (s_0, \sigma_0) \models \neg P$, entailing that $\mathbf{S}, s_0 \models [\Gamma_0]\neg P$. As $\Gamma_0$ is unique, this is chosen as next update. It should be easy to see that $\mathbf{G}$ is also a solution to $\mathbf{T}$ over $\mathbf{S} \otimes \Gamma_0$, where $\Gamma_1$ is chosen. Further re-application of $\mathbf{T}$ loops the system.[17]

### 4.5 Three agent types

Transition rules may be used to provide general characterizations of agent behavior determined by belief. Rules with a doxastic trigger will be referred to as *decision rules*, by sets of which an abundance of possible *agent types* may be defined. Of interest are the following three, corresponding to three types of human behavior relevant to the bystander effect (Table 1).

---

[15] Note the analogy with numerical equations; for both $2 + x = 5$ and $\{2 + x = 5, 4 + x = 7\}$, $x = 3$ is the (unique) solution.

[16] The definition is altered to suit transition rules using $[X]_i$ "modalities" by suitable replacing $[X]$ with $[X]_i$ and $[\Gamma]$ with $[\Gamma]_i$ throughout.

[17] For a definition of *system*, see (Rendsvig 2013a).

**Table 1** Decision rules specifying three agent types

| First Responder | City Dweller | Hesitator |
|---|---|---|
| $B_i A \rightsquigarrow [X]_i I_i$ | $B_i A \rightsquigarrow [X]_i E_i$ | $K_i A \rightsquigarrow [X]_i I_i$ |
| $B_i \neg A \rightsquigarrow [X]_i E_i$ | $B_i \neg A \rightsquigarrow [X]_i E_i$ | $B_i A \wedge \neg K_i A \rightsquigarrow [X]_i O_i$ |
| | | $B_i \neg A \rightsquigarrow [X]_i E_i$ |

Denote by $\mathcal{F}_{1i}$ and $\mathcal{F}_{2i}$ resp. the upper and lower first responder rule indexed for $i$, and set $\boldsymbol{F}_i := \{\mathcal{F}_{1i}, \mathcal{F}_{2i}\}$ and treat $\mathcal{C}_{1i}, \mathcal{C}_{2i}, \mathcal{H}_{1i}, \mathcal{H}_{2i}, \mathcal{H}_{3i}, \boldsymbol{C}_i$ and $\boldsymbol{H}_i$ in a similar manner

The First Responder will intervene if she believes there is an accident, otherwise not. First Responders thus reflect the normally expected, but not witnessed, behavior in relation to emergencies. A City Dweller will evade the scene no matter what his beliefs, hereby reflecting the media's grim picture of the "apathetic" urban citizen, ignoring the murder of Kitty Genovese. Finally, a Hesitator will choose to observe if she believes but does not *know* that there is an accident, and will else evade. Hereby the Hesitator rules capture (part of, see below) the behavior used as explanation for the bystander effect [e.g. by Latané and Darley (1968) when they write "it is likely that an individual bystander will be considerably influenced by the decisions he perceives other bystanders to be taking" (p. 216)]. Presently, focus will be on Hesitators, with comments on First Responders and City Dwellers. The latter two are subjects of Sect. 6.

### 4.6 Possible choices

To implement either of the rule sets, a suitable set of doxastic programs for $X$ to range over must be specified. It seems natural to assume that when an agent is intervening, then this is epistemically unambiguous for all agents. When $b$ and $c$ see $a$ choose either to observe or evade, it seems more plausible that they cannot tell these actions apart, as neither action has an observable, distinguishing mark.[18] It is assumed that agents find it more plausible that others evade than that they observe. [19] In sum, these consideration give rise to the APM $\mathbf{E}_{1i}$ of Fig. 6.

$\mathbf{E}_{1i}$ does not facilitate simultaneous choice, in the sense that it does not contain a solution to e.g. $\{\mathcal{H}_{1a}, \mathcal{H}_{1b}\}$ over $\mathbf{S}_1$. Combining, however, a copy of $\mathbf{E}_{1i}$ for each of $a$, $b$ and $c$ while respecting the doxastic links in an intuitive way may easily be done. Specifically, a combined APM $\mathbf{E}_1$ may be obtained by taking the reflexive, transitive

---

[18] Unless evading entails leaving the scene or observation is performed in a non-discrete manner. Often this is not the case, though: "Among American males it is considered desirable to appear poised and collected in times of stress. ... If each member of a group is, at the same time, trying to appear calm and also looking around at the other members to gauge their reactions, all members may be led (or misled) by each other to define the situation as less critical than they would if alone. Until someone [intervenes], each person only sees other nonresponding bystanders, and ... is likely to be influenced not to act himself." (Latané and Darley 1968, p. 216); "... Apparent passivity and lack of concern on the part of other bystanders may indicate that they feel the emergency is not serious, but it may simply mean that they have not yet had time to work out their own own interpretation or even that they are assuming a bland exterior to hide their inner uncertainty and concern." (Latané and Rodin 1969, p. 199).

[19] The second quote in the previous note seems to indicate the plausibility of this assumption.
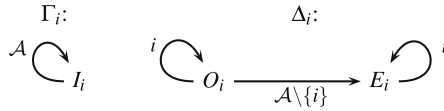
**Fig. 6** The APM $\mathbf{E}_{1i}$, representing the three moves available to $i$ as well as the doxastic perception of these for the remaining agents. State names specify postconditions; all preconditions are $\top$. If $i$ chooses to intervene, $\Gamma_i$ will be next APM choice, whereas if $i$ chooses to either observe or evade, $\Delta_i$ will be used, with actual event resp. $O_i$ or $E_i$



Type I (1x)     Type II (3x)     Type III (3x)     Type IV (1x)
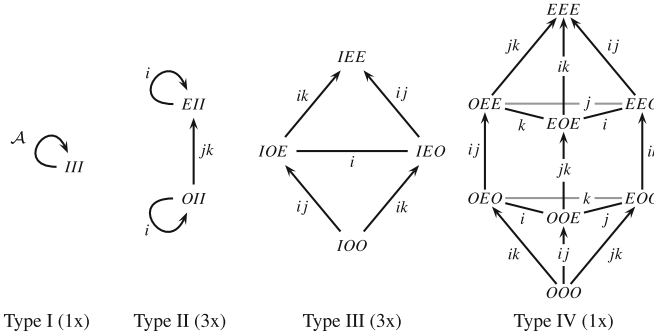
**Fig. 7** The APM $\mathbf{E}_1$, representing simultaneous move programs, omitting (most) reflexive and transitive *arrows*; states are labeled with postconditions, with $IOE$ representing $\langle \top; I_i O_j E_k \rangle$, etc.; all preconditions are $\top$. Notice the rise in dimensions; Type I is a point, Type IV is a cube

closure of the *Cartesian graph product* $\mathbf{E}_{1a}\Box\mathbf{E}_{1b}\Box\mathbf{E}_{1c}$ (see e.g. Hammack et al. 2011 for definition) and specifying pre- and postconditions as follows: for $(s, t, u) \in \mathbf{E}_{1a}\Box\mathbf{E}_{1b}\Box\mathbf{E}_{1c}$, let $pre(s, t, u)_{\mathbf{E}_1} = \top$ and $post(s, t, u) = post(s)_{\mathbf{E}_{1a}} \wedge post(t)_{\mathbf{E}_{1b}} \wedge post(u)_{\mathbf{E}_{1c}}$. The resulting APM $\mathbf{E}_1$ has eight mutually disconnected components of four types, see (Fig. 7).

In Fig. 7, the Type I model is obtained by the Cartesian product of $\Gamma_i$, $\Gamma_j$ and $\Gamma_k$; Type II from $\Gamma_j$, $\Gamma_k$ and $\Delta_i$; Type III from $\Gamma_i$ and $\Delta_j$, $\Delta_k$ and Type IV from $\Delta_i$, $\Delta_j$, $\Delta_k$. Notice that reflexive and transitive closure is required to ensure that all $\leq_i$'s are pre-orders. Euclidean closure is not required, but is doxastically reasonable: in e.g. the Type III component, $i$ should not be able to distinguish between $IOE$ and $IEO$, nor consider either more plausible.

Doxastic programs over $\mathbf{E}_1$ identical to each of the four sub-model types give rise to the desired solution space $\mathbb{E}_\Gamma$. Let $\mathbb{E}$ be the set consisting of all pointed $\mathbf{E}_1$ models, and let $\mathbb{E}_\Gamma$ be the set of all doxastic programs over models in $\mathbb{E}$ such that each $\Gamma \in \mathbb{E}_\Gamma$ contains exactly one of the type I-IV sub-models. Then $\mathbb{E}_\Gamma$ contains a unique solution to every combination of the three agent types. (recall, def. of solution set: for every $s \in S$.)

Over $\mathbf{S}_1$ and $\mathbb{E}_\Gamma$, the Type I program is the unique solution to $\boldsymbol{F}_1 \cup \boldsymbol{F}_2 \cup \boldsymbol{F}_3$; the Type II program with $i = 2$ and $\sigma_0 = EII$ is the unique solution to $\boldsymbol{C}_1 \cup \boldsymbol{F}_2 \cup \boldsymbol{F}_3$ – with $\sigma_0 = OII$, it is the unique solution to $\boldsymbol{H}_1 \cup \boldsymbol{F}_2 \cup \boldsymbol{F}_3$; the Type III program with $i = 1, j = 2, k = 3$ and $\sigma_0 = IEO$ is unique solution to $\boldsymbol{F}_1 \cup \boldsymbol{C}_2 \cup \boldsymbol{H}_3$; the Type IV program with $\sigma_0 = OOO$ is unique solution to $\boldsymbol{H}_1 \cup \boldsymbol{H}_2 \cup \boldsymbol{H}_3$. Interestingly, the

Type IV program with actual state $OOO$ reflects an implicit norm-based pluralistic ignorance: though every agent is observing, each perceive the situation as one where they are the only ones doing it.

## 4.7 Choosing to observe

Given the above, the system based on $\mathbf{S_1}$ with $\boldsymbol{H}_1 \cup \boldsymbol{H}_2 \cup \boldsymbol{H}_3$ as rules over $\mathbf{E_1}$ dictates the type IV program with $\sigma_0 = OOO$ as next APM choice. That is, every agent chooses to observe. The ensuing EPM $\mathbf{S_2} := \mathbf{S_1} \otimes \Gamma_{IV}$ contains 128 states, but is easily described. Take $\Gamma_{IV}$ and replace every state with a complete copy of $\mathbf{S_1}$'s frame, connect two states from two different copies $[(s, \sigma_1) \leq_i (s', \sigma_2)]$ iff $s = s'$ and $\sigma_1 \leq_i \sigma_2$, and finally take the reflexive-transitive closure. The new actual world is $[(s_0, \sigma_0), OOO]$, satisfying

$$A, \ O_a \wedge O_b \wedge O_c \text{ and } \bigwedge_{i \in \mathcal{A}} \left( K_i O_i \wedge B_i \bigwedge_{i \in \mathcal{A} \setminus \{i\}} E_i \right).$$

The latter captures the post-factual effects of the mentioned norm-based pluralistic ignorance of $\Gamma_{IV}$: all have the belief that *they individually were the only ones to observe, while all others evaded*. Importantly, had EPM $\mathbf{E_{1}}_i$ on the previous page been such that agents perceived actions according to their own decision rules, i.e. found observation more plausible than evasion in others, all would have correct beliefs about the actions of others.

## 5 Action interpretation and social proof

Albeit all agents have formed the belief that their co-witnesses evaded, none has formed any beliefs rationalizing these choices: All are still doxastically indifferent between whether the others believe $A$ or $\neg A$. Neither does any agent have means of deducing others' beliefs, given the introduced formal framework. Such a deduction would require e.g. the ability to rationalize by *forward induction*, which requires information from both past play and future possibilities (Benthem 2014) couched in a game framework representing preferences, rationality, etc. (see e.g. Rendsvig and Hendricks 2013 for an implementation). Though structures akin to game trees may be defined using EPMs, APMs and protocols (Benthem et al. 2007; Dégremont 2010), a simpler, more superficial construct may be used to facilitate the reasoning. The suggested approach utilizes an 'inverse' version of decision rules, brute forcing conclusions about belief from observations about action.

In making decisions, our beliefs about the relevant state of affairs dictate our action, up to error and human factors. Hence the route from beliefs to actions is often functional. As this function will often not be injective, moving from actions to beliefs is not as straightforward, since multiple different belief states may result in the same action. Having to provide a rationalization of a given action will therefore often include abductive reasoning. An abductive hypothesis to rationalize an action allows inferring

as explanation of the observed action a previous belief state of the acting agent. Below, such hypotheses are called *interpretation rules*.

5.1 Interpretation rules

An *interpretation rule* is a doxastic proposition $\varphi \rightarrow [\mathbf{S}]B_i\psi$, with $\varphi$ called the *basis* and $\psi$ the *content*, with the underlying idea that on the basis of an action (e.g. $E_i$), agents may deduce something about the content of $i$'s beliefs (e.g. that $B_i\neg A$).

Doxastic propositions involving the modality of the consequent are given by

$$([\mathbf{S}]\chi)_{\mathbf{S}'} := \{s' \in S' : \exists s \in S \text{ such that } s \in s' \text{ and } (\mathbf{S}, s) \models \chi\},$$

where $S$ is the domain of $\mathbf{S}$, and where $s \in s'$ means that $s$ is a *predecessor*[20] of $s'$. Hence $[\mathbf{S}]\chi$ is true in $(\mathbf{S}', s')$ just in case $s'$'s predecessor in $\mathbf{S}$ was a $\chi$-world. The modality is included to respect the temporal aspect introduced by updates, and $\mathbf{S}$ is to be substituted with the EPM based on which $i$ made the choice in question.

A set of interpretation rules may in general be implemented using an APM where the preconditions of each state is a conjunction of interpretation rules with different bases with a conjunct for each action to be interpreted. Hereby each state represents a different hypothesis regarding the acting agent's type, i.e. how the agent made decisions. The plausibility order then specifies the 'abductive hierarchy' of such hypotheses.[21]

To simplify, agents are given only *one* hypothesis about types, the hypothesis also being *correct* in the sense that the interpretation rules are (close to) the converse of the transition rules that are in fact applied. Hence the interpretation rule model APM $\mathbf{E_2}_{ij}$ that determines how agents $\mathcal{A}\backslash\{j\}$ interprets the actions of $j$ is a one state model. Let $\rho_j \in \mathbf{E_2}_j$ and set

$$
\begin{aligned}
pre(\rho_j) := \; & E_j \rightarrow [\mathbf{S_1}]B_j\neg A && \wedge \\
& I_j \rightarrow [\mathbf{S_1}]K_j A && \wedge \\
& O_j \rightarrow [\mathbf{S_1}]B_j A \wedge \neg K_j A
\end{aligned}
$$

Applying such rules for all agents may be done by sequential application of $\mathbf{E_2}_j$ for each $j \in \{a, b, c\}$ on $\mathbf{S_2}$ (and the resulting models).[22] Call the APU product $\mathbf{S_3}$.

---

[20] When constructing APU products, a state in the product model is an ordered pair $(s, \sigma)$ of a state $s$ and action $\sigma$. In this pair, $s$ may again be such a pair. Say that a *predecessor* of $s'$ is any $s$ that occurs in any of the ordered pairs of $s'$, including $s'$ itself.

[21] It is possible to give agents a choice of interpretation by invoking transition rules with interpretation rules as possible solutions. In the present, agents are given *no choice of interpretation*, and this construction is consequently skipped for simplicity.

[22] Each $\mathbf{E_2}_j$ functions as a *truthful public announcement* of $pre(\rho)$, for which the order of announcements does not matter (Baltag and Smets 2009): states are deleted, the remaining orderings staying as previous. Deleting simultaneously or in some sequence makes no difference.

## 5.2 Belief-based pluralistic ignorance

By application of interpretation rules, agents establish beliefs about each others previous beliefs, but even though the applied interpretation rules were correct, the obtained beliefs are wrong: $\mathbf{S_1} \models B_b A$, but $\mathbf{S_3} \models B_a[\mathbf{S_1}]B_b \neg A$. This is a direct consequence of the mis-perception of actions occurring in $\mathbf{E_1}$.[23]

In the actual world of $\mathbf{S_3}$, the agents are in a state of belief-based pluralistic ignorance with respect to $A$: $\mathbf{S_3} \models \bigwedge_{i \in \mathcal{A}}(B_i A \wedge \bigwedge_{j \in \mathcal{A}\setminus\{i\}} B_i B_j \neg A)$, cf. the definition in Hansen (2011), ch. 6. To see this, notice what happens when $a$ interprets the actions of $b$ and $c$ over $\mathbf{S_2}$. In $\mathbf{S_2}$, the most plausible copy of $\mathbf{S_1}$ (Fig. 4) is the one in which all states satisfy $O_a \wedge E_b \wedge E_c$ (Fig. 7). Of these 16 states, only 4 satisfy both $pre(\rho_b)$ and $pre(\rho_c)$, namely the successors of $\sigma_2, \tau_2, \sigma_7$ and $\tau_7$, and only the first is in $a$'s plausibility cell relative to $(s_0)_{\mathbf{S_2}}$. As all other states in this $\mathbf{S_1}$-copy are deleted upon update with $\mathbf{E_{2b}}; \mathbf{E_{2c}}$, it follows that $a$'s plausibility cell $\mathcal{B}_a[(s_0)_{\mathbf{S_3}}]$ contains only the state $((((s_0, \sigma_2), OEE), \rho_b), \rho_c)$, which in turns satisfies $A \wedge B_b \neg A \wedge B_c \neg A$. Hence $\mathbf{S_3} \models B_a(A \wedge B_b \neg A \wedge B_c \neg A)$. Analogous reasoning for $b$ and $c$ shows that $(s_0)_{\mathbf{S_3}}$ is a state of pluralistic ignorance w.r.t. $A$.

Again, importantly, had EPM $\mathbf{E_{1i}}$ been defined so that agents considered observation more plausible than evasion, this state of pluralistic ignorance would *not* have arisen.

## 5.3 Social proof

In the portrayal of the bystander effect, witnesses alter their beliefs following their mutual act of orientation, and in the light of the newly obtained information that no one else believes that there is cause for alarm, concludes that no intervention is required. To represent the revised beliefs of agents [24], introduce a new operator $SB_{i|G}$, representing the beliefs of agent $i$ when socially influenced by her beliefs about the beliefs of agents from group $G$. $SB_{i|G}$ is defined using simple majority 'voting' with a self-bias tie-breaking rule: let

$$s \in (SB_{i|G}\varphi)_{\mathbf{S}} \text{ iff } \alpha + |\{j \in G : s \in (B_i B_j \varphi)_{\mathbf{S}}\}| > \beta + |\{j \in G : s \in (B_i B_j \neg\varphi)_{\mathbf{S}}\}|$$

with tie-breaking parameters $\alpha, \beta$ given by

$$\alpha = \begin{cases} \frac{1}{2} & \text{if } s \in (B_i \varphi)_{\mathbf{S}} \\ 0 & \text{else} \end{cases} \qquad \beta = \begin{cases} \frac{1}{2} & \text{if } s \in (B_i \neg\varphi)_{\mathbf{S}} \\ 0 & \text{else} \end{cases}$$

---

[23] Though time has passed, beliefs have not changed, and this is known to all: $\mathbf{S_3} \models \bigwedge_{i \in \mathcal{A}} K_i(\bigwedge_{j \in \mathcal{A}}[\mathbf{S_1}]B_j\varphi \to B_j\varphi)$ for $\varphi \in \{A, \neg A\}$.

[24] Strictly speaking, in the present model agents do not *revise* their beliefs. An additional operator is instead introduced to facilitate comparison with private beliefs. A belief revision policy may easily be defined using decision rules to the effect that agents update their beliefs under the suitable circumstances, see (Rendsvig 2013b).

This definition leaves agent $i$'s 'social beliefs' w.r.t. $\varphi$ undetermined [(i.e. $\neg(SB_{i|G}\varphi \vee SB_{i|G}\neg\varphi)$)] iff both $i$ is agnostic whether $\varphi$ and there is no strict majority on the matter.

Applying the notion of social belief to $A$ in $\mathbf{S_3}$, it is easily seen that $\mathbf{S_3} \models \bigwedge_{i \in \mathcal{A}} SB_{i|\mathcal{A}}\neg A$. That is, upon incorporating social proof, all agents 'socially believe', contrary to their private beliefs, that no accident occurred.

## 5.4 Action under influence

Notice that none of the three agent types introduced so far will change their action if presented again prompted to intervene, observe or evade. First Responders will again intervene, City Dwellers will again choose to evade, and Hesitators will again, irrespective of social proof, choose to observe.

To make Hesitators pay heed to the observation they chose to make, their decision rules are changed (in the ensuing section, a fusion of the two types is defined). Let an 'influenced' agent act in accordance with the following rules:

**Influenced** :
$$SB_{i|G}A \rightsquigarrow [X]_i I_i$$
$$SB_{i|G}\neg A \rightsquigarrow [X]_i E_i$$

Note that an Influenced agent acts like a First Responder who bases her actions on social beliefs.

A Hesitator-now-turned-Influenced presented with the choice to intervene, observe or evade (as given by $\mathbb{E}_\Gamma$) will choose to evade. More precisely, if $a$, $b$ and $c$ are Influenced, the unique next APM choice will be the Type IV program with $\sigma_0 = EEE$. The actual world in the ensuing EPM $\mathbf{S_4}$, the final step of the model, will then satisfy

$$A \wedge \bigwedge_{i \in \mathcal{A}} B_i A \wedge \bigwedge_{i \in \mathcal{A}} SB_{i|G}\neg A \wedge \bigwedge_{i \in \mathcal{A}} E_i.$$

The last conjunct is an (unfortunate) end condition, as specified in Sect. 3.1.6. Hereby, informational dynamics leading to an observable bystander effect has been modeled.

## 6 Comparison to empirical studies

The presented sequence of models, transition rules and updates conjoined captures important informational aspects of the observable bystander effect, given that the model is accepted. As presented, the sequence may be regarded as one possible execution of a broader, implicit system. Other runs of this system may be constructed by varying parameters relevant to the bystander effect.

In this section, the effect of changing two parameters will be presented. The first change is of agent types, where it is seen that for non-mixed groups, both City Dwellers and (Influenced) Hesitators will produce the observable bystander effect. The second

**Table 2** Re-specification of agent types

| City Dweller | Hesitator | Influenced Hesitator | First Responder |
|---|---|---|---|
| $B_i A \rightsquigarrow [X]_i E_i$ | $K_i A \rightsquigarrow [X]_i I_i$ | $K_i A \rightsquigarrow [X]_i I_i$ | $B_i A \rightsquigarrow [X]_i I_i$ |
| $B_i \neg A \rightsquigarrow [X]_i E_i$ | $B_i \neg A \rightsquigarrow [X]_i E_i$ | $B_i \neg A \rightsquigarrow [X]_i E_i$ | $B_i \neg A \rightsquigarrow [X]_i E_i$ |
| | $B_i A \wedge \neg K_i A \rightsquigarrow [X]_i O_i$ | $\neg K_i A \wedge \neg(SB_{i\|G} A \vee SB_{i\|G} \neg A) \rightsquigarrow [X]_i O_i$ | |
| | | $O_i \wedge SB_{i\|G} A \rightsquigarrow [X]_i I_i$ | |
| | | $O_i \wedge SB_{i\|G} \neg A \rightsquigarrow [X]_i E_i$ | |

variation is group size, and it is shown that of non-mixed populations, only (Influenced) Hesitator behavior varies as a function of group size.

Let us briefly outline the implicit system before changing parameters. The system has initial state $\mathbf{S_0}$ in Sect. 3.1.6, where everybody knows nothing has happened and end conditions either $\bigvee_{i \in \mathcal{A}} I_i$ or $\bigwedge_{i \in \mathcal{A}} E_i$. $\mathbf{S_0}$ is updated with the occurrence of the accident, $\mathbf{E_0}$ (Fig. 5) resulting in $\mathbf{S_1}$ (Sect. 3.2.7), where all believe an accident has occurred, while having no information about others' beliefs. Apart from adding further agents to the population, these steps will remain fixed.[25] Next, agents make a first decision over $\mathbb{E}_\Gamma$ and $\mathbf{S_1}$ is updated with the next APM choice.[26] Depending on agent types, the run might end at $\mathbf{S_2}$. If not, the interpretation rule model of Sect. 5.1 is applied for all agents, and a second decision is made based on the outcome, possibly involving the aggregation of the perceived beliefs of others'. Again, if the system does not satisfy one of the end conditions, it will continue, in which case the interpretation rule model is re-applied (suitably altered to accommodate the temporal shift), followed by decisions, etc.

In the run described in the previous sections, two *different* agent types were used. For the first choice made, agents were assumed to be Hesitators, making them choose to observe. For their second choice, they were assumed to be Influenced, making them act on their social beliefs.[27] To facilitate comparison of models, this 'mixed' type may be properly defined as *Influenced Hesitators* (Table 2):

Notice that Influenced Hesitators behave as a mixture of Hesitators (first three rules) and First Responders (last two), but who take social proof into account. Notice the difference between third rule for Hesitators and the same for Influenced Hesitators. The latter requires that Influenced Hesitators have undetermined social beliefs before they choose to observe. The altered First Responder rules (rows four and five) capture that if the agents has observed and have determined social beliefs, observation gives way to intervention or evasion.[28]

---

[25] Concerning $\mathbf{E_0}$, it should be obvious how the APM must be altered to include further agents, while maintaining complete higher-order ignorance.

[26] Again, it should be obvious how $\mathbb{E}_\Gamma$ may be altered to accommodate for a larger population.

[27] The shift was made to ease the exposition. Influenced agents require the notion of social beliefs, not necessary for Hesitators' first choice.

[28] The requirement that $i$ must have observed before acting on social beliefs ensures that agents do not intervene immediately after seeing the accident (a private belief that $A$ would imply that $SB_{i\|G} A$, as agents then hold no beliefs regarding others' beliefs).

**Table 3** Summary of end conditions and times as a function of agent type and group size

| | 1 | 2 | 3 | $k \geq 4$ |
|---|---|---|---|---|
| FR | $I_2$ | $I_2$ | $I_2$ | $I_2$ |
| CD | $E_2$ | $E_2$ | $E_2$ | $E_2$ |
| H | – | – | – | – |
| IH | $I_3$ | $I_3$ | $E_3$ | $E_3$ |

## 6.1 Varying types and group size

Table 3 summarizes the end conditions and the EPM number where they arise relative to agent types and group size. For end conditions, $I_k$ and $E_k$ represent $\bigvee_{i \in \mathcal{A}} I_i$ and $\bigwedge_{i \in \mathcal{A}} E_i$ respectively being satisfied in EPM $\mathbf{S_k}$, with $k$ rising as described in the previous sections. '–' means that no end conditions are met.

In words, for any group size, First Responders will intervene in $\mathbf{S_2}$, i.e. immediately following the accident. At the same time, City Dwellers will evade the scene, no matter the group size. 'Simple' Hesitators, as defined in Table 2, will never reach an end conclusion, as they will never come to either believe there is no accident, or know that there is one. That all these three agent types' actions are invariant over group size is due to their inherently *non-social* decision rules. The 'social', or Influenced, Hesitators will however change their behavior according to group size: they will intervene immediately if the group size is small enough for their private belief not to be 'overridden' by social proof. If the group size is 3 or above, Influenced Hesitators will conclude, by the mis-perception of others' choice to observe as an act of evasion and the resulting state of pluralistic ignorance, that enough agents believe that no accident occurred for themselves to be 'socially convinced' that this is the case. Consequently, they will choose to evade in $\mathbf{S_3}$ for any group size of 3 or above.

At group size 2 these agents still decide to intervene because they use their own belief to break the tie between what they perceive as an even split on whether an accident is happening.[29]

## 6.2 Comparison to empirical studies

Running the system with each of the four agent types may be considered as providing four different models of the bystander effect, each of which may be compared to empirical results to evaluate consistency with data. Table 3 allows for only a simple comparison, checking whether end conditions as a function of group size correlates properly with the observed.

A wide variety of studies have been performed on the inhibiting effect of the presence of others in situation requiring intervention (see e.g. Latané and Nida 1981 for a meta-study). Many of these have different, more specific foci, e.g. the role of diffusion of responsibility, friendship, gender, and more. As the focus of this paper is the second step of the bystander effect, only studies on the effect of social proof on the perception

---

[29] Cf. the tie-breaking rule used in the definition of social beliefs.

of the accident are relevant. Alas, no one such has been found that provides suitable data for all group sizes. Comparison can therefore only be made using multiple studies invoking different experimental settings.

Inherently, the presented models are *deterministic*, while experimental data provides information about the *percentage* of individuals who intervene. Given this, data will *not* be correctly matched. To evaluate tendency of correctness, acceptance or rejection of models are therefore based not on strict consistency, but on the weak requirement that a model must correctly match the binary experimental end conditions in strictly more than 50 % of cases.

## 6.3 Smoky room and the rejection of FR, CD and H

The classic 'smoky room' experiment of Latané and Darley (1968), specifically designed to test the hypothesis of the second step of the explanation of the bystander effect (see p. 2), has served as a strong guide for the construction of the models, and provides data which allows the rejection of three of them. In the study, groups of size 1 or 3 where sat in a waiting room, completing questionnaires. The groups of size 3 either consisted of 1 individual naive to the experiment and 2 of the experimenters' confederates, or 3 naive subjects. While completing the questionnaire, smoke was introduced to the room through a visible vent, ambiguously indicating either an emergency (e.g. fire) or not (e.g. steam). As the possible accident will have dire consequences for the subjects themselves, the experimenters assumed that no diffusion of responsibility arose.

The experiment was stopped when either one agent intervened, or after six minutes of smoke introduction and questionnaire completion, at which point smoke was heavy. Compared to the model, these end conditions are identified with $\bigvee_{i \in \mathcal{A}} I_i$ and $\bigwedge_{i \in \mathcal{A}} E_i$, respectively.

Of the subjects that were alone, 75 % reported the smoke, a number high enough to warrant the rejection of the City Dweller model, which would have it that all evade. Likewise, the Hesitator model is rejected, as it would have it that individuals would continuously observe (instead of completing their questionnaires). Both the First Responder and Influenced Hesitator models score better than 50 %.

With 2 confederates in the room, only 1 in 10 naive subjects intervened. With 75 % of individuals intervening when alone, it should be expected that 98 % of groups of size 3 with three naive subjects would intervene if individuals acted independently, [30] but only 38 % of the 8 groups did so.

The First Responder model is 10 % correct in the confederates condition[31] and 38 % in the naive group condition, hereby falling below the 50 % mark. The Influence Hesitators model does better: it is 90 % correct of the confederates condition[32] and 62 % correct in the naive group condition. Hence, it fairs better than 50 % overall.

---

[30] See Latané and Darley (1968) for calculation of hypothetical baseline based on the alone condition.

[31] In a mixed population model, using City Dweller agents for the two confederates.

[32] Again in a mixed population model, using City Dweller agents for the two confederates.

Interestingly, the response time for intervention in the three naive subjects case was considerably longer than the single subject case, indicating (1) that individuals in groups did pay heed to social proof before acting, and (2) that in many cases (38 %), social proof from only two peers was not enough to preclude intervention.

6.4 Is One Additional Witness Enough for Intervention Inhibition?

The smoky room study only compares groups of size 1 and 3, whereby it does not supply sufficient data to evaluate the IH model for group sizes of 2 or above 3. To evaluate the model for group size of 2, another classic experiment may be consulted, namely the 'lady in distress' case of Latané and Rodin (1969). Three conditions where tested in this experiment: with a lone, naive subject, with one naive subject and one confederate, and with two naive subjects, again filling out questionnaires. In a simulated accident, the female interviewer faked a fall in an easily accessible and audible adjacent room. The fall was indicated by a loud crash, a scream and subsequent moaning of complaints and hurt. Contrary to the smoky room, this accident is ambiguous between either a serious accident (e.g. broken leg) or a not-so-serious one (weakly sprained ankle).

In the first condition, 70 % of the alone subjects intervened, with a strong drop to 7 % when an inactive confederate was introduced. With two naive subjects, 91 % of groups would be expected to intervene if subjects acted independently, whereas only 40 % of such groups in fact did so.

These percentages strongly contradict the IH model for group size 2, as the expectation is that neither of the two agents would be sufficiently influenced by each other to not act on their private beliefs. Both would therefore intervene following observation. This makes the model incorrect in 60 % of cases, making it worse than a random bet.

Partly, the model may misfire as the experiment does not conform to the pluralistic ignorance explanation of the bystander effect. Specifically, the experiment does not preclude the possibility of a mix of social proof and diffusion of responsibility effects, given that the accident in question did not put the subjects in faked danger. An experiment precluding diffusion effects may be conjectured to show a higher degree of intervention, yielding a better fit.

The meta-study (Latané and Nida 1981) strongly indicates that determining social influence occurs in groups of size 2. Summarizing 33 studies with face-to-face interaction, the effective individual probability of helping was 50 %, with an effective individual response rate in groups only 22 %. Most of these studies involved groups of 2.[33] Hence, it seems that the model misfires when it comes to groups of size 2.

The obvious parameter to tweak for a better fit is the self-biased majority voting definition of social beliefs, which does not put enough weight on the other in the 2 person case. Changing this to one favoring the perceived beliefs of the other would yield a better model for the 2 subject case, while it would not alter the results for the

---

[33] How well these individual studies conform to the pluralistic ignorance explanation of the bystander effect has not been checked.

| Table 4 Summary of end conditions as a function of agent type and group size for others-biased social beliefs | | 1 | 2 | 3 | $k \geq 4$ |
|---|---|---|---|---|---|
| | **IH′** | $I_3$ | $E_3$ | $E_3$ | $E_3$ |

group size 3 case. Table 4 summarizes the effect of the Influenced Hesitator model run using a tie-breaking rule favoring the *opposite* belief of ones own.[34]

Though this model does not fair very well on the data from the smoky room and lady in distress studies, it does at least do better than a random bet.

## 7 Conclusions

Accepting the individual modeling steps as reasonable explications of the reasoning steps occurring in bystander effect-like scenarios, the constructed dynamics allows for a number of conclusions about such phenomena:

– The sub-optimal choice for all to evade is *not* a consequence of "apathetic" agents: City Dweller evasion is not influenced by group size.
– The sub-optimal choice for all to evade is a direct result of considering social proof in a state of proposition-based pluralistic ignorance: Influenced Hesitators with *correct* beliefs about their peers beliefs would choose to intervene.
– Proposition-based pluralistic ignorance arises due to norm-based pluralistic ignorance: that all agents assume others are evading when they themselves are observing is a necessary condition for the state of proposition-based pluralistic ignorance to arise.
– Subjects *do not* incorporate social proof by self-biased majority voting, but rather the opposite.

Several venues further for both formal and empirical research present themselves. As no pure agent type group fits data very well, two possibilities are worth investigating. First, how well will models with mixed groups perform? That not all subjects chose to intervene in the single agent case seems to indicate that at least some behave as City Dwellers; that some chose to intervene in the three agent case indicates that some act as First Responders. With suitable proportions of each agent type, a model may be produced which will match data more closely with an average of end conditions of runs based on random picks from the mixed population. To fit both population mix and the social influence parameter, a data set from a large-scale smoky room-style study would be required.

Finding implementable resolution strategies for the pluralistic ignorance state could be of benefit, if these turn out to work in practice. Some such have been suggested in the literature; in the study of Schroeder and Prentice (1998), information on the subject diminished the alcohol consumption among college students. How information should provoke changes in agent type in the present framework is an open question. A

---

[34] I.e., interchanging the $\alpha$, $\beta$ tie-breaking parameters. Alternatively, social beliefs could be defined by weighing others' perceived beliefs higher than one's own, or by moving to a threshold rule requiring e.g. perceived agreement with all peers as done in Seligman et al. (2013), Christoff and Hansen (2013).

shorter term strategy for obtaining help is suggested by Cialdini (2007): Single out an individual and ask *only her* for help. If an agent is singled out, the inaction of others will no longer be eligible as a source of information about the event. Hence the agent is forced to act on her private beliefs, in which case both Influenced Hesitator models predicts intervention. Of formal studies, Proietti and Olsson (2013) show how a state of pluralistic ignorance state may be dissolved by a series of announcements of private beliefs heard by matrix neighbors. Specifying an agent type replicating the behavior and varying only the network structure of the model might provide further insights into the fragility of the phenomenon.

For a complete model of the bystander effect, both the first (noticing the event) and third step (assuming responsibility) of the explanation provided in the introduction must be modeled. The former may rest less on information processing than features of physical space: as more people are present, less may notice the event e.g. due to obscured line of sight. Modeling the third step may require a more expressive logical framework in which beliefs regarding agent types may be held: if all agents falsely believe a First Responder is present, all may believe that intervention is required while no-one will take action.

# References

Baltag, A., & Smets, S. (2008). A qualitative theory of dynamic interactive belief revision. In G. Bonanno, W. van der Hoek, & M. Wooldridge (Eds.), *Logic and the foundations of game and decision theory (LOFT 7), number Loft 7 in texts in logic and games* (Vol. 3, pp. 9–58). Amsterdam: Amsterdam University Press.

Baltag, A., Smets, S., et al. (2009). Talking your way into agreement: Belief merge by persuasive communication. In M. Baldoni (Ed.), *Proceedings of the second multi-agent logics, languages, and organisations federated workshops, CEUR workshop proceedings* (Vol. 494, pp. 129–141). Italy: Turin.

Benthem, J. V. (2007). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, *17*(2), 129–155.

Benthem, J. V. (2014). *Logic in games*. Cambridge, MA: MIT Press.

Benthem, J. V., Gerbrandy, J., & Pacuit, E. (2007). Merging frameworks for interaction: DEL and ETL. In D. Samet (Ed.), *Proceedings of TARK 2007* (pp. 72–81).

Bicchieri, C., & Fukui, Y. (1999). The great illusion: Ignorance, informational cascades, and the persistence of unpopular norms. *Business Ethics Quarterly*, *9*(1), 127–155.

Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal logic*. Cambridge: Cambridge University Press.

Bolander, T., & Birkegaard, M. (2011). Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, *21*(1), 9–34.

Christoff, Z., & Hansen, J. U. (2013). A two-tiered formalization of social influence. In D. Grossi & O. Roy (Eds.), *Logic, rationality, and interaction, lecture notes in computer science* (pp. 68–81). New York: Springer.

Cialdini, R. (2007). *Influence: The psychology of persuasion*. UK: HarperCollins Publishers.

Dégremont, C. (2010). *The temporal mind: Observations on the logic of belief change in interactive systems*. PhD thesis, University of Amsterdam.

Demey, L. (2011). Some remarks on the model theory of epistemic plausibility models. *Journal of Applied Non-Classical Logics*, *21*(3–4), 375–395.

 Springer

Ditmarsch, H. V., & Kooi, B. (2008). Semantic results for ontic and epistemic change. In G. Bonanno, W. van der Hoek & M. Wooldridge (Eds.), *Logic and the foundations of game and decision theory (LOFT 7)* (pp. 87–117), Amsterdam University Press.

Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about knowledge*. Cambridge, MA: The MIT Press.

Halbesleben, J. R., & Buckley, M. R. (2004). Pluralistic ignorance: Historical development and organizational applications. *Management Decision*, *42*(1), 126–138.

Hammack, R., Imrich, W., & Klavzar, S. (2011). *Handbook of product graphs, discrete mathematics and its applications* (2nd ed.). Boca Raton, FL: CRC Press.

Hansen, J. U. (2011). *A logic toolbox for modeling knowledge and information in multi-agent systems and social epistemology*. PhD thesis, Roskilde University.

Hansen, P. G., Hendricks, V. F., & Rendsvig, R. K. (2013). Infostorms. *Metaphilosophy*, *44*(3), 301–326.

Hendricks, V. F., & Rasmussen, J. L. (2012). *Nedtur! Finanskrisen forstået filosofisk*. Copenhagen: Gyldendal Business.

Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, *10*(3), 215–221.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?*. New York: Appleten-Century-Crofts.

Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, *89*(2), 308–324.

Latané, B., & Rodin, J. (1969). A lady in distress: Inhibiting effects of friends and strangers on bystander intervention. *Journal of Experimental Social Psychology*, *5*, 189–202.

Miller, D., & McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, *53*, 298–305.

Myers, D. G. (2012). *Social psychology* (11th ed.). New York, NY: McGraw-Hill.

Prentice, D., & Miller, D. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, *64*, 243–256.

Proietti, C., & Olsson, E. J. (2013). A DDL approach to pluralistic ignorance and collective belief. *Journal of Philosophical Logic*, *March*, 1–17.

Rendsvig, R. K. (2013a). Aggregated beliefs and informational cascades. In D. Grossi & O. Roy (Eds.), *Logic, rationality, and interaction, lecture notes in computer science* (pp. 337–341). Berlin: Springer.

Rendsvig R. K. (2013b) Look and learn: Extracting information from actions of others, http://staff.science.uva.nl/~ulle/esslli-workshop-2013/papers/rendsvig.pdf .

Rendsvig, R. K., & Hendricks, V. F. (2013). Social proof in extensive games. (under review).

Schroeder, C. M., & Prentice, D. A. (1998). Exposing pluralistic ignorance to reduce alcohol use among college students. *Journal of Applied Social Psychology*, *28*(23), 2150–2180.

Seligman, J., Girard, P., & Liu, F. (2013). Logical dynamics of belief change in the community. Semantic results for ontic and epistemic change. In G. Bonanno, W. V. D. Hoek, & M. Wooldridge (Eds.), *Logic and the foundations of game and decision theory (LOFT 7), number Loft 7 in texts in logic and games* (Vol. 3, pp. 87–117). Amsterdam: Amsterdam University Press.

Westphal, J. D., & Bednar, M. K. (2005). Pluralisitc ignorance in corporate boards and firms' strategic persistence in response to low firm performance. *Administrative Science Quarterly*, *50*, 262–298.

# Paper II

# Model Transformers for Dynamical Systems of Dynamic Epistemic Logic

Rasmus K. Rendsvig[1,2]

[1] Information Quality Research Group, Lund University, Lund, Sweden
[2] Center for Information and Bubble Studies, University of Copenhagen,
Copenhagen, Denmark
rendsvig@gmail.com

**Abstract.** This paper takes a dynamical systems perspective on the semantic structures of dynamic epistemic logic (DEL) and asks the question which orbits DEL-based dynamical systems may produce. The class of dynamical systems based directly on action models produce very limited orbits. Three types of more complex model transformers are equivalent and may produce a large class of orbits, suitable for most modeling purposes.

**Keywords:** dynamic epistemic logic, dynamical systems, model transformers, protocols, modeling.

## 1 Introduction

When modeling socio-epistemic phenomena, working with the temporally local models of dynamic epistemic logic (DEL) is both a blessing and a bane. It is a blessing as both epistemic state models and their updates are small relative to a fully explicated epistemic temporal structure. This eases both model construction and comprehension. It is a bane as the small models are incomplete: each is an individual time-step while we seek to model temporally extended dynamics. To form a 'complete model', we must specify the 'temporal glue' that ties individual epistemic states together to dynamics.

This 'temporal glue' is often presented informally in the DEL literature by way of a natural language problem description, typically involving conditional tests to determine which update to apply. Methodologically, this leaves modelers with a small gap: when modeling information dynamics using the semantic tools of DEL, what *mathematical object* shall we identify as *the model* of our target phenomenon?

It is an advantage of the DEL approach that a full sequential model need not be specified from the outset, but a drawback that a complete formalization of the problem under investigation is missing. Ideally, such 'complete models' should be both

1. Computably tractable (for each step), and
2. Informative (model the problem, not just describe the solution).

The first *desideratum* is for implementation purposes. By the second, it is sought that eventual implementations are interesting: models that formalize problems without requiring they be solved first, allows one to draw informative conclusions about the modeled phenomena. The informal approach is typically informative.

This paper suggests a *dynamical systems* approach to specifying 'complete models' of information dynamics and provides some preliminary results.[1] As a (discrete time) dynamical system consists of only a state space $X$ and a map $\tau : X \longrightarrow X$ iteratively applied, the future development of the dynamics depend only on the current state and the map $\tau$. Dynamical systems thus provide a formal container for dynamical models in the local spirit of DEL. This stands in contrast to the only formal alternative, *DEL protocols* [3], which define dynamics globally. This approach is discussed in Section 3.

Dynamical systems are simple but may therefore also be limiting. E.g., if one's chosen model transformer class contains only action models, then the set of scenarios that can be modeled is very narrow: the same action model will be reapplied by the dynamical system, scenarios such as the well-known Muddy Children example [10] are among the unrepresentable phenomena. This provides a motivation for seeking broader classes of model transformers, the topic of Section 5. Three methods for defining complex model transformers are defined, being *multi-pointed action models*, *programs* and *problems*. The main technical results compare these approaches with respect to the orbits they can produce when used in dynamical systems.

## 2   DEL Preliminaries

Let be given a finite, non-empty set of propositional atoms $\Phi$ and a finite, non-empty set of agents, $\mathcal{A}$.

**Definition 1 (Kripke Model).** *A* Kripke model *is a tuple* $M = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket)$ *where*

$\llbracket M \rrbracket$ *is a non-empty set of states;*

$R : \mathcal{A} \longrightarrow \mathcal{P}(S \times S)$ *is an accessibility function;*

$\llbracket \cdot \rrbracket : \Phi \longrightarrow \mathcal{P}(S)$ *is a valuation function.*

*A pair* $(M, s)$ *with* $s \in \llbracket M \rrbracket$ *is called an* epistemic state.

**Definition 2 (Language, Semantics).** *Where* $p \in \Phi$ *and* $i \in \mathcal{A}$, *define a language* $\mathcal{L}$ *by*

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi$$

*with non-propositional formulas evaluated over epistemic state* $(M, s)$ *by*

$$(M, s) \models K_i\varphi \text{ iff } \forall t \in R_i(s), (M, t) \models \varphi.$$

---

[1] The approach to dynamical systems taken here thus differs from that [14], which mainly seeks modal logical descriptions of dynamical system concepts.

With a normal modal logical language like $\mathcal{L}$, the natural notion of equality of epistemic states is *bisimulation*:

**Theorem 1 (Hennessy-Milner, [4], Thm.2.24).** *Let $M$ and $M'$ be image-finite, i.e., $\forall s \in [\![M]\!], \forall i \in \mathcal{A}$, the set $\{t : (s,t) \in R_i\}$ is finite. Then for all $s \in [\![M]\!], s' \in [\![M']\!]$, $s$ and $s'$ are modally equivalent iff $(M,s)$ and $(M',s')$ are bisimilar.*

When working with finite models, $\mathcal{L}$ is strong enough to distinguish any two non-bisimilar models:

**Theorem 2 ([11], Thm.32).** *Let $(M,s)$ and $(M',s')$ be finite epistemic states that are not n-bisimilar. Then there exists $\delta \in \mathcal{L}$ such that $(M,s) \models \delta$ and $(M',s') \not\models \delta$.*

Dynamics are introduced by transitioning from one epistemic state to the next:

**Definition 3 (Model Transformer).** *Let $\mathcal{M}$ be the set of epistemic states based on $\mathcal{A}$. A* model transformer *is a (possibly partial) function $\tau : \mathcal{M} \longrightarrow \mathcal{M}$.*

Several model transformers have been suggested in the literature, the most well-known being public announcement, $!\varphi$ [12]. Primary to this paper is the rich class of action models [2] with postconditions [8].

**Definition 4 (Action Model).** *An* action model *is a tuple $\Sigma = ([\![\Sigma]\!], \mathsf{R}, \mathsf{pre}, \mathsf{post})$ where*

  $[\![\Sigma]\!]$ *is a finite, non-empty set of* actions*;*

  $\mathsf{R} : \mathcal{A} \longrightarrow \mathcal{P}([\![\Sigma]\!] \times [\![\Sigma]\!])$ *is an* accessibility *function;*

  $\mathsf{pre} : [\![\Sigma]\!] \longrightarrow \mathcal{L}$ *is a* precondition *function;*

  $\mathsf{post} : [\![\Sigma]\!] \longrightarrow \{\bigwedge_{i=0}^{n} \varphi_i \not\models \bot : \varphi_i \in \{\top, p, \neg p : p \in \Phi\}\}$ *is a* postcondition *function.*

*A pair $(\Sigma, \sigma)$ with $\sigma \in [\![\Sigma]\!]$ is called an* epistemic action*.*

The precondition of an action $\sigma$ specifies the conditions under which $\sigma$ is executable; the postconditions specify how $\sigma$ sets the values of select atoms. If $\mathsf{post}(\sigma) = \top$, then $\sigma$ changes nothing.

An epistemic state is informationally updated with an epistemic action by taking their product:

**Definition 5 (Product Update).** *The* product update *of epistemic state $(M,s) = ([\![M]\!], R, [\![\cdot]\!], s)$ with epistemic action $(\Sigma, \sigma) = ([\![\Sigma]\!], \mathsf{R}, \mathsf{pre}, \mathsf{post}, \sigma)$ is the epistemic state*

$$(M \otimes \Sigma, (s, \sigma)) = ([\![M \otimes \Sigma]\!], R', [\![\cdot]\!]', (s, \sigma))$$

*where*

  $[\![M \otimes \Sigma]\!] = \{(s, \sigma) \in [\![M]\!] \times [\![\Sigma]\!] : (M,s) \models \mathsf{pre}(\sigma)\}$

  $R'_i = \{((s, \sigma), (t, \tau)) : (s,t) \in R_i \text{ and } (\sigma, \tau) \in \mathsf{R}_i\}$

  $[\![p]\!]' = \{(s, \sigma) : s \in [\![p]\!], \mathsf{post}(\sigma) \not\models \neg p\} \cup \{(s, \sigma) : \mathsf{post}(\sigma) \models p\}$.

In combination, an epistemic action $(\Sigma, \sigma)$ and product update $\otimes$ thus define a model transformer. Denote the class of such transformers by $\boldsymbol{\Sigma}$. Each $\tau \in \boldsymbol{\Sigma}$ has the following pleasant property:

**Fact (Bisimulation Preservation).** $\forall \tau \in \boldsymbol{\Sigma}$, if $(M, s)$ and $(M', s')$ are bisimilar, then so are $\tau(M, s)$ and $\tau(M', s')$.

$\boldsymbol{\Sigma}$ is a very powerful class: for any finite epistemic state $(M, s)$, it contains a transformer that will map $(M, s)$ to any other finite epistemic state $(M', s')$, as long as no agents with empty access in $M$ has non-empty access in $M'$ and as long as $M$ and $M'$ differ only in the truth value of a finite number of atoms. The restrictions are due to the 'and'-condition used in defining $R_i'$ in product update and the finite conjunction used in defining postcondition maps. If the directed relation given by these restrictions holds from $(M, s)$ to $(M', s')$, then call the transition from the first to the second *reasonable*:

**Definition 6 (Reasonable Transition).** *Let* $(M, s) = (\llbracket M \rrbracket, R, V, s)$ *and* $(M', s') = (\llbracket M' \rrbracket, R', V', s')$ *be two epistemic states. Then the transition from* $(M, s)$ *to* $(M', s')$ *is* reasonable *iff*

1. *it* preserves insanity*: there exists a submodel* $M^{\mathsf{s}}$ *of* $M$ *such that* $s \in \llbracket M^{\mathsf{s}} \rrbracket$ *and* $\forall i \in \mathcal{A}$, *if* $R_i' \neq \emptyset$, *then* $R_i$ *is serial in* $M^{\mathsf{s}}$, *and*
2. *it* invokes *finite ontic change*:

$$\{p : \llbracket p \rrbracket \neq \emptyset \text{ and } \llbracket p \rrbracket \neq \llbracket M \rrbracket\}$$
$$\cup \{p : \llbracket p \rrbracket = \emptyset\} \setminus \{p : \llbracket p \rrbracket' = \emptyset\}$$
$$\cup \{p : \llbracket p \rrbracket = \llbracket M \rrbracket\} \setminus \{p : \llbracket p \rrbracket' = \llbracket M' \rrbracket\}$$

*is finite.*

**Theorem 3 (Arbitrary Change, [8], Prop.3.2).** *Let the transition from finite* $(M, s)$ *to finite* $(M', s')$ *be reasonable. Then there exists a* $(\Sigma, \sigma) \in \boldsymbol{\Sigma}$ *such that* $(M, s) \otimes (\Sigma, \sigma)$ *and* $(M', s')$ *are bisimilar.*

## 3  DEL Protocols

One framework which could be used to construct 'complete models' is *DEL protocols* [3,7,13,15].

**Definition 7 (DEL Protocol).** *Let* $\boldsymbol{\Sigma}^*$ *be the set of all finite sequences of transformers* $\tau \in \boldsymbol{\Sigma}$. *A set* $\mathsf{P} \subseteq \boldsymbol{\Sigma}^*$ *is a (uniform)* DEL protocol *iff* $\mathsf{P}$ *is closed under non-empty prefixes.*

A DEL protocol specifies which model transformers *may* be executed at a given time—whether they *can* be executed depends on the model transformers, e.g. their preconditions.

Where $\mathsf{P}$ is a DEL protocol and $\sigma = (\tau_1, ..., \tau_n) \in \mathsf{P}$, set $(M, s)^{\sigma} := \tau_n \circ \cdots \circ \tau_1(M)$. From an initial model $(M, s)$ and time 0, a DEL protocol $\mathsf{P}$ produces a set of possible evolutions to each time $n$, namely

$\{(M,s)^\sigma : \mathsf{len}(\sigma) = n\}$. Notice that $\mathsf{len}(\sigma) = n$ does not imply that $(M,s)^\sigma$ exists: one of the transformers from $\sigma$ may have been unexecutable at some earlier stage.

DEL protocols are dismissed as suitable for constructing 'complete models' as the results will be unexecutable, incorrect or uninformative. To see this, assume that some phenomenon that involves multiple model transformers $T = \{\tau_1, ..., \tau_n\}$, as e.g. Muddy Children does.

If the DEL protocol used is $T^*$ (the set of all finite strings sequences of transformers from $T$) a very nice model is obtained: it is applicable to multiple initial states with varying mud distributions, and it may accordingly be used to obtain answers to questions about e.g. how the scenario unfolds as a function of the number of muddy children. Alas, $T^*$ is infinite and as a model therefore unexecutable: given some initial state $(M,s)$ it will not be possible to run $T^*$ on $(M,s)$ in finite time as the input to any function that is to determine the set $\{(M,s)^\sigma : \mathsf{len}(\sigma) = 1\}$ will be infinite.

To obtain an executable model, $T^*$ could be pruned to obtain a finite DEL protocol $\mathsf{T} \subseteq T^*$, e.g. by setting some upper bound on the length of $\sigma \in \mathsf{T}$. The risk associated with this move (pruning) is that the model becomes useless or uninformative: if the upper bound is set too low, the model will terminate too soon and not provide a correct output; to ensure the upper bound high enough, the problem must have been solved beforehand, leading to an uninformative model. In the extreme case where the only included maximal $\sigma$ is 'the correct one' given some natural language protocol and initial state, a descriptive model is produced, but such a 'gold in, gold out' model is of little interest from an investigative perspective.

## 4   DEL and Dynamical Systems

Given Theorem 3, one might expect that dynamical systems based on the class of action models $\boldsymbol{\Sigma}$ would allow modeling of a plethora of phenomena. Surprisingly, not even even simple and well-known epistemic puzzles such as Muddy Children can be modeled by this class. To see this, let us first clarify the notion of dynamical system.

As standardly defined [6], a dynamical system is a tuple $D = (X, T, \mathcal{E})$ where $X$ is set, called the *state space*, $T \subseteq \mathbb{R}$ is a *time set* which forms an additive semi-group $(t_1, t_2 \in T \Rightarrow t_1 + t_2 \in T)$ and $\mathcal{E} : X \times T \to X$ is an *evolution map* satisfying that $\mathcal{E}(x,0) = 0$ and $\mathcal{E}(\mathcal{E}(x,t_1),t_2) = \mathcal{E}(x, t_1 + t_2)$.

To obtain a state space for DEL-based dynamical systems, it is natural, given Theorem 1, to equate bisimilar epistemic states, and let the state space consist of each bisimulation type's smallest representative. For an epistemic state $(M,s)$, this representative is given by $(M,s)$'s *generated submodel rooted at $s$'s bisimulation quotient* $(M[s]/\rho^M, [s]_\rho^M)$, see [11], Sec. 3.6. Setting

$$\mathbf{M} := \{(M[s]/\rho^M, [s]_\rho^M) : (M,s) \text{ is an epistemic state}\},$$

a class is obtained that contains a canonical representative of each epistemic state, each unique up to isomorphism.

As DEL updates are discrete and non-invertible, the suitable time set for a DEL-based dynamical system is $\mathbb{Z}_+$. The evolution function of any dynamical system $D = (X, \mathbb{Z}_+, \mathcal{E})$ with time set $\mathbb{Z}_+$ may be defined by the iterations of a function $e : X \to X$ by $\mathcal{E}(x, n) = e^n(x)$. Given the chosen state space, the suitable class of such functions $e$ is the set of model transformers $\tau : \mathbf{M} \to \mathbf{M}$, denoted by $\mathbf{T}$.

Given these considerations, the following definition of DEL-based dynamical systems is obtained:

**Definition 8 (DEL-based Dynamical System).** *A* DEL-based dynamical system *is a pair* $D = (\mathbf{X}, \tau)$ *where* $\mathbf{X} \subseteq \mathbf{M}$ *and* $\tau : \mathbf{X} \to \mathbf{X}$.
*The* orbit *of D from* initial state $x_0 \in \mathbf{X}$ *is the sequence* $o(D, x_o) = (\tau^n(x_0))_{n \in \mathbb{Z}_+}$.

**Remark.** Given an epistemic action $\tau \in \mathbf{\Sigma}$, $x \in \mathbf{M}$ does not imply that $\tau(x) \in \mathbf{M}$. There will however be a $x' \in \mathbf{M}$ that is bisimilar to $\tau(x)$. Given Fact 1, each $\tau \in \mathbf{\Sigma}$ may be identified with a $\tau' \in \mathbf{T}$ by if $\tau(x) = (M, s)$, then $\tau'(x) = (M[s]/\rho^M, [s]_\rho^M)$. Henceforth, when executing an epistemic action $(\Sigma, \sigma)$ in $x \in \mathbf{M}$, it is thus assumed that $x \otimes (\Sigma, \sigma) \in \mathbf{M}$.

It is immediately clear that any dynamical system $D = (\mathbf{X}, \tau)$ with $\tau \in \mathbf{\Sigma}$ will be limited in its orbits. In particular, where $s_0$ is the actual state in the initial epistemic state $x_0$ and $\sigma_0$ is the actual state of $\tau$, then for any $n$, the actual state of $\tau^n(x_0)$ will be of the form $(...(s_0, \sigma_0), ..., \sigma_0)$. Consequently, any phenomenon that involves the occurrence of more than one actual action is unmodelable. As most phenomena do involve shift in the performed action, e.g. by a shift in the announcement made, there is a motivation for seeking out a more general class of model transformers.

## 5   Complex Model Transformers

The limitation of DEL-based dynamical systems does not stem from action models, but rather from the fact that their usage is not controlled. This problem is solved by DEL protocols or update streams; simply specify at which time which action model should be executed. However, this requires a description of the evolution before execution, leaving little of the local DEL spirit intact.

A natural way to specify which transformer should be applied next that still remains local in spirit is by using a map $\pi : \mathbf{M} \longrightarrow \mathbf{T}$. Composing such a $\pi$ with the model transformers it picks at each epistemic state is then again a model transformer $\tau_\pi : \mathbf{M} \longrightarrow \mathbf{M}$ given by $\tau_\pi(x) = \pi(x)(x)$.

To be interesting from modeling and implementation perspectives, such $\pi$ must be finitely representable. This puts constraints on the dynamical systems definable, but, as will be shown, the restriction is still to a vast class of such systems.

We focus on three ways of specifying maps $\pi$, each picking model transformers from $\mathbf{\Sigma}$. The choice to restrict attention to maps picking transformers from

$\Sigma$ is warranted by Theorem 3: As basic transformers, this class has sufficient transformational power to construct a rich class of dynamical systems.

The first type is closely related to the (knowledge-based) programs known from interpreted systems [10], though defined to specify transformers based on the global, epistemic state rather than specifying sub-actions based on agents' local states:[2]

**Definition 9 (Program).** *A (finite, deterministic,$(\mathcal{L}, \Sigma)$) program is a finite set of formula-transformer pairs*

$$P = \{(\varphi_i, \tau_i) : \varphi_i \in \mathcal{L}, \tau_i \in \Sigma\}$$

*where $\forall i, j$ if $\varphi_i \neq \varphi_j$ and $(\varphi_i, \tau_j), (\varphi_j, \tau_j) \in P$, then $\mathbf{M} \models \varphi_i \wedge \varphi_j \rightarrow \bot$.*
*Each program $P$ gives rise to a model transformer $\tau_P$ given by $\tau_P(x) = \tau_i(x)$ if $x \models \varphi_i$ and $(\varphi_i, \tau_i) \in P$. Denote this class by $\mathbf{P}$.*

Each program may be read as a set of conditional tests of the form **if** $\varphi_i$, **do** $\tau_i$, in form similar to the informal specifications often used in DEL literature.

The explicit specification of programs stands in contrast with the implicit specification of the second transformer type, *problems*, where each instruction may be read **if** $\varphi_i$, **obtain** $\psi_i$. Problems as defined here are related to *epistemic planning problems*, also know from the DEL literature [5].

**Definition 10 (Problem).** *A (finite $(\mathcal{L}, \Sigma)$) problem is a pair*

$$\Pi = (Q, \Sigma_\Pi)$$

*where $Q = \{(\varphi_i, \psi_i) : \varphi_i, \psi_i \in \mathcal{L}\}$ is a finite set of formula-formula pairs and $\Sigma_\Pi \subset \Sigma$ is a finite set of model transformers with an associated strict order $<$.*

*A solution to $\Pi = (Q, T)$ at epistemic state $x$ is a model transformer $\tau \in T$ such that $\forall (\varphi_i, \psi_i) \in Q$, if $x \models \varphi_i$, then $\tau(x) \models \psi_i$. Denote the set of solution to $\Pi$ at $x$ by $\Pi(x)$.*

*Each problem $\Pi$ gives rise to a model transformer $\tau_\Pi$ given by $\tau_\Pi(x) = \min_< \Pi(x)$. Denote this class by $\mathbf{\Pi}$.*

The model transformer $\tau_\Pi$ is defined using the strict order $<$ on $\Sigma_\Pi$ to ensure that $\tau_\Pi$ is a function: nothing in the definition ensures that $|\Pi(x)| \leq 1$.

The last model transformer type to be considered is a slight generalization of action models [1], where each such may have multiple actual states. In the definition it is required, non-standardly, that the preconditions of the actual states must be mutually exclusive. This is to ensure that executing a multi-pointed action model using product update remains a single-pointed epistemic state.

---

[2] Programs based on agents' local states is also at least to some degree feasible in a DEL setting, using *parallel action model composition* [9].

**Definition 11 (Multi-Pointed Epistemic Actions).** *A (finite, deterministic) multi-pointed epistemic action is an epistemic action $(\Sigma, \sigma)$ with $\sigma$ replaced by a finite, non-empty set $S \subseteq [\![\Sigma]\!]$, where for each $\sigma, \sigma' \in S$, if $\sigma \neq \sigma'$, then $\mathbf{M} \models pre(\sigma) \wedge pre(\sigma') \rightarrow \bot$.*

*Applied using product update, each $(\Sigma, S)$ is a model transformer $\tau : (M \otimes \Sigma, (s, S)) \mapsto ([\![M \otimes \Sigma]\!], R', [\![\cdot]\!]', (s, \sigma_i))$ where $(M, s) \models pre(\sigma_i)$. Denote this class by $\boldsymbol{\Sigma}^+$.*

With mutually exclusive preconditions, a multi-pointed action model $(\Sigma, S)$ encodes a map $\pi : \mathbf{M} \longrightarrow \mathbf{T}$ with image $\{(\Sigma, \sigma) : \sigma \in S\}$ by $\pi(x) = (\Sigma, \sigma)$, $x \models pre(\sigma)$.

## 6  Results

Note initially that DEL-based dynamical systems fair better than DEL protocols in regard to executability and informativity. DEL-based dynamical systems resting on either a program or a multi-pointed action model are step-wise computable, as both transformer types are finite and therefore require only check of a finite set of formulas at each $(M, s)$. The case for problems must be checked against [5]. Moreover, DEL-based dynamical systems will provide informative models: once a system is defined, one may start investigating how its orbits behave as a function of initial state without having pre-solved the encoded problem.

The first main result shows that dynamical systems based on the class $\boldsymbol{\Pi}$ of problem-based model transformers can model any reasonable, deterministic, finite or cyclic sequence of finite epistemic states. Problem-based dynamical systems can thus model a large class of phenomena.

The proof of Proposition 1 is by brute force. The construction results in a large, cumbersome problem fully pre-encoding the target orbit. For many modeling purposes, far more economical complex model transformers will do.

**Definition 12 (Finite Variation, Deterministic).** *Let $\overline{x} = (x_0, x_1, ...)$ be a sequence of epistemic states from $\mathbf{M}$. $\overline{x}$ has* finite variation *iff*

1. *$\overline{x}$ is finite, or*
2. *$\exists n, m, k \in \mathbb{Z}_+ \backslash \{0\} : x_k = x_{k+m}$ for all $k \geq n$.*

*$\overline{x}$ is* deterministic *iff if $x_k, x_{k+1}, x_m \in \overline{x}$ and $x_k = x_m$, then $x_{m+1} \in \overline{x}$ and $x_{k+1} = x_{m+1}$.*

**Proposition 1 (Arbitrary Orbits).** *Let the sequence $\overline{x} = (x_0, x_1, ...)$ of finite epistemic states be deterministic, with finite variation and where the transition between each $x_i$ and $x_{i+1}$ is reasonable. Then there exists a dynamical system $D = (\mathbf{M}, \tau_\Pi)$ with $\tau_\Pi \in \boldsymbol{\Pi}$ such that $o(D, x_0) = \overline{x}$.*

*Proof.* By constructing a problem $\Pi = (Q, \Sigma_\Pi)$ that gives rise to the sought $\tau_\Pi$.

For each $x_i, x_j \in \overline{x}$, $x_i \neq x_j$, let $\delta_{i,j}$ be a formula that distinguishes $x_i$ from $x_j$ such that $x_i \models \delta_{i,j}$ and $x_j \not\models \delta_{i,j}$; this $\delta_{i,j}$ exists by Theorem 2. As $\overline{x}$ has

finite variation, $\delta_i := \bigwedge_{j:x_j \in \overline{x} \setminus \{x_i\}} \delta_{i,j}$ is a formula that distinguishes $x_i$ from all other $x_j \in \overline{x}$. For each $x_i, x_{i+1} \in \overline{x}$, let $\tau_i \in \Sigma$ be a model transformer such that $\tau_i(x_i) = x_{i+1}$; this exists by Theorem 3.

Let $Q$ be the smallest set that for each $x_i, x_{i+1} \in \overline{x}$ contains $(\delta_i, \delta_{i+1})$. Let $\Sigma_\Pi$ be the smallest set that for each $x_i, x_{i+1} \in \overline{x}$ contains $\tau_i$. Both $Q$ and $\Sigma_\Pi$ are finite by the assumption of finite variation, so $\Pi = (Q, \Sigma_\Pi)$ is a finite program, so $\tau_\Pi$ is a model transformer.

That $o(D, x_0) = \overline{x}$ when $D = (\mathbf{M}, \tau_\Pi)$ is shown by induction on $x_n$:
*Base:* $\tau_\Pi{}^0(x_0) = x_0$. *Step:* Assume $\tau_\Pi{}^n(x_0) = x_n$. If $\overline{x} = (x_0, ..., x_n)$, then $o(D, x_0) = \overline{x}$ as $(\delta_n, \varphi) \notin Q$ for any $\varphi$, by determinism of $\overline{x}$, so $\tau_\Pi(x_n)$ is undefined. If $x_{n+1} \in \overline{x}$, then $(\delta_n, \delta_{n+1}) \in Q$ and $\tau_n \in \Sigma_\Pi$. By construction, $\Pi(x_n) = \tau_n$, so $\tau_\Pi(x) = x_{n+1}$.                               □

**Proposition 2 (Problem Orbit Properties).** *Let $o(D, x_0) = \overline{x}$ with $D = (\mathbf{M}, \tau_\Pi)$, $\tau_\Pi \in \mathbf{\Pi}$. Then $\overline{x}$ is deterministic and for each $x_i, x_{i+1} \in \overline{x}$, the transition from $x_i$ to $x_{i+1}$ is reasonable.*

*Proof.* $\overline{x}$ is deterministic as $\tau_\Pi$ is a function; each transition is reasonable as $x_{i+1} = \tau(x_i)$ for some $\tau \in \Sigma$.

Propositions 1 and 2 cannot be strengthened to a characterization result as not all problem-based dynamical system have finite variation:

**Proposition 3 (Infinite Variation).** *There exists a dynamical system $D = (\mathbf{M}, \tau_\Pi)$ with $\tau_\Pi \in \mathbf{\Pi}$ such that $o(D, x_0)$ does not have finite variation.*

*Proof.* Let $D = (\mathbf{M}, \tau_\Pi)$ with problem $\Pi = (\{(\top, \top)\}, \{(\Sigma, \sigma_1)\})$. This trivial problem has unique solution $(\Sigma, \sigma_1)$ for all $(M, s) \in \mathbf{M}$. Hence, for all $x \in \mathbf{M}$, $\tau_\Pi(x) = (M, s) \otimes (\Sigma, \sigma_1)$.

Let $M$ and $\Sigma$ given by

$$M: \quad s: \ \overset{s:}{\circlearrowleft \boxed{\phantom{x}}} \overset{t:}{\longleftrightarrow \boxed{p}} \qquad \Sigma: \quad \overset{\sigma_1:}{\circlearrowleft} \overset{}{\boxed{\neg p, \top}} \overset{\sigma_2:}{\longrightarrow \boxed{\neg p, p}} \overset{\sigma_3:}{\longrightarrow \boxed{p, \top}} \circlearrowright$$
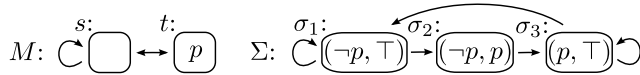
Then $o(D, (M, s))$ does not have finite variation: for each iteration of $\tau_\Pi$, the state not satisfying $p$ will split, inserting a new $p$ state as it's child with $\sigma_2$:

$$(M, s) \otimes (\Sigma, \sigma_1): \quad \overset{s\sigma_1:}{\circlearrowleft \boxed{\phantom{x}}} \overset{s\sigma_2:}{\longrightarrow \boxed{p}} \overset{t\sigma_3:}{\longrightarrow \boxed{p}}$$
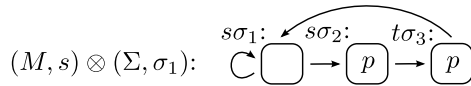
All other states have only one child, with $\sigma_3$.

In all further applications of $(\Sigma, \sigma_1)$, the circular structure seen in $(M, s) \otimes (\Sigma, \sigma_1)$ is preserved, only with an additional $p$ state. No two such models are bisimilar, and hence the orbit does not have finite variation.                               □

The second main result shows that also program-based dynamical systems and dynamical systems based on multi-pointed action models can produce a vast class of orbits.

**Proposition 4 (Equivalence).** *Let $\overline{x} = (x_0, x_1, ...)$ be a sequence of epistemic states. Then*

1. $\exists \tau_\Pi \in \mathbf{\Pi}$ *such that for* $D = (\mathbf{M}, \tau_\Pi)$, $o(D, x_0) = \overline{x}$.

   $\Uparrow$

2. $\exists \tau_P \in \mathbf{P}$ *such that for* $D = (\mathbf{M}, \tau_P)$, $o(D, x_0) = \overline{x}$.

   $\Updownarrow$

3. $\exists \tau_{\Sigma^+} \in \mathbf{\Sigma}^+$ *such that for* $D = (\mathbf{M}, \tau_{\Sigma^+})$, $o(D, x_0) = \overline{x}$.

*If $\overline{x} = (x_0, x_1, ...)$ has finite variation and $x_0$ is finite, then the three statements are equivalent.*

*Proof.*
**Case: 2. $\Rightarrow$ 1.** Let $D = (\mathbf{M}, \tau_P)$, $\tau_P \in \mathbf{P}$ with $o(D, x_0) = \overline{x} = (x_0, x_1, ...)$ be given.

Construct a problem $\Pi = (Q, \Sigma_\Pi)$ as follows: Let $Q$ be the smallest set that for each $(\varphi_i, \tau_i) \in P$ contains $(\varphi_i, \top)$. Let $\mathbf{\Sigma}_\Pi$ be the smallest set that for each $(\Sigma, \sigma) \in \mathbf{\Sigma}_P$ contains $(\Sigma, \sigma^*)$ identical to $(\Sigma, \sigma)$ in all respects except that $pre(\sigma^*) = pre(\sigma) \wedge \varphi_i$. As $P$ is finite, $\Pi = (Q, \Sigma_\Pi)$ is a finite problem; $\tau_\Pi$ is a model transformer as the $\varphi_i$'s of $P$ are mutually exclusive.

Then $o((\mathbf{M}, \tau_\Pi), x_0) = o((\mathbf{M}, \tau_P), x_0)$: Assume $x_i, x_{i+1} \in \overline{x}$. Then $x_{i+1} = \tau(x_i)$ for some $\tau = (\Sigma, \sigma)$ such that for some $\varphi$, $(\tau, \varphi) \in P$. Hence for some $\varphi, (\tau, \varphi) \in P$, it holds that $x_i \models \varphi$. Given the preconditions and that $(\varphi, \top) \in Q$, $\tau^* = (\Sigma, \sigma^*) \in \mathbf{\Sigma}_\Pi$ will be the only solution to $\Pi$ at $x_i$. As $x_i \models \varphi$, $\tau^*(x_i) = \tau(x_i)$.

Assume $\overline{x} = (x_0, ..., x_n)$ is finite. Then either $x_n \not\models \varphi_i$ for all $(\varphi_i, \tau_i) \in P$ or if $x_n \models \varphi_i$ for $(\varphi_i, (\Sigma, \sigma)) \in P$, then $x_n \not\models pre(\sigma)$. In the first case, $x_n \not\models \varphi_i$ for all $(\varphi_i, \top) \in Q$; in the second, $x_n \not\models pre(\sigma^*)$. In either case, $\tau_\Pi(x_n)$ is undefined.

**Case: 2. $\Rightarrow$ 3.** Let $D = (\mathbf{M}, \tau_P)$, $\tau_P \in \mathbf{P}$ with $o(D, x_0) = \overline{x} = (x_0, x_1, ...)$ be given. Let $\mathbf{\Sigma}_\Pi$ be as in the case 2. $\Rightarrow$ 1. Define a multi-pointed action model $(\Sigma^+, S)$ by $\Sigma^+ = \biguplus \{ \Sigma : (\Sigma, \sigma^*) \in \mathbf{\Sigma}_\Pi \}$ and $S = \{ \sigma^* : (\Sigma, \sigma^*) \in \Sigma_\Pi \}$. Let $\tau_{\Sigma^+}$ be the associated model transformer.

Then $o((\mathbf{M}, \tau_{\Sigma^+}), x_0) = o((\mathbf{M}, \tau_P), x_0)$: Assume $x_i, x_{i+1} \in \overline{x}$. Then $x_{i+1} = \tau(x_i)$ for some $\tau = (\Sigma, \sigma)$ such that for some $\varphi$, $(\tau, \varphi) \in P$. Hence for some $\varphi, (\tau, \varphi) \in P$, it holds that $x_i \models \varphi \wedge pre(\sigma)$, so by construction, $x_i \models pre(\sigma^*)$. Hence only the submodel $(\Sigma, \sigma^*)$ of $\Sigma^+$ is executable at $x_i$, so $\tau_{\Sigma^+}(x_i) = \tau_P(x_i)$.

If $\overline{x} = (x_0, ..., x_n)$ is finite, then either $x_n \not\models \varphi_i$ for all $(\varphi_i, \tau_i) \in P$ or if $x_n \models \varphi_i$ for $(\varphi_i, (\Sigma, \sigma)) \in P$, then $x_n \not\models pre(\sigma)$. In the first case, $x_n \not\models pre(\sigma^*)$ for all $(\Sigma, \sigma^*) \in \Sigma^+$; in the second, $x_n \not\models pre(\sigma^*)$. In either case, $\tau_{\Sigma^+}(x_n)$ is undefined.

**Case: 3. $\Rightarrow$ 2.** Let $D = (\mathbf{M}, \tau_{\Sigma^+})$, $\tau_{\Sigma^+} \in \mathbf{\Sigma}^+$ with $o(D, x_0) = \overline{x} = (x_0, x_1, ...)$ be given. Let the $\Sigma^+$ of $\tau_{\Sigma^+}$ be $\Sigma^+ = (\Sigma, S)$ and create from it a set of $|S|$ single-pointed action models $A = \{ (\Sigma, \sigma) : \sigma \in S \}$. Create a program $P = \{ (pre(\sigma), (\Sigma, \sigma)) : (\Sigma, \sigma) \in A \}$. $P$ is both finite and deterministic.

Then $o((\mathbf{M}, \tau_P), x_0) = o((\mathbf{M}, \tau_{\Sigma^+}), x_0)$: Assume $x_i, x_{i+1} \in \overline{x}$. Then $x_i \models pre(\sigma)$ for exactly one $\sigma \in S$. As $(pre(\sigma), (\Sigma, \sigma)) \in P$, $\tau_P(x_i) = \tau_{\Sigma^+}(x_i)$.

If If $\overline{x} = (x_0, ..., x_n)$ is finite, then $x_n \not\models pre(\sigma)$ for all $\sigma \in S$. Hence for all $(\varphi, \tau) \in P$, $x_n \not\models \varphi$, so $\tau_P(x_n)$ is undefined.

**Case:** 1. $\Rightarrow$ 2., *if $\overline{x} = (x_0, x_1, ...)$ has finite variation and $x_0$ is finite*: Let $D = (\mathbf{M}, \tau_\Pi)$, $\tau_\Pi \in \mathbf{\Pi} = (Q, \Sigma_\Pi)$ with $o(D, x_0) = \overline{x} = (x_0, x_1, ...)$ having finite variation. Brute force construct a program using characteristic formulas: let $\delta_i$ be the characteristic formula of $x_i \in \overline{x}$. For each pair $x_i, x_{i+1} \in \overline{x}$, there is a unique $\tau_i \in \Sigma_\Pi$ such that $\tau_i(x_i) = x_{i+1}$. Let $P = \{(\delta_i, \tau_i) : x_i \in \overline{x}\}$. As $\overline{x}$ has finite variation, $P$ is finite and gives rise to a model transformer $\tau_P$.

Then $o((\mathbf{M}, \tau_P), x_0) = o((\mathbf{M}, \tau_\Pi), x_0)$: Assume $x_i, x_{i+1} \in \overline{x}$. Then $(\delta_i, \tau_i) \in P$, so $\tau_P(x_i) = x_{i+1}$. If $\overline{x} = (x_0, ..., x_n)$ is finite, then by Proposition 2, for no $x_i, i < n$ is $x_i = x_n$. Hence $(\delta_n, \tau) \notin P$, for any $\tau$. Hence $\tau_P(x_n)$ is undefined.                                            $\square$

**Corollary 1 (Orbit Properties).** *For any dynamical system $D = (\mathbf{M}, \tau)$ with $\tau \in \mathbf{P} \cup \mathbf{\Sigma}^+$ and any $x_0 \in \mathbf{M}$, $o(D, x_0)$ is deterministic and for each $x_i, x_{i+1} \in \overline{x}$, the transition from $x_i$ to $x_{i+1}$ is reasonable.*

*Proof.* Let $D$ be as described. By Proposition 4 there exists a $D' = (\mathbf{M}, \tau_\Pi)$, $\tau_\Pi \in \mathbf{\Pi}$, that recreates $o(D, x_0)$. The corollary then follows from Proposition 2.

# 7    Conclusion

The main contributions are

▷ that although dynamical systems defined using epistemic action models can produce only very limited orbits, dynamical systems that control when particular action models are used may produce orbits sufficient for most modeling purposes, and

▷ that the three methods for controlling which action models are applied are equivalent under the presented conditions.

The first result shows that DEL-based dynamical systems provide a rich framework for producing mathematically specified models of information dynamics. The latter shows that there are multiple ways of extending the DEL toolbox compatible with modeling using dynamical systems.

It would be interesting to make an in-depth comparison between DEL protocols and DEL-based dynamical systems, comparing the orbits they may produce and under which conditions such might be equivalent. Two considerations here involve the finite nature of DEL protocols, guaranteeing finite variation not guaranteed by DEL-based dynamical systems, and the 'bisimulation respecting' behavior of DEL-based dynamical systems, which is not necessarily followed by DEL protocols. Obtaining such results could be used to link DEL-based dynamical systems with Epistemic Temporal Logic via the results in [3].

Moreover, it would be interesting to investigate any deeper relationship between dynamic epistemic logic and dynamical systems; the latter field is well-developed, and one could envision that methods and results may be transferable.

# References

1. Baltag, A., Moss, L.S.: Logics for Epistemic Programs. Synthese 139(2), 165–224 (2004)
2. Baltag, A., Moss, L.S., Solecki, S.: The Logic of Public Announcements, Common Knowledge, and Private Suspicions (extended abstract). In: Proc. of the Intl. Conf. TARK 1998, pp. 43–56. Morgan Kaufmann Publishers (1998)
3. van Benthem, J., Gerbrandy, J., Hoshi, T., Pacuit, E.: Merging Frameworks for Interaction. Journal of Philosophical Logic 38(5), 491–526 (2009)
4. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press (2001)
5. Bolander, T., Birkegaard, M.: Epistemic planning for single- and multi-agent systems. Journal of Applied Non-Classical Logics 21(1), 9–34 (2011)
6. Broer, H.W., Takens, F.: Preliminaries of Dynamical Systems Theory. In: Hasselblatt, B., Broer, H.W., Takens, F. (eds.) Handbook of Dynamical Systems, North-Holland, vol. 3 (2010)
7. Dégremont, C.: The Temporal Mind: Observations on the logic of belief change in interactive systems. PhD thesis, University of Amsterdam (2010)
8. van Ditmarsch, H., Kooi, B.: Semantic Results for Ontic and Epistemic Change. In: Bonanno, G., van der Hoek, W., Wooldridge, M. (eds.) Logic and the Foundations of Game and Decision Theory (LOFT 7). Texts in Logic and Games, vol. 3, pp. 87–117. Amsterdam University Press (2008)
9. van Eijck, J., Sietsma, F., Wang, Y.: Composing models. Journal of Applied Non-Classical Logics 21(3-4), 397–425 (2011)
10. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning About Knowledge. The MIT Press (1995)
11. Goranko, V., Otto, M.: Model Theory of Modal Logic. In: Blackburn, P., van Benthem, J., Wolter, F. (eds.) Handbook of Modal Logic, Elsevier (2008)
12. Plaza, J.A.: Logics of public communications. In: Emrich, M.L., Pfeifer, M.S., Hadzikadic, M., Ras, Z.W. (eds.) Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems, pp. 201–216 (1989)
13. Rodenhäuser, B.: A logic for extensional protocols. Journal of Applied Non-Classical Logics 21(3-4), 477–502 (2011)
14. Sarenac, D.: Modal Logic for Qualitative Dynamics. In: Roy, O., Girard, P., Marion, M. (eds.) Dynamic Formal Epistemology. Synthese Library, vol. 351, pp. 75–101. Springer (2011)
15. Wang, Y.: Epistemic Modelling and Protocol Dynamics. Doctoral thesis, Universiteit van Amsterdam (2010)

# Paper III

# Intensional Protocols for Dynamic Epistemic Logic

**Hanna S. van Lee · Rasmus K. Rendsvig ·
Suzanne van Wijk**

**Abstract** In dynamical multi-agent systems, agents are controlled by protocols. In choosing a class of formal protocols, an implicit choice is made concerning the types of agents, actions and dynamics representable. This paper investigates one such: An intensional protocol class for agent control in Dynamic Epistemic Logic (DEL), called 'DEL dynamical systems'. After illustrating how such protocols may be used in formalizing and analyzing information dynamics, the types of epistemic temporal models that they may generate are characterized. This facilitates a formal comparison with the only other formal protocol framework in Dynamic Epistemic Logic, namely the extensional 'DEL protocols'. The paper is concluded with a conceptual comparison, highlighting modeling tasks where DEL dynamical systems are natural.

**Keywords** dynamic epistemic logic, multi-agent systems, protocols, epistemic temporal logic, dynamical systems

## 1 Introduction

In logically modeling dynamics in multi-agent systems – whether by global-perspective frameworks like Interpreted Systems [21] and Epistemic Temporal Logic [36], or by local-perspective frameworks like Dynamic Epistemic Logic [4] – the dynamics rely on *protocols*: control mechanisms that determine which actions may occur when.

---

Hanna S. van Lee
Center for Information and Bubble Studies, University of Copenhagen

Rasmus K. Rendsvig
Theoretical Philosophy, Lund University

Suzanne van Wijk
Independent

Protocols take a plethora of forms, ranging from natural language descriptions, over pseudo-code renderings, to fully formalized representations. Moreover, protocol specifications may vary in their fundamental structure. Specifically, one may distinguish between *extensional* protocols and *intensional* protocols.[1]

Extensional protocols are *temporal*: They consult an *external clock* to specify which actions are available for execution at a given time of a run of a system. Roughly speaking, an extensional protocol is a set of sequences of actions that allows the execution of action $a$ at time $t$ if $a$ is on the $t$th position of a sequence in the protocol. Abstractly, think of a function assigning to each natural number a set of allowed actions.

Intensional protocols, in contrast, are *conditional*: They consult the *current state of the system* to specify which actions are available for execution *now*. Roughly speaking, a conditional protocol is a set of "if $\varphi$, then do $a$" statements. Such a statement – or rule – allow the execution of action $a$ *now* if the current state satisfies the test condition $\varphi$. Abstractly, think of a function assigning to each possible state of the system some set of allowed actions.

Both extensional and intensional protocols *qua* protocols have been investigated in the epistemic agency literature, but mainly in different paradigms: Where the Interpreted Systems literature has favored intensional protocols [21, 44, 34], the literature on protocols in Dynamic Epistemic Logic has favored extensional protocols [13, 18, 29, 30, 43, 41].

There is, however, no formal reason to avoid intensional protocols in the Dynamic Epistemic Logic setting. In fact, such protocols may be both intuitive and compact in representation. Moreover, the mathematical basis and logical theory for intensional protocols enjoys established results, albeit not cast as results concerning protocols (cf. Sec. 1.2 on related literature).

This paper concerns intensional protocols for Dynamic Epistemic Logic (DEL). In particular, it investigates intensional protocols as represented by *multi-pointed action models* applied iteratively. By this, the paper takes a *discrete-time dynamical systems* perspective on protocols for information dynamics. The resulting intensional protocols are referred to as *DEL dynamical systems*.

The overarching question of the paper is how such intensional protocols relate to their closest extensional relative, namely the DEL protocols of van Benthem, Gerbrandy, Hoshi and Pacuit, [13]. The main motivation for this question is a wish to clarify similarities and differences in the implicit assumptions and restrictions inherent in the two frameworks. This, in turn, is motivated by a desire to understand up- and downsides of protocol frameworks from a design and modeling perspective.

Methodologically, the main comparison is achieved by characterizing the types of Epistemic Temporal Logic (ETL) models generatable by intensional protocols coded as DEL dynamical systems, and compare the resulting ETL

---

[1] The terms and distinction is adopted from Parikh and Ramunjam, see [36, Sec. 2.2].

properties with those previously obtained for extensional DEL protocols by van Benthem, Gerbrandy, Hoshi and Pacuit.

This methodology has a two-fold incentive, foremost of which is that ETL models provide an assumption-free common point of reference between the two protocol forms, thus allowing a comparison of induced properties.[2] This is desirable as the fundamental difference between extensional and intensional protocols – that one relies on an external clock whereas the other reacts to the current state – makes it difficult to compare the protocol frameworks directly. In particular, then the structure of an extensional protocol is not, in general, enough to determine whether the resulting sequence of models may be obtained from an intensional protocol. The second aspect of the motivation is that the methodology as a by-product relates DEL dynamical systems to ETL models, thus yielding results illuminating the former, on which there has been a recent interest, cf. Sec. 1.2 on related literature.

1.1 Structure of the Paper

**Section 2** defines core DEL components as well as intensional protocols ("DEL dynamical systems") and extensional protocols ("DEL protocols"). These are informally compared and contrasted by example. Finally, it is illustrated how DEL dynamics may be seen as producing ETL models.

**Section 3** presents ETL models and eight structural properties of key relevance to the paper.

**Section 4** formally defines how to generate ETL models from DEL dynamical systems and contains a first result: For an ETL model to be generatable by a DEL dynamical system, it must necessarily satisfy specific seven of the eight structural properties, but not necessarily the eighth.

**Section 5** concerns the other direction: Constructing DEL dynamical systems that will generate a given ETL model. It will be shown that if an ETL model possesses all eight structural properties, then this is sufficient for a suitable DEL dynamical system to exist.

Jointly, the results of Sections 4 and 5 *almost* yield a characterization of the ETL models generatable by DEL dynamical systems, but not quite.

**Section 6** restricts attention to a subclass of DEL dynamical systems and a subclass of ETL models: When a DEL dynamical system is *image-finite* and *concluding*, it generates an *image-finite* and *concluding* ETL model. In this case, a proper characterization is obtained: The eight properties are both necessary and sufficient.

**Section 7** moves the attention to *non-deterministic* intensional protocols, implemented by running several (deterministic) DEL dynamical systems in

---

[2] It is also a point of reference for other frameworks, like interpreted systems or extensive games with imperfect information, cf. the motivation in [13].

parallel. The motivation is a tighter correspondence with the methodology of extensional DEL protocols, of which only special cases are deterministic. The section presents weaker, necessary properties of ETL models generated by *families* of DEL dynamical systems.

**Section 8** contains the main comparison of intensional and extensional protocols for DEL, based on the differences in structural properties of generatable ETL models. The section thus compares and discusses the present results with those of van Benthem, Gerbrandy, Hoshi and Pacuit [13].

**Section 9** concludes with open questions.

1.2 Related Literature

The two main bodies of research to which this paper relates is that on protocols for DEL and that on DEL and dynamical systems.

The first body of research – on protocols for DEL – comprises [18,29,30, 43,41,11] and van Benthem, Gerbrandy, Hoshi and Pacuit's 2009 [13]. All papers in this collection uses extensional protocols in the style of Parikh and Ramanujam [36], to various ends. Of special interest to the present is [13]: In [13], the authors investigate which classes of ETL models one may generate using action models, product update and extensional protocols. Their results are illuminating in elucidating epistemic and logical properties inherent in the DEL methodology. The approach and results of [13] are presented and discussed throughout.

The second body of research – on DEL and dynamical systems – considers the iterated application of DEL model transformers on sets of pointed Kripke models *as* a dynamical system.[3] This idea was first explicitly put in play in *Exploring the Iterated Update Universe* by T. Sadzik in 2006.[4] The paper investigates frame conditions for action models that guarantee a stabilizing orbit modulo bisimulation, drawing on conceptual ideas from van Benthem, advanced in 2002 [8].

Since, various papers have honed in on long-run behavior of iterated announcements (see e.g. [1,7,17]) without explicit ties to dynamical systems. In

---

[3] For the reader interested so interested, there also exists a body of literature taking the converse perspective, using logics to describe qualitative aspects of long-run behavior. On this approach, logic meets dynamical systems by the latter playing the role of semantics to the former. Papers falling in this category, detailing logics of *dynamical topological systems*, include Kremer and Mints' 2007 handbook chapter [33] (on research from 1997 onwards by e.g. Artemov [2] and the authors of [32]) and several recent papers by Fernandéz-Duque [23,22,24]; Sarenac's paper from 2011 [42] exploring modal logical approaches to describing *iterated function systems*; and finally van Benthem's work in [8,10], outlining various possible logical approaches to fixed points and limit cycles of dynamical systems by applying fixed-point and oscillation operators galvanized by modal $\mu$-calculus. The latter two papers additionally provide an excellent bridge between the high abstraction level approach to logic and dynamical systems of this note and the micro-perspective literature in the main text.

[4] Available as Report PP-2006-26, Institute for Logic, Language, and Computation, University of Amsterdam.

[38, 39], intensional protocols are used to model information dynamics, using constructions similar to (one-step) *planning problems* of [16]. In these papers, no link is made to either protocol nor dynamical systems literature. In [40], the protocol format of [38, 39] is cast in dynamical systems terms and related to the multi-pointed action models [4, 3] and the *knowledge-based programs* of [21]. In [31], iterated dynamics are construed in a topological setting, investigating objects satisfying the common definition of a dynamical system: A compact, metric space under the action of a continuous function.[5]

In neither of the mentioned papers is DEL investigated as a dynamical system *qua* its role as protocol defining, nor has the resulting sequences been related to ETL models or extensional protocols.

## 2 Protocols for DEL

In this section, standard notions from Dynamic Epistemic Logic are introduced together with intensional and extensional DEL protocols. The reader is referred to the excellent literature on the topic of epistemic logic and DEL for more information and philosophical interpretation: See e.g. [28, 21, 4, 6, 19, 9, 5, 14, 12, 20].

### 2.1 Pointed Kripke Models and Language

Let be given a countable, non-empty set of propositional atoms $\Phi$ and a finite, non-empty set of agents, $I$. Throughout the paper, it will be assumed that these sets remain fixed.

A **Kripke model** is a tuple $M = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket)$ where

$\llbracket M \rrbracket$ is a countable, non-empty set of **states**;

$R : I \longrightarrow \mathcal{P}(\llbracket M \rrbracket \times \llbracket M \rrbracket)$ assigns to each agent $i$ an **accessibility relation** $R(i)$, also denoted $R_i$;

$\llbracket \cdot \rrbracket : \Phi \longrightarrow \mathcal{P}(\llbracket M \rrbracket)$ is a **valuation**, assigning to each atom an extension of states.

A pair $(M, s)$ with $s \in \llbracket M \rrbracket$ is called a **pointed Kripke model**. Throughout, the pair $(M, s)$ is written $Ms$.

Where $p \in \Phi$ and $i \in I$, define a language $\mathcal{L}_{(\Phi, I)}$ by

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i \varphi$$

with non-propositional formulas evaluated over pointed Kripke model $Ms$ by

$$Ms \models \Box_i \varphi \text{ iff for all } t \in \llbracket M \rrbracket, sR_i t \text{ implies } Mt \models \varphi,$$

and standard propositional semantics.

---

[5] In the present paper, this topological augmentation is not made. The reason is that the arguments used in comparing sequences of models obtained through some protocol to ETL models rely on structural features of concrete models: The required moves between concrete models and abstract, quotient models would thus add superfluous steps.

2.2 Action Models and Product Update

In Dynamic Epistemic Logic, dynamics are introduced by transitioning between pointed Kripke models from some set $X$ using a possibly partial map $f : X \longrightarrow X$. Such a map is often referred to as a *model transformer*. Many model transformers have been suggested in the literature, the most well-known being *truthful public announcement*, $!\varphi$ [37]. Truthful public announcements are a special case of a rich class of model transformers, here referred to as the class of *clean maps*.

In essence, a clean map $f$ is given by $f(x) = c(x \otimes a)$ with specific term $a$, product $\otimes$ and restricting operation $c$. The term $a$ is based on a *deterministic multi-pointed action model*, defined below.[6] Intuitively, one may think of a such as a set of program lines, each of the form "If $\varphi_i$, then do $a_i$", where the preconditions $\varphi_i$ are mutually exclusive. When "run" on a pointed Kripke model $x$, the program checks if $x$ satisfies any $\varphi_i$. If so, it executes action $a_i$ (a sub-action of $a$) on $x$, obtaining the result $x \otimes a$. If not, the the product of $x$ and $a$ is undefined. Finally, the operation $c$ removes redundant states. Their usage is exemplified in Section 2.6.

Define an **action model** as a tuple $\Sigma = (\llbracket \Sigma \rrbracket, \mathsf{R}, pre, post)$, sharing language $\mathcal{L}_{(\Phi, I)}$ with models in $X$, where

$\llbracket \Sigma \rrbracket$ is a countable, non-empty set of **actions** $\sigma$;

$\mathsf{R} : I \longrightarrow \mathcal{P}(\llbracket \Sigma \rrbracket \times \llbracket \Sigma \rrbracket)$ assigns an **accessibility relation** $\mathsf{R}(i)$ to each index $i \in I$, with $\mathsf{R}(i)$ denoted $\mathsf{R}_i$;

$pre : \llbracket \Sigma \rrbracket \longrightarrow \mathcal{L}_{(\Phi, I)}$ assigns to each action a **precondition**, specifying the conditions under which $\sigma$ is executable;

$post : \llbracket \Sigma \rrbracket \longrightarrow \mathcal{L}_{(\Phi, I)}$ assigns to each action a **postcondition** (a conjunctive clause[7] over $\Phi$, or $\top$). The postcondition specifies whether $\sigma$ changes the values of select atoms.

A pair $(\Sigma, \Gamma)$ with $\emptyset \neq \Gamma \subseteq \llbracket \Sigma \rrbracket$ is a **multi-pointed action model**; $(\Sigma, \Gamma)$ is also written $\Sigma\Gamma$. If $\Gamma$ is a singleton $\{\sigma\}$, then $\Sigma\Gamma$ is called **single-pointed** and is written $\Sigma\sigma$. If $X \models pre(\sigma) \wedge pre(\sigma') \to \bot$ for each $\sigma \neq \sigma' \in \Gamma$, then $\Sigma\Gamma$ is called **deterministic** over $X$, for $X$ a set of pointed Kripke models. The term deterministic is used as the requirement ensures that at most one designated action from $\Gamma$ "survives" when $\Sigma\Gamma$ is applied to a pointed Kripke model $Ms \in X$ using **product update** $\otimes$. The product $Ms \otimes \Sigma\Gamma$ is the

---

[6] Action models and product update was introduced in [4]. The extension to multi-pointed action models came with [3]. The present version of postconditions is inspired by [20] and the usage of deterministic models by [40].

[7] I.e. a conjuction of literals, where a literal is an atomic proposition or a negated atomic proposition.

pointed Kripke model $(\llbracket M\Sigma \rrbracket, R', \llbracket \cdot \rrbracket', s')$ with

$$\llbracket M\Sigma \rrbracket = \{(s,\sigma) \in \llbracket M \rrbracket \times \llbracket \Sigma \rrbracket : Ms \models pre(\sigma)\}$$
$$R' = \{((s,\sigma),(t,\tau)) : (s,t) \in R_i \text{ and } (\sigma,\tau) \in \mathsf{R}_i\}, \text{ for all } i \in I$$
$$\llbracket p \rrbracket' = \{(s,\sigma) : s \in \llbracket p \rrbracket, post(\sigma) \not\models \neg p\} \cup \{(s,\sigma) : post(\sigma) \models p\}, \text{ for all } p \in \Phi$$
$$s' = (s,\sigma) : \sigma \in \Gamma \text{ and } Ms \models pre(\sigma)$$

If $Ms$ does not satisfy the precondition of any action $\sigma$ in $\Gamma$ or if $\Sigma\Gamma$ is not deterministic over $\{Ms\}$, then product is undefined.

In the product $Ms \otimes \Sigma\Gamma$, there may be states that are not reachable from the point $(s,\sigma)$ via any collection of relations. Such states are, for present purposes, superfluous: They neither affect the formulas satisfied at $(s,\sigma)$ nor the set of models with which $(Ms \otimes \Sigma\Gamma, (s,\sigma))$ is bisimilar. As it is later convenient to work with correspondence between structures up to isomorphism, in the current paper such superfluous states are always deleted.

Superfluous states are deleted by regarding only the substructure of any pointed Kripke model $Ms$ that is connected to the actual state $s$. This substructure is denoted $C(Ms)$ and is defined as follows:

Let $R^*$ be the reflexive, transitive and symmetric closure of $\{R_i\}_{i \in I}$. Let $R^*(s)$ be the set of states reachable from $s$ via $R^*$, i.e., $R^*(s) := \{s' \in \llbracket M \rrbracket : (s,s') \in R^*\}$. Then the **connected component** of $Ms$ is the *unpointed* substructure $C(Ms) := (\llbracket M \rrbracket_{|R^*(s)}, R_{|R^*(s)}, V_{|R^*(s)})$. With $s' \in \llbracket C(Ms) \rrbracket$, $C(Ms)s'$ is thus again a pointed Kripke model. In particular, $C(Ms)s$ is bisimilar to $Ms$.[8],[9]

2.3 Intensional Protocols: DEL Dynamical Systems

The most general class of maps – the intensional protocols – of interest in the following may now be defined as follows:

**Definition 1 (Clean Map)** Let $X$ be a set of pointed Kripke models. A **clean map on** $X$ is any possibly partial model transformer $f : X \longrightarrow X$ given by $f(x) = C(x \otimes \Sigma\Gamma)s'$, for all $x \in X$, with $\Sigma\Gamma$ a multi-pointed action model deterministic over $X$.

Defining intensional protocols using mappings, it is required that also their domain and range be specified:

**Definition 2 (DEL dynamical system)** A **DEL dynamical system** is a pair $(X, f)$ where $X$ is a set of pointed Kripke models and $f$ is a clean map on $X$. A **pointed DEL dynamical system** $(X, f, x)$ is augmented with an **initial model** $x \in X$, assumed to be connected.

---

[8] This is not a *bisimulation contraction* (cf. [27]): $C(Ms)s$ need not be bisimulation minimal.

[9] The authors apologize for the cumbersome notation: It is useful when later working with connected components in unpointed epistemic, temporal structures.

The **orbit** of $(X, f, x)$ is the (possibly finite) sequence $x, f(x), f(f(x)), ...$. Misusing notation if $f^k(x)$ is undefined for some $k \in \mathbb{N}$, the orbit is denoted $\langle f^k(x) \rangle_{k \in \mathbb{N}}$.

*Remark 1* This definition of a DEL dynamical system is restrictive. A broader definition would allow $f$ to be any bisimulation-preserving map.

2.4 Extensional Protocols: DEL Protocols

In [13], two types of DEL protocols are defined, one allowing the protocol to vary from state to state of the initial model and one where the protocol is "common knowledge":

**Definition 3 (DEL Protocol)** Let $\mathbb{E}$ be the class of all $\mathcal{L}_{(\Phi, I)}$ single-pointed action models. Let $\mathbb{E}^*$ be the class of all finite sequences of elements from $\mathbb{E}$. A set $\mathsf{P} \subseteq \mathbb{E}^*$ is a **DEL protocol** iff $\mathsf{P}$ is closed under non-empty prefixes. Let $Ptcl(\mathbb{E})$ denote the class of all DEL protocols.

Let $Ms$ be a pointed Kripke model. A **state-dependent DEL protocol** on $Ms$ is a map

$$\mathsf{p} : [\![M]\!] \longrightarrow Ptcl(\mathbb{E}).$$

If $\mathsf{p}$ is constant over $[\![M]\!]$, i.e., if for all $s, t \in [\![M]\!], \mathsf{p}(s) = \mathsf{p}(t)$, then $\mathsf{p}$ is a **uniform DEL protocol**.

A DEL protocol specifies which pointed action models *may* be executed at a given time – whether they *can* be executed then again depends on the preconditions of the designated action. Their usage is exemplified in Section 2.6.

2.5 An Initial Comparison

Although DEL protocols and DEL dynamical systems invoke the same rudimentary changes by using action models, they differ vastly in structure. In particular, where every DEL dynamical systems encodes a deterministic[10] protocol – by virtue of being defined as a mapping – DEL protocols may be non-deterministic. Roughly, DEL dynamical systems may be correlated with state-dependent and uniform DEL protocols in the following manner:

▷ A DEL dynamical system is analogous to a deterministic, uniform DEL protocol: A DEL protocol $\mathsf{P} \subseteq \mathbb{E}^*$ for which all sequence $\varsigma, \varsigma' \in \mathsf{P}$, either $\varsigma$ is a prefix of $\varsigma'$ or *vice versa*.
▷ A non-deterministic, uniform DEL protocol is analogous to a family of DEL dynamical system, executed in parallel on the *same* pointed Kripke model.

---

[10]  In the sense that given any input state (pointed Kripke model), the protocol outputs at most a single resulting state.

▷ A non-deterministic, non-uniform DEL protocol is analogous to a family of DEL dynamical systems, executed in parallel on *different* pointed Kripke models, all of which are identical up to the choice of designated point.

In the present, dealing with non-uniform DEL protocols or their DEL dynamical systems counterparts will be omitted.

Without going through ETL models, DEL dynamical systems and uniform DEL protocols may be related, showing that the orbits obtainable from DEL dynamical systems is a sub-class of those obtainable using DEL protocols:

**Proposition 1** *Let $(X, f, x)$ be a pointed DEL dynamical system. Then there exists a singleton uniform DEL protocol that produces the orbit of $f$ from $x$.*

*Proof* At each iteration, the clean map $f$ is − in effect − going to execute a single-pointed action model. Copying the sequence of thusly executed action models provides a uniform DEL protocol. For details, see Appendix.

The converse of Proposition 1 does not hold: There exists pointed Kripke models with associated singleton uniform DEL protocols that produce sequences of pointed Kripke models not duplicatable by any DEL dynamical system.[11] This is is a consequence of DEL protocols being *extensional*: Not only do they consult the information inherent in the present model to determine ensuing actions, but also the current *time*, exogenously provided by the sequential nature of the protocol. This information is not available to DEL dynamical systems and can therefore not be used in guiding dynamics.

As remarked in the introduction, this feature makes it difficult to compare DEL protocols and DEL dynamical systems directly: The structure of the DEL protocol may not be enough to determine whether the resulting sequence of pointed Kripke models may be obtained as the orbit of a DEL dynamical system. Hence the current approach, a comparison using ETL models.

2.6 Example: The Muddy Children Puzzle

To illustrate the differences in use of DEL dynamical systems *qua* intensional protocols and uniform DEL protocols *qua* extensional protocols, two such formal protocols of the classic Muddy Children Puzzle, well-known in the DEL literature (see e.g. [26,19]), are presented. As a simplified version of the puzzle is sufficient for present purposes, attention is restricted to the case with three children.

The puzzle initiates with a partial description of an epistemic state:

---

[11] An example is the following: Let a two-state pointed Kripke model $Ms$ with $Ms \models p \wedge q$ and $Mt \models p \wedge \neg q$ be given. Let $\mathsf{p}(s) = \{(!p), (!p, !p \wedge q)\}$ with $!\varphi$ the truthful public announcement of $\varphi$. Then $\mathsf{p}$ on $Ms$ produces the sequence $(Ms, Ms, M's)$ with $\llbracket M' \rrbracket = \llbracket M \rrbracket \backslash \{t\}$. No clean map can duplicate this sequence: As $f(Ms) = Ms$, the system has reached a fixed point from which it will never deviate to produce $M's$.

$M$ :
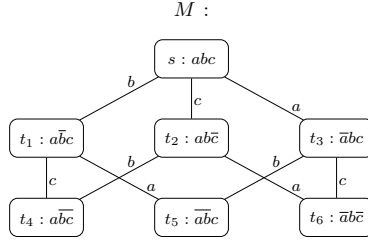


**Fig. 1** Unpointed Kripke model $M$ representing the initial situation of the Muddy Children Puzzle. Each state specifies which children are muddy, where for all $i \in \Phi$, $\bar{i} := \neg i$, that is, "child $i$ is not muddy". Labelled relations between states represent ignorance of the children. Reflexive relations are omitted.

Three brilliant children have been playing outside. During play, each may have obtained a muddy forehead. Each can tell whether or not others have muddy foreheads, but cannot tell this of themselves. Upon returning home from play, an adult of unspecified gender informs the children that at least one of them is muddy.

Following standard practice in DEL, this partial description is modeled as an unpointed Kripke model for a language $\mathcal{L}_{(\Phi,I)}$ with the set of agents $I = \{a, b, c\}$ and the set of atoms $\Phi = \{a, b, c\}$ with $i \in \Phi$ read "child $i$ is muddy". The unpointed model $M$ is illustrated in Figure 1. A pointed Kripke model results when a designated state is determined: This corresponds to fixing which children became muddy during play. Denote the set of resulting pointed Kripke models $X_M$.

The puzzle specification continues by the adult detailing a protocol by which the children should update the initial epistemic state:

"Concurrently with this metronome," the adult instructs, "repeatedly and simultaneously announce aloud whether or not you know whether or not you are muddy."

By means of a suitable model of this protocol, it is desirable to be able to answer the main question of the puzzle, namely:

If there are $n$ muddy children, how many times does the metronome have to tick before all three children know whether or not they are muddy?

As the uniform DEL protocol and the DEL dynamical systems protocol will share the same informational actions, these will be introduced first.

*2.6.1 Muddy Children: Announcements*

As standard, each of the announcements made is treated as a truthful public announcement, cf. [37]. A truthful public announcement of the formula $\varphi$ may be modeled using a single-pointed action model with a single action with $\varphi$ as precondition.

Each epistemic announcement is modeled using the same singleton single-pointed action model, changing only the precondition. Build the formulas for the group announcements as follows:

1. Interpret the $\Box_i$ modality as reading "child $i$ knows that...", and denote the operator by $K_i$.
2. Let $know_i$ be short for $K_i i \vee K_i \neg i$. If $know_i$ is true, then child $i$ knows whether he or she is muddy or not.
3. Let $know_S$ for $S \subseteq I$ be the formula $\bigwedge_{i \in S} know_i \wedge \bigwedge_{i \notin S} \neg know_i$. Then $know_S$ states that exactly the children in $S$ know their status.
4. For each $S \subseteq I$, let $\Sigma_S \sigma_S = (\{\sigma_S\}, \mathsf{R}_S, pre_S, post_S, \sigma_S)$ be the singleton single-pointed action model with $pre(\sigma_S) = know_S$, $post(\sigma_S) = \top$ (as the announcement makes no changes to atomic valuations), and $\mathsf{R}_S(i) = \{(\sigma_S, \sigma_S)\}$ each $i \in I$. As in the initial Kripke model, the epistemic relations are thus equivalence relations.

*2.6.2 Muddy Children: Intensional Protocol*

Notice that the instructions of the parent in the natural language protocol are already provided in an intensional (conditional) form. Essentially, the parent instructs the children to follow the rules

"If you know whether you are muddy, then announce so.", and

"If you don't know, then announce so."

Aggregated to rules for the group, the antecedents in these conditional rules are exactly the preconditions of the actions in the $\Sigma_S \sigma_S$ models. As these preconditions are pairwise jointly unsatisfiable over any set of pointed Kripke models and the models are disjoint, their union is a deterministic multi-pointed action model: Let $\Sigma \Gamma = ([\![\Sigma]\!], \mathsf{R}, pre, post, \Gamma)$ with, for $* \in \{[\![\Sigma]\!], \mathsf{R}, pre, post, \}$, $* = \cup_{S \subseteq I} *_S$ and $\Gamma = \cup_{S \subseteq I} \{\sigma_S\}$.

Let $X$ be a superset of the muddy children models $X_M$ of Fig. 1, closed under the operation $\otimes \Sigma \Gamma$. With $f$ the clean map on $X$ based on $\Sigma \Gamma$, $(X, f)$ is a DEL dynamical system. Moreover, applied to any $x \in X_M \subseteq X$, $f$ implements the desired protocol and produces, tractably and in finite time, an answer to the puzzle. Figure 2 illustrates this for the case of three muddy children.

With this implementation, the intensional protocol may straightforwardly be applied to pointed models differing in other respects then the number of muddy children, e.g., with different initial announcements of the parent.

*2.6.3 Muddy Children: Extensional Protocol*

Constructing an extensional protocol for the Muddy Children given some initial Kripke model is straightforward: Simply run the intensional protocol above on the initial model, taking note which designated actions' preconditions were satisfied when and encode this sequence as a extensional protocol. The resulting extensional protocol will induce the motions appropriate for the given

$Ms:$          $f^1(Ms):$          $f^2(Ms):$
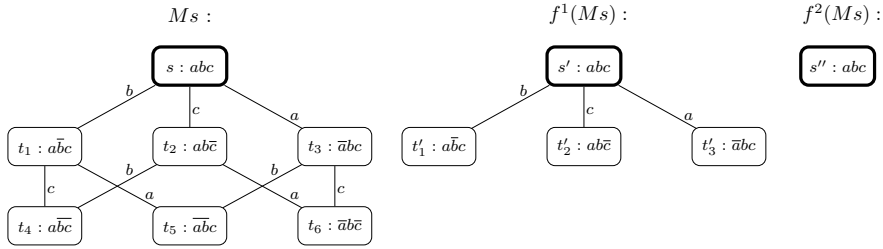


**Fig. 2** The three models of the Muddy Children Puzzle in the case of three muddy children. For any $r' \in \llbracket f^1(Ms) \rrbracket$, let $r' := (r, \sigma_\emptyset)$ and let $s'' := ((s, \sigma_\emptyset), \sigma_\emptyset)$. Points are distinguished by a thick contour. It can be seen that after two rounds of question and answers all children know whether or not they are muddy. Furthermore, $f^k(Ms)$ is isomorphic to $f^2(Ms)$ for all $k \geq 2$.

initial model. However, the protocol will not be useful in answering the question of the puzzle: It is a once-off solution for the given Kripke model only, constructed with knowledge of the answer sought.

A more informative extensional DEL protocol may be constructed, but it requires a countably infinite representation: Assume to construct an extensional DEL protocol that will adequately encode the natural language instructions, is applicable to any model in $X_M$ and presumes no prior knowledge of the developing information dynamics. The set of relevant announcements is, as above, $\{\Sigma_S \sigma_S \colon S \subseteq I\}$. For the announcement made at the first time step, the protocol must allow $\Sigma_S \sigma_S$ for each $S \subseteq I$, seeing that no information about the development of the dynamics may be assumed. Similarly, each possible announcement must be allowed to follow the first, etc. Hence, only satisfactory extensional protocol is $\mathsf{P} = \{\Sigma_S \sigma_S \colon S \subseteq I\}^*$. This set is countably infinite.

This protocol facilitates finding an answer to the Muddy Children Puzzle: For a given initial model $Ms \in X_M$, find the actions that the protocol allows to be executed at time 1. These are all the actions models $\Sigma_S \sigma_S$ for which the length 1 sequence $\langle \Sigma_S \sigma_S \rangle$ is in $\mathsf{P}$ (i.e., all the actions the protocol allows at time 1). For each of these, calculate the product $Ms \otimes \Sigma_S \sigma_S$. As the preconditions are, in the current example, mutually inconsistent, only one such model will be well-defined. The result is exactly $f(Ms)$, for $f$ the intensional protocol given above. For time 2, take all the models produced at time 1 – in this case $\{f(Ms)\}$ – and execute on each of them all the actions in the continuations of the sequence from which that model stems. This produces a second set of pointed Kripke models – in this case $\{f(f(Ms)\}$. This process thus leads to a model in which all children will announce that they know whether they are muddy.

*Remark 2* If one is interested in implementing a DEL protocol to seek computational assistance in puzzle solving, the countably infinite representations required for the extensional protocol may prove cumbersome.[12]

---

[12] It was not suggested in [13] that DEL protocols be implementable nor that they are suited for modeling purposes.

### 3 Epistemic Temporal Logic

The run of a pointed DEL dynamical system may be recorded as a sequence of pointed Kripke models. Using information from the action models, an insight of [13] was that this may naturally be regarded as a temporal, modal structure, a so called *Epistemic Temporal Logic model.*

ETL models, introduced in [36], form a both simple and general framework. Such models allow the representation of epistemic and temporal interplay, and allow doing so without in an assumption-free manner. Hence, in generating ETL models from DEL dynamics, any structural properties (e.g., Synchronicity, Perfect Recall) shared by the generated ETL models are features induced by the DEL operations. Thus, characterizing the classes of generatable ETL models elucidates assumptions implicit in DEL dynamics about epistemic and temporal interplay. This is a main conceptual insight of [13].

An ETL model is a temporal forest with additional modal (epistemic) relations between notes. With $E^*$ the set of all finite sequences of elements from the set $E$, an **ETL model** for the language $\mathcal{L}_{(\Phi,I)}$ is a tuple $\mathcal{H} = (E, H, \mathtt{R}, V)$ where

$E$ is a set of **events** $e$;

$H \subseteq E^*$ is a set of **histories**, closed under non-empty prefixes;[13]

$\mathtt{R} : I \longrightarrow \mathcal{P}(H \times H)$ is a map assigning to each agent an **accessibility relation** $\mathtt{R}(i)$, written $\mathtt{R}_i$;

$V : \Phi \longrightarrow \mathcal{P}(H)$ is a **valuation**.

In contrast with pointed Kripke models, ETL models are not equipped with actual states. To obtain a tighter connection between DEL dynamical system orbits and ETL models, the latter is augmented to include *multiple points.* Figure 3 on the following page illustrates such an augmented ("saturated") ETL model and its relation to the orbit of a DEL dynamical system.

As in Sec. 2.2, let $\mathtt{R}^*$ be the reflexive, symmetric and transitive closure of $\mathtt{R}$, and let $\mathtt{R}^*(h) := \{h' \in H \colon (h, h') \in \mathtt{R}^*\}$. Then define the **connected component** of $h \in H$ in $\mathcal{H}$ – denoted $C(\mathcal{H}h)$ – as the restriction of $\mathcal{H}$ to $\mathtt{R}^*(h)$, i.e., let $C(\mathcal{H}h) := (H_{|\mathtt{R}^*(h)}, \mathtt{R}_{|\mathtt{R}^*(h)}, V_{|\mathtt{R}^*(h)})$. If $\underline{h} \in H_{|\mathtt{R}^*(h)}$, then $C(\mathcal{H}h)\underline{h}$ is a pointed Kripke model.

Finally, define the ETL structures of interest as follows:

**Definition 4 (Saturated ETL Model)** Let $\mathcal{H} = (E, H, \mathtt{R}, V)$ be an ETL model. Let $\underline{H} \subseteq H$ be a set of histories closed under prefixes, called **points**. The pair $(\mathcal{H}, \underline{H})$ is **saturated** iff for all $h \in H$, the connected component $C(\mathcal{H}h)$ contains a unique point $\underline{h}$ from $\underline{H}$.

---

[13] It is overall assumed that any ETL model contains no redundant events relative to the model's set of histories. That is, for any ETL model $\mathcal{H}$, it holds that any event $e \in E$ is either a history ($e \in H$) or part of a history ($\exists h \in H$ such that $he \in H$).
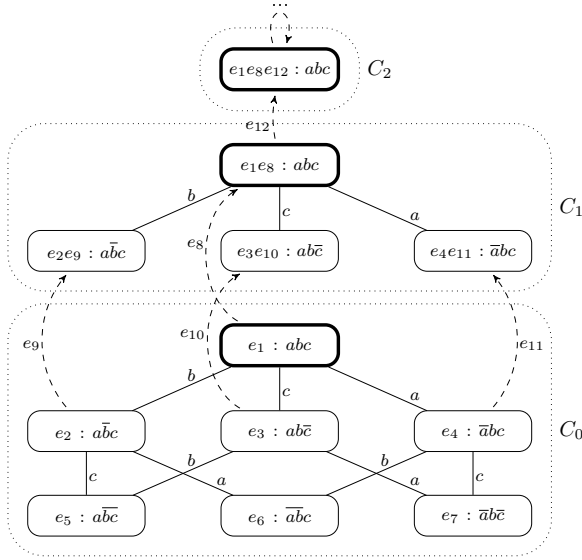
**Fig. 3**   A saturated ETL model $(\mathcal{H}, \underline{H})$. Time flows upwards where labelled dashed lines represent events. Connected components are marked by dotted circles, points by thick contours. Each pointed connected component is isomorphic to a pointed Kripke model from Figure 2 on page 12. Moreover, a history $h'$ is the successor of $h$ in $\mathcal{H}$ iff the Fig. 2 counterpart of $h'$ is a state $(s, \sigma)$ for $s$ the counterpart of $h$.

*Remark 3* The addition of points to ETL models is vital for the ensuing constructions: When computing the $k+1$th element of an orbit of a DEL dynamical system, any clean map uses information from the designated point of the $k$th element. Thus, it is essential information for the evolution of a DEL dynamical system which point of a Kripke model is designated. Hence, to structurally relate ETL models to orbits of DEL dynamical systems (each element of which is pointed), it is necessary to add a notion of points to the former.

## 3.1 ETL Isomorphism

For simplicity of arguments, saturated ETL models are identified up to isomorphism. This allows arguments without repeated references to bisimulation contractions or other specific representatives. In the definition of isomorphism between ETL models, note that the temporal structure of the models is also preserved:

**Definition 5 (ETL Isomorphism)** Let saturated ETL models $(\mathcal{H}, \underline{H}) = (E, H, \mathtt{R}, V, \underline{H})$ and $(\mathcal{H}', \underline{H}') = (E', H', \mathtt{R}', V', \underline{H}')$ be given. Let $f : E \longrightarrow E'$. For $h = e_0...e_n \in E^*$, let $f(h) := f(e_0)...f(e_n)$. The map $f$ is an **ETL isomorphism** iff $f$ is a bijection and for all $h \in H$, $h' \in H'$

1. $h \in H$ iff $f(h) \in H'$, and $h \in \underline{H}$ iff $f(h) \in \underline{H}'$,
2. $h\mathtt{R}_i h'$ iff $f(h)\mathtt{R}'_i f(h')$, for all $i \in I$,
3. $h \in V(p)$ iff $f(h) \in V'(p)$, for all $p \in \Phi$.

$(\mathcal{H}, \underline{H})$ and $(\mathcal{H}', \underline{H}')$ are **ETL isomorphic** iff there exists an ETL isomorphism between their domains.

In the remainder, "ETL isomorphism" and "isomorphism" are used interchangeably.

3.2 ETL Model Properties

When generating an ETL model from a DEL dynamical system, the resulting forest will inherit a set of properties. Some stem from the graph theoretic nature of action models, product update and the associated pruning to connected components of clean maps, some from the workings of pre- and postconditions, and yet some stock from the functional *modus operandi* of dynamical systems. Def. 6 lists the eight properties of main relevance to this paper.

Of these, Synchronicity, Perfect Recall and Local No Miracles are well-known . In models interpreted epistemically, they roughly require, respectively, that agents know the current time, never forget what they have learned (though new uncertainty may be introduced), and that events carry the same information in all states in the same "context" – in each connected component. These were identified by van Benthem et al. [13] to be inherited in any ETL model generated using sequences of actions models (see Sec. 8 for discussion and comparison).[14]

Connected Time-Steps is almost the converse of Synchronicity: It requires that a time-step contains at most one connected component. It relates to the deletion of superfluous states by clean maps.

From the use of product update *simpliciter* comes Precondition Describable while Postcondition Describable results from using action models with postconditions. The formula $\delta_e$ required to exist by Precondition Describable describes exactly those histories $h$ in a connected component in the ETL model on which $e$ is executed. This is required in order to formulate the precondition of the action $\sigma_e$ that will 'simulate' event $e$ in the DEL dynamical system. The formula in Postcondition Describable describes the propositional change due to $e$. This is necessary for the postcondition of $\sigma_e$ based on $e$ to be well-formed.

In property Component Collection Describable, it is demanded that any (bisimulation-closed) collection of pointed connected components is definable by a formula. The existence of these formulas is required to ensure that the preconditions for each $\sigma \in \varGamma$ are well-defined such that the right action points may be executed at the right times.

The three formula properties Precondition Describable, Postcondition Describable and Component Collection Describable are given as existence requirements without listing criteria that ensure their satisfaction.

Related to the functional nature of DEL dynamical systems is Point Bisimulation Invariance. As this property is somewhat tricky, it is here explained:

---

[14] Among the four core properties of [13], the authors also included Atomic Permanence, stating that atom values do not change. Working with action models with postconditions, this no longer applies.

Take two connected components, $C(\mathcal{H}h_1)$ and $C(\mathcal{H}h_2)$, from the saturated ETL model $(\mathcal{H}, \underline{H})$. Each component will then have a single designated point. Let $\underline{h}$ be the designated point of $C(\mathcal{H}h_1)$ and $\underline{h}'$ that of $C(\mathcal{H}h_2)$. The property then states the following: If the points $\underline{h}$ and $\underline{h}'$ are bisimilar, then if two other histories, say $h_3$ and $h_4$, from respectively $C(\mathcal{H}h_1)$ and $C(\mathcal{H}h_2)$ are also bisimilar, then $h_3$ and $h_4$ will be extended by exactly the same events. The property thus reflects two aspects of clean maps. First, that they are mappings: When applied to identical elements (pointed Kripke models that have bisimilar points), then the same action is executed on these elements. Second, it reflects the workings of the preconditions in action models: If the same action model is executed on any two pointed Kripke models, then if *any* points in those two models are bisimilar, then they will be treated equally under the product with the action model. The second aspect is the content of the weaker property Local Bisimulation Invariance of [13], to which Point Bisimulation Invariance is related in Sec. 8.

Write $Ms \leftrightarrow Nt$ iff the pointed Kripke models $Ms, Nt$ are *bisimilar*, cf. [15]. With $len(h)$ denoting the *length* of history $h$, $he$ denoting the sequence extending history $h$ with event $e$ and $h \sqsubseteq h'$ denoting that $h$ is a prefix of $h'$, consider the properties formally:

**Definition 6 (Saturated ETL Model Properties)** Let $\mathcal{H} = (E, H, \mathtt{R}, V, \underline{H})$ be a saturated ETL model. $\mathcal{H}$ satisfies

Synchronicity *iff* $\forall h, h' \in H$, if $h\mathtt{R}_i h'$, then $len(h) = len(h')$;

Perfect Recall *iff* $\forall h, h' \in H, \forall e, e' \in E : he, h'e' \in H$, if $he\mathtt{R}_i h'e'$, then $h\mathtt{R}_i h'$;

Local No Miracles *iff* $\forall h, h', h_1, h_2 \in H, \forall e, e' \in E : he, h'e' \in H$, if $h\mathtt{R}_i h'$, $h_1 e\mathtt{R}_i h_2 e'$ and $h\mathtt{R}^* h_1$, then $he\mathtt{R}_i h'e'$;

Connected Time-Steps *iff* $\forall h, h' \in H$, if $len(h) = len(h')$, then $h\mathtt{R}^* h'$;

Precondition Describable *iff* $\forall e \in E$, the exists a $\delta_e \in \mathcal{L}_{(\Phi, I)}$ such that if there is a $h'e \in H$, then for all $h \in H$, if $h'\mathtt{R}^* h$, then $C(\mathcal{H}h)h \models \delta_e$ iff $he \in H$;

Postcondition Describable *iff* $\forall he \in H$, there exists a $\delta_{D_e} \in \mathcal{L}_{(\Phi, I)}$ such that $\delta_{D_e} \models D_e$ for $D_e$ the union
$\{p \in \Phi : h \notin V(p), he \in V(p)\} \cup \{\neg q : q \in \Phi, h \in V(q), he \notin V(q)\}$;

Component Collection Describable *iff* $\forall A \subseteq \underline{H}$ such that $\underline{h} \in A$ and $\underline{h}' \leftrightarrow \underline{h}$ implies $\underline{h}' \in A$ and such that $\forall \underline{h}, \underline{h}' \in A$, $\underline{h} \sqsubseteq \underline{h}'$ or $\underline{h}' \sqsubseteq \underline{h}$, there exists a $\varphi_A \in \mathcal{L}_{(\Phi, I)}$ such that $C(\mathcal{H}\underline{h})\underline{h} \models \varphi_A$ iff $\underline{h} \in A$;

Point Bisimulation Invariance *iff* $\forall h_1, h_2, h_3, h_4 \in H$, if $C(\mathcal{H}h_1)\underline{h} \leftrightarrow C(\mathcal{H}h_2)\underline{h}'$ and $C(\mathcal{H}h_1)h_3 \leftrightarrow C(\mathcal{H}h_2)h_4$, then $h_3 e \in H$ iff $h_4 e \in H$.

## 4 Generated ETL Models and their Properties

A saturated ETL model is generated from an initial pointed Kripke models $x$ and a clean map $f$ by, essentially, recording the orbit of $f$ from $x$ as a temporal

structure: The states of $x$ become histories of length 1 and states of $f^k(x)$ become histories of length $k + 1$; the actual state in each $f^k(x)$ becomes an ETL model point; epistemic relations and valuations are directly transferred. Formally:

**Definition 7 (Generated Structure)** For any pointed DEL dynamical system $(X, f, x)$, its **generated structure** is the tuple $(E, H, \mathtt{R}, V, \underline{H})$ given by

$$
\begin{aligned}
E &:= \{e_\sigma : \sigma \in [\![\Sigma]\!]_k \text{ for some } k \in \mathbb{N}\} \\
&\quad \text{for } [\![\Sigma]\!]_0 := [\![x]\!] \text{ and } [\![\Sigma]\!]_{k+1} := \{\sigma : (s, \sigma) \in [\![f^{k+1}(x)]\!]\} \\[4pt]
H &:= \{\gamma(s) : s \in [\![f^k(x)]\!] \text{ for some } k \in \mathbb{N}\} \\
&\quad \text{with } \gamma : [\![\Sigma]\!] \longrightarrow E \text{ given by } \gamma(\sigma) = e_\sigma \\
&\quad \text{and for } s = ((\sigma_1, \sigma_2), ..., \sigma_n) \text{ use } \gamma(s) := \gamma(\sigma_1)\gamma(\sigma_2)...\gamma(\sigma_n) \\[4pt]
\mathtt{R}_i &:= \{(h, h') \in H \times H : \gamma^{-1}(h) R_i \gamma^{-1}(h')\} \text{ for all } i \in I \\[4pt]
V(p) &:= \{h \in H : \exists k \in \mathbb{N}, \gamma^{-1}(h) \in [\![p]\!]_k\} \\[4pt]
\underline{H} &:= \{\underline{h} : \exists k \in \mathbb{N}, \ f^k(x) = Ms \text{ and } \underline{h} = \gamma(s)\}
\end{aligned}
$$

If $(E, H, \mathtt{R}, V, \underline{H})$ is isomorphic to a saturated ETL model $(\mathcal{H}', \underline{H}')$, then $(X, f, x)$ **generates** $(\mathcal{H}', \underline{H}')$.

*Property 1* For any DEL dynamical system, the structure generated is a saturated ETL model: $H$ is indeed closed under prefixes and it is saturated as for all $h \in H$, $C(\mathcal{H}h)$ shares a unique $\underline{h}$ with $\underline{H}$.

The first main result furnishes a set of properties that any DEL dynamical system generated ETL model will necessarily satisfy:

**Proposition 2** *If saturated ETL model $(\mathcal{H}, \underline{H})$ is generated by a pointed DEL dynamical system, then $(\mathcal{H}, \underline{H})$ satisfies seven of the eight properties of Def. 6, namely Synchronicity, Perfect Recall, Local No Miracles, Connected Time-Steps, Precondition Describable, Postcondition Describable, and Point Bisimulation Invariance.*

*Proof* All proofs may be found in Appendix starting on page 29.

The second result shows that the last property of Def. 6 is indeed only a contingent feature of some generated ETL models:

**Proposition 3** *Not all saturated ETL models generated by pointed DEL dynamical systems are Component Collection Describable.*

*Proof* See Appendix.

## 5 From ETL Model to Dynamical System

For certain ETL models, there exists DEL dynamical systems that will generate them. The following result lists sufficient conditions of an ETL model to be generatable:
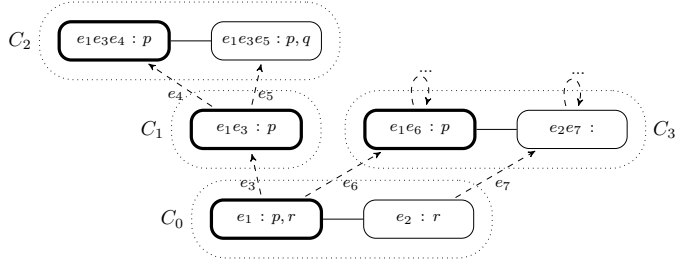
**Fig. 4** An ETL model $(\mathcal{H}, \underline{H})$ with two saturated component branches. Connected components $C_0$, $C_1$ and $C_2$ form one saturated component branch **b**. The infinite set consisting of components $C_0$, $C_3$, etc. form another saturated component branch **b'** .

**Proposition 4** *If* $(\mathcal{H}, \underline{H})$ *is a saturated ETL model that satisfies all eight properties of Def. 6, then there exists a pointed DEL dynamical system that generates* $(\mathcal{H}, \underline{H})$.

The proof, which may be found in the Appendix, rests on the idea of regarding an ETL model as a collection of *saturated component branches*, illustrated in Figure 4. Each such branch is a sequence of pointed Kripke models and hence potentially the orbit of a DEL dynamical system.

To obtain the notion of a saturated component branch, decompose ETL model into *branches*, lump these together in connected components and saturate:

**Definition 8 (Branches)** A **branch** of an ETL model $\mathcal{H} = (E, H, \mathtt{R}, V)$ is a set $b \subseteq H$ that

1. has a unique **root**, i.e., contains a unique history that has length 1;
2. is maximal with unique extension: If $h \in b$ and $he \in H$ for some $e \in E$, then $|\{he' : he' \in b\}| = 1$;
3. is closed under finite prefixes.

The **component branch** of $b$ is the sequence $\mathbf{b} = \mathbf{b}_1, \mathbf{b}_2, ...$ of connected components, ordered according to history length with prefix $(C(\mathcal{H}h))_{h \in b}$, extended to be either either **maximal** in $H$ $(\exists k \in \mathbb{N} \forall h' \in \mathbf{b}_k \neg \exists e \in E : h'e \in H)$ or **infinite** $(\forall k \in \mathbb{N} \exists h' \in \mathbf{b}_k \exists e \in E : h'e \in \mathbf{b}_{k+1})$. A **saturated component branch** is a pair $(\mathbf{b}, \underline{H})$ with $\underline{H} \subseteq H$ a set of **points** closed under finite prefixes such that every component in **b** has exactly one point.

Enumerating $\underline{H}$ by history length, the following link to pointed Kripke models is obtained:

*Property 2* For saturated component branch $(\mathbf{b}, \underline{H})$, the pair $(\mathbf{b}_k, \underline{h}_k)$ is a pointed Kripke model.

Moreover, the construction emphasizes how select ETL models have a strong resemblance to DEL dynamical system orbits:

*Property 3* If saturated ETL model $(\mathcal{H}, \underline{H})$ has property Connected Time-Steps, then $\mathcal{H}$ has a unique component branch **b**.

Jointly, these two properties allow us to illustrate the proof methodology of Proposition 4: Take an ETL model $\mathcal{H}$ that has Connected Time-Steps and is saturated by points $\underline{H}$. Envision the model as a component branch $\mathbf{b}$ saturated by $\underline{H}$. Extract from this the sequence of pointed Kripke models $(\mathbf{b}_k, \underline{h}_k)_{k \in \mathbb{N}}$. For each $k$, find an action model that transforms $(\mathbf{b}_k, \underline{h}_k)$ into $(\mathbf{b}_{k+1}, \underline{h}_{k+1})$. Join all these action models into a deterministic multi-pointed action model and construct its clean map $f$. Then $(\{(\mathbf{b}_k, \underline{h}_k) : k \in \mathbb{N}\}, f, (\mathbf{b}_1, \underline{h}_1))$ is a pointed DEL dynamical system that generates $(\mathcal{H}, \underline{H})$. Full details may be found in the Appendix.

The relation between DEL dynamical systems and ETL models taht do not have Connected Time-Steps is discussed in Sec. 7.

## 6 Characterization: Image-finite and Concluding

Propositions 2 and 4 do not quite yield a characterization result pertaining to the ETL model generatable by DEL dynamical systems. This is due to the fact that Component Collection Describable is not implied for ETL models generated by an DEL dynamical system when working with a normal, finitary modal logical language, as shown by Proposition 3.

Imposing two restrictions on ETL models and DEL dynamical systems yields a characterization result. Both are finiteness assumptions. The first an assumption of *image-finiteness* for the modal relations:

A binary relation $B \subseteq A \times A$ is **image-finite** iff the set $\{y : (x, y) \in B\}$ is finite for all $x \in A$. On sets of image-finite structures, the **Hennessy-Milner Theorem** ensures that bisimilarity and modal equivalence relate exactly the same models, cf. e.g. [15, 27]. The assumption is therefore natural from a modal logical point of view. The notion may be applied to DEL dynamical systems and ETL models: Call a pointed DEL dynamical system $(X, f, x)$ image-finite if both $x$ and the action model of $f$ are image-finite for all $I$-indexed relations. This ensures that $f^k(x)$ is image-finite for all $i \in I$, all $k \in \mathbb{N}$. An ETL model is image-finite if all its $I$-indexed relations are image-finite.

The second restriction concerns the temporal evolution, which is required to show finite variety:

**Definition 9 (Concluding DEL Dynamical System)** A pointed DEL dynamical system $(X, f, x)$ is **periodic** iff $f^k(x) = f^{k+m}(x)$ for some $k \geq 0$, $m > 0$. It **terminates** iff for some $k \in \mathbb{N}$, $f^k(x)$ is undefined. If it does either, it is said to **conclude**.

**Definition 10 (Concluding ETL Model)** A point $\underline{h} \in \underline{H}$ of a saturated ETL model $(\mathcal{H}, \underline{H})$ is **repeating** if there exists points $\underline{h}', \underline{h}'' \in \underline{H}$ with $\underline{h} \sqsubseteq \underline{h}' \sqsubset \underline{h}''$ and $C(\mathcal{H}\underline{h}')\underline{h}' \leftrightarrows C(\mathcal{H}\underline{h}'')\underline{h}''$. A point $\underline{h}$ is **finite** if there exists a point $\underline{h}'$ with $\underline{h} \sqsubseteq \underline{h}'$ while there is no $e \in E$ for which $\underline{h}'e \in H$. The model $(\mathcal{H}, \underline{H})$ **concludes** if every point in $\underline{H}$ is either repeating or finite.

Restricting attention to the classes of image-finite and concluding DEL dynamical systems and ETL models, a proper characterization result exists:

**Theorem 1** *A saturated ETL model $(\mathcal{H}, \underline{H})$ is image-finite, concluding and satisfies all eight properties of Def. 6 if, and only if, it is generatable by an image-finite and concluding pointed DEL dynamical system.*

*Proof Left-to-right:* The existence of a generating DEL dynamical system is guaranteed by Proposition 4. The constructions in the proof of Prop. 4 moreover ensure that the DEL dynamical system is both image-finite and concluding.

   *Right-to-left:* Proposition 2 ensures that the model will satisfy all eight properties, except maybe Component Collection Describable. Lemma 1 ensures the model is image-finite and concluding, which by Lemma 2 ensures that it does satisfy Component Collection Describable. Proofs of both lemmas are provided in the Appendix.

**Lemma 1** *If there exists an image-finite and concluding pointed DEL dynamical system that generates $(\mathcal{H}, \underline{H})$, then $(\mathcal{H}, \underline{H})$ is image-finite and concluding.*

**Lemma 2**  *If a saturated ETL model $(\mathcal{H}, \underline{H})$ is image-finite, concluding and satisfies Connected Time-Steps, then $(\mathcal{H}, \underline{H})$ is Component Collection Describable.*

*Remark 4* The converse of Lemma 2 does not hold, as Component Collection Describable does not imply image-finiteness.[15]

## 7 Non-Deterministic Intensional Protocols

In the previous sections, the ETL models regarded have been limited to single component branches as this is a requirement to be generatable from a DEL dynamical system – or a deterministic extensional DEL protocol. Extensional DEL protocols are in general non-deterministic and may therefore generate ETL models with multiple component branches, as e.g. the ETL model in Fig. 4. To facilitate comparison, this section is dedicated to non-deterministic intensional protocols, implemented as families of DEL dynamical systems running in parallel.

**Definition 11 (Component Branch Sub-Model)** Let $\mathcal{H} = (E, H, \mathtt{R}, V)$ be an ETL model and let $(\mathbf{b}, \underline{H})$ be a terminal component branch obtained from $\mathcal{H}$. The **component branch sub-model** of $\mathcal{H}$ given by $(\mathbf{b}, \underline{H})$ is then $\mathcal{H}_\mathbf{b} = (E_\mathbf{b}, H_\mathbf{b}, \mathtt{R}_{|H_\mathbf{b}}, V_{|H_\mathbf{b}}, \underline{H})_{i \in I}$ such that $E_\mathbf{b} = \{e \in E : e \in \mathbf{b}$ or $\exists h \in \mathbf{b}, he \in \mathbf{b}\}$, $H_\mathbf{b} = \{h \in H : h \in \mathbf{b}\}$ and $_{|H_\mathbf{b}}$ denotes restriction.

---

[15] Being Component Collection Describable does not imply being image-finite: Let the model have two components in one component branch. Let the root component $\mathbf{b}_0$ be image-infinite and satisfy $p$ at $\mathbf{b}_0\underline{h}$. Let $\mathbf{b}_1\underline{h'}$ satisfy $\neg p$. Then the model satisfies Component Collection Describable with $\varphi_{\{\underline{h}\}} := p$, $\varphi_{\{\underline{h'}\}} := \neg p$ and $\varphi_{\{\underline{h},\underline{h'}\}} := p \wedge \neg p$, but it is not image-finite.

*Property 4* If $\mathcal{H}$ is an ETL model and $(\mathbf{b}, \underline{H})$ obtained from $\mathcal{H}$ is both terminal and saturated, then $(\mathcal{H}_{\mathbf{b}}, \underline{H})$ is a saturated ETL model.

To (re-)produce ETL models that consist of more than one component branch, a family of dynamical systems each generating a terminal component branch of the ETL model is used. The complete ETL model is obtained by taking the union of all ETL component branches.

**Definition 12 (ETL Model Union)** Given a countable family of saturated ETL models $\{(\mathcal{H}_j, \underline{H}_j)\}_{j \in J}$ with each $(\mathcal{H}_j, \underline{H}_j) = (E_j, H_j, \mathtt{R}_j, V_j, \underline{H}_j)$ for $j \in J$, their (unpointed) **union model** is $U_J = (E_J, H_J, \mathtt{R}_J, V_J)$ with $\star_J := \bigcup_{j \in J} \star_j$ for $\star \in \{E, H, \mathtt{R}, V\}$.

An ETL model $\mathcal{H}$ is **generated by a family of pointed DEL dynamical systems** $\{(X_k, f_k, x_k)\}_{k \in K}$ iff each $(X_k, f_k, x_k)$ generates a saturated ETL model $(\mathcal{H}_k, \underline{H}_k)$ such that $\mathcal{H}$ is the union model of$\{(\mathcal{H}_k, \underline{H}_k)\}_{k \in K}$. The family $\{(X_k, f_k, x_k)\}_{k \in K}$ is **minimal** in generating $\mathcal{H}$ iff no proper subset of the family also generates $\mathcal{H}$.

**Lemma 3** *Let $\{(X_k, f_k, x_k)\}_{k \in K}$ be minimal in generating $\mathcal{H}$ and let $(X_k, f_k, x_k)$ generate the saturated ETL model $(\mathcal{H}_k, \underline{H}_k)$. Then $(\mathcal{H}_k, \underline{H}_k)$ is the component branch sub-model for some terminal component branch $\mathbf{b}$ of $\mathcal{H}$.*

*Proof* See Appendix.

**Theorem 2** *Let an image-finite and concluding ETL model $\mathcal{H}$ be given. $\mathcal{H}$ is generatable up to ETL isomorphism by a family of image-finite, concluding pointed DEL dynamical systems, if, and only if, there exists a saturation of each terminal component branch $\mathbf{b}$ of $\mathcal{H}$ such that $(\mathcal{H}_{\mathbf{b}}, \underline{H})$ satisfies all eight properties of Def. 6.*

*Proof* See Appendix.

7.1 Persistence Under Union

The properties in Theorem 2 are associated with the component branches of the ETL model, instead of the ETL model itself. However, several of the eight properties are not inherited from component branches to union model: Some are not defined for unpointed structures, and some are simply not robust under union. In the following final set of results linking DEL dynamical systems and ETL models, properties definable for general, unpointed ETL models are detailed.

**Lemma 4** *The saturated ETL model properties Synchronicity, Perfect Recall and Postcondition Describable persist under ETL model union. I.e.: Let $\{(\mathcal{H}_j, \underline{H}_j)\}_{j \in J}$ be a countable set of saturated ETL models. If all $(\mathcal{H}_j, \underline{H}_j)$ satisfy either of the mentioned properties, then the (unsaturated) union model $U_J$ satisfies that property.*

*Proof* See Appendix.

*Property 5* Local No Miracles, Precondition Describable, Point Bisimulation Invariance and Connected Time-Steps do not persist under union.

*Proof* See Appendix.

Though neither Local No Miracles nor Point Bisimulation Invariance persist under union, weaker versions of each property do recur in the union model: See Proposition 5 below. Both properties persevere as they are independent of structure outside a given component. They are therefore not affected by union. Local Bisimulation Invariance originates from [13] and is further discussed in Sec. 8.

**Definition 13 (ETL Model Properties)** An unsaturated ETL model $\mathcal{H} = (E, H, \mathtt{R}, V)$ satisfies

Very Local No Miracles *iff* $\forall h, h', h_1, h_2 \in H, \forall e, e' \in E : he, h'e' \in H$, if $h\mathtt{R}_i h'$, $h_1 e \mathtt{R}_i h_2 e'$, $he\mathtt{R}^* h_1 e$ and $h\mathtt{R}^* h_1$, then $he\mathtt{R}_i h'e'$;

Local Bisimulation Invariance *iff* for all $h, h' \in H$, $e \in E$, if $h$ and $h'$ are bisimilar, $h\mathtt{R}^* h'$ and $he \in H$, then $h'e \in H$.

**Proposition 5** *If an ETL model $\mathcal{H}$ is generated by a family of pointed DEL dynamical systems (possibly neither image-finite nor concluding), then $\mathcal{H}$ satisfies Synchronicity, Perfect Recall and Postcondition Describable and Very Local No Miracles and Local Bisimulation Invariance.*

*Proof* See Appendix.


## 8 Protocol Comparison

This paper is in line with the approach of van Benthem et al. [13] in investigating the generative power of DEL dynamical systems with respect to the class of ETL models. In this section, the above results are compared to those obtained in [13] relating DEL protocols to ETL models.


8.1 Generating ETL Models from DEL Protocols

Generating ETL models from DEL protocols is somewhat simpler than from sets of DEL dynamical systems. Unsaturated ETL forests are generated directly from a DEL protocol, without e.g. first defining saturated component branches. For the special case of uniform DEL protocols, an ETL model is generated from an initial pointed Kripke model as follows, cf. [13]:[16]

---

[16] The method for generating an ETL model from a state-dependent DEL protocol has a slightly more complex definition. As uniform DEL protocols is the case closest to the cases for DEL dynamical systems dealt with in this paper, the reader is referred to [13] for the definition for state-dependent DEL protocols.

**Definition 14 (ETL Model Generated from a Uniform DEL Protocol)** Let $\mathsf{p}$ be a uniform DEL protocol for the pointed Kripke model $Ms$, let $\rho = \rho_1...\rho_n \in \mathsf{p}(s)$ and let $(Ms)^\rho := (Ms \otimes \rho_1)... \otimes \rho_n$. The generated ETL model of $Ms$ and $\mathsf{p}$ is $\mathcal{H} = (E, H, \mathtt{R}, V)$ with $(H, \mathtt{R}, V) = \bigcup_{\rho \in \mathsf{p}} (Ms)^\rho$.

*Remark 5* Notice that no restriction to connected components is required posterior to taking products.

8.2 ETL Properties from DEL Protocols

The properties of ETL models generated from DEL protocols [13] results in a list of properties not identical to that of Def. 6. But there is overlap: Synchronicity, Perfect Recall and Local No Miracles. The remaining properties from [13] are Local Bisimulation Invariance, Propositional Stability and Finite Executions:

**Definition 15 (ETL Model Properties of [13])** An ETL model $\mathcal{H} = (E, H, \mathtt{R}, V)$ satisfies

Propositional Stability *iff* for all propositional formulas $p$ and for all $h \in H, e \in E$ such that $he \in H$, it holds that $h \in V(p)$ iff $he \in V(p)$;

Finite Executions *iff* for each $n$, for each $e \in E$, the set $\{h : he \in H$ and $len(h) = n\}$ is finite.

*Remark 6* The property Propositional Stability is required as [13] concerns action models *without postconditions*.[17] There is a comment on the resulting difference below.

Before relating DEL protocols to DEL dynamical systems and the ETL model properties they induce, recall the main results of [13].

**Theorem (Main Representation Theorem of [13])** *1) If an ETL model is generated by a uniform DEL protocol, then it satisfies the five properties Propositional Stability, Local Bisimulation Invariance, Synchronicity, Perfect Recall and Local No Miracles.*
*2) If an ETL model satisfies the six properties Finite Executions, Propositional Stability, Local Bisimulation Invariance, Synchronicity, Perfect Recall and Local No Miracles, then it is generatable by some uniform DEL protocol.*

**Theorem (Theorem 2 of [13])** An ETL model is generatable by a state-dependent DEL protocol iff it satisfies Propositional Stability, Synchronicity, Perfect Recall and Local No Miracles.

---

[17] Equivalently in the current setting would be action models with $\mathsf{post}(\sigma) = \top$ for all events $\sigma$.

## 8.3 Discussion and Comparison of DEL Protocols and DEL Dynamical Systems

With results established for both DEL dynamical systems and extensional DEL protocols, these may now be compared, first on a technical level concerning the induced properties, and second from a modeling perspective.–
For both DEL dynamical systems and DEL protocols the generated ETL model satisfies the core DEL properties Synchronicity, Perfect Recall and Local No Miracles. This is no surprise, as these properties – as was mentioned in Sec. 3.2 – stem from the very nature of product update. Beyond these, differences emerge:

*Connected Time-Steps:*

An ETL model generated using a single DEL dynamical system has connected time-steps as a consequence of the use of the restriction to connected components. This property does not survive model union, and is therefore not inherited by ETL models generated by families of DEL dynamical systems, cf. Remark 5. DEL protocols do not induce the property in generated ETL models, irrespective of whether such are defined using a restriction to connected components or not: DEL protocols may contain several sequences of action models, producing disjoint new time steps.

*Finite Executions vs. Precondition Describable:*

Finite Executions (referred to as "the finiteness assumption" in [13]) is meant to ensure the existence of the precondition formula of the action model event $\sigma_e$ for each ETL event $e$. Thus, it shares a role with the abstract Precondition Describable, but is weaker than this direct existence requirement. It is conjectured that a compilation error occurred post-submission of [13], omitting further requirements.[18]

*Propositional Stability vs. Postcondition Describable:*

That the theorems of van Benthem and co-authors include Propositional Stability is a result of their use of action models without postconditions. DEL dynamical systems limited to complex model transformers built over the same

---

[18] Finite Executions is not enough to guarantee the existence of suitable preconditions formulas: Let a single-agent ETL model $\mathcal{H}$ be given with histories of length 1 divided into two disconnected $\mathtt{R}_1$-components, $H_1$ and $H_1'$ with $e \in H_1$, and $e' \in H_1'$. Let the sub-model $H_1, H_1'$ be non-image-finite and non-bisimilar but let $(H_1, e)$ $(H_1', e')$ be modally equivalent. Such pointed Kripke model exist, cf. e.g. [15, Ex. 2.23, p. 68]. Let the set of histories of length 2 be given by $\{ee^*\}$ and let $\mathcal{H}$ contain no further histories. Then $\mathcal{H}$ satisfies Finite Executions (and the other properties), but there exists no suitable precondition formula for $\sigma_{e^*}$ as $e$ and $e'$ are modally equivalent, but $e^*$ only executed on $e$. An additional requirement of image-finiteness would solve this problem.

class of action models would generate ETL models also satisfying this property. Conversely, it is hypothesized that any ETL model generated by a DEL protocol defined over action models with postconditions would satisfy the abstract requirement of being Postcondition Describable by exhibiting only finite atomic change between successive histories.

### Component Collection Describable:

The Component Collection Describable requirement ensures the existence of suitable preconditions for the designated actions of the multi-pointed action model underlying the clean map, which control the temporal flow of the dynamical system when seeking to build a particular ETL model. This is not needed when working with DEL protocols, as the temporal occurrence of events is exogenously given. The requirement is not inherited by every ETL model build from a DEL dynamical system, but is implied when the system is aptly finite.

### Local vs. Point Bisimulation Invariance:

Whether generated by a uniform DEL protocol, single DEL dynamical system or a family of DEL dynamical systems, the resulting ETL model satisfies Local Bisimulation Invariance. This is due to the nature of preconditions in product update. Any saturated ETL model generated by a single DEL dynamical system satisfies the stronger property of Point Bisimulation Invariance, which also involves a temporal component, reflecting that clean maps are mappings acting on the points of pointed Kripke models. As the temporal invariance is defined on points, it is lost when moving to unsaturated models, exactly as these are unpointed. As DEL protocols react to an external clock rather than to the structure of the current pointed Kripke model, such protocols do not induce this strong version of bisimulation invariance.

### Protocols for Modeling Information Dynamics:

As mentioned in the introduction, logical modeling of dynamics in multi-agent systems relies on protocols as control mechanisms. With multiple protocol frameworks available, the question naturally arises which, if any, is better suited for a given modeling task. From a conceptual point of view, the property Point Bisimulation Invariance reflects that intensional protocols represented as clean maps are mappings, and hence output equivalent values given equivalent inputs. Again, this is the aspect of Point Bisimulation Invariance *not* shared with Local Bisimulation Invariance. In many multi-agents settings, this property is natural. E.g., in extensive games with imperfect information of game theory, it is standardly assumed that agents have *knowledge of their own actions*, i.e., that if two histories belong to the same information cell of

the agent, then the agent will choose the same action in the two nodes.[19] Such knowledge of own actions is guaranteed in DEL dynamical systems, but not in extensional DEL protocols.

Extensional DEL protocols are very handy for encoding extensional protocols: In case one wishes to answer a question concerning how a particular sequence of actions will influence a given initial model, then directly specifying that sequence of actions is a straightforward formalization. Dynamics thusly run on an external clock is not possible using DEL dynamical systems.[20]

In case one seeks to model an intensional protocol, possibly applicable to more than a single initial model, then DEL dynamical systems enjoys a particular benefit: As exemplified using the Muddy Children example, intensional natural language protocols may in a natural way be encoded as clean maps. Finally, as also exemplified, the clean map may be an incomparably more succinct representation of the protocol than the extensional DEL protocol counterpart.

## 9 Conclusion

In logically modeling dynamics in multi-agent systems, the main control mechanism of agents' actions are protocols. In choosing a class of protocols in which to cast a model, an implicit choice of agency, actions and dynamics is thus made. In this paper, an implementation of intensional protocols as DEL dynamical systems have been investigated. On the technical side, the type of epistemic temporal models that DEL dynamical systems may generate have been characterized. Conceptually, DEL dynamical systems *qua* intensional protocols have been compared to the main protocol framework in Dynamic Epistemic Logic, namely the extensional DEL protocols of [13]. In conclusion, DEL dynamical systems do not severely restrict the class of ETL models generatable. In addition, there are several situations where intensional protocols seem a reasonable choice over extensional protocols when modeling dynamic information phenomena using Dynamic Epistemic Logic. Given the popularity of intensional protocols in other multi-agent paradigms, it is surprising that such have not previously been systematically investigated for Dynamic Epistemic Logic.

---

[19] See e.g. [35] where this is assumed by definition or [25], where the requirement is denoted the *Ex Interim Condition*.

[20] A clock may however be build into DEL dynamical system: This may be done by working in an extended language with atomic propositions denoting the current time and using postconditions to make time run. For finite time sequences, this may be encoded using a finite model.

article was funded by the Swedish Research Council through the framework project 'Knowledge in a Digital World' (Erik J. Olsson, PI). We would like to thank Alexandru Baltag and Thomas Bolander for valuable comments on elements in this paper, as well as the participants of the 2016 CADILLAC workshop, held in Christiania, Denmark, and participants of the 2016 LogiCIC workshop, held in Amsterdam, the Netherlands.

## References

1. T. Ågotnes, H. van Ditmarsch, and Y. Wang. True lies. *Synthese*, May 2017.
2. S. Artemov, J. Davoren, and A. Nerode. Modal Logics and Topological Semantics for Hybrid Systems. Technical report, Cornell University, 1997.
3. A. Baltag and L. S. Moss. Logics for Epistemic Programs. *Synthese*, 139(2):165–224, 2004.
4. A. Baltag, L. S. Moss, and S. Solecki. The Logic of Public Announcements, Common Knowledge, and Private Suspicions (extended abstract). In *Proc. of the Intl. Conf. TARK 1998*, pages 43–56. Morgan Kaufmann Publishers, 1998.
5. A. Baltag and B. Renne. Dynamic epistemic logic. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2016 edition, 2016.
6. A. Baltag and S. Smets. A Qualitative Theory of Dynamic Interactive Belief Revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 9–58. Amsterdam University Press, 2008.
7. A. Baltag and S. Smets. Group Belief Dynamics Under Iterated Revision: Fixed Points and Cycles of Joint Upgrades. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '09, pages 41–50, New York, NY, USA, 2009. ACM.
8. J. van Benthem. "One is a Lonely Number": Logic and Communication. In Z. Chatzidakis, P. Koepke, and W. Pohlers, editors, *Logic Colloquium '02*, Lecture Notes in Logic, 27, pages 95–128. Association for Symbolic Logic, 2002.
9. J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011.
10. J. van Benthem. Oscillations, Logic, and Dynamical Systems. In S. Ghosh and J. Szymanik, editors, *The Facts Matter*, pages 9–22. College Publications, 2016.
11. J. van Benthem and C. Dégremont. Bridges between dynamic doxastic and doxastic temporal logics. In *Logic and the Foundations of Game and Decision Theory–LOFT 8*, pages 151–173. Springer, 2010.
12. J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, 2006.
13. J. van Benthem, J. Gerbrandy, T. Hoshi, and E. Pacuit. Merging Frameworks for Interaction. *Journal of Philosophical Logic*, 38(5):491–526, 2009.
14. J. van Benthem and S. Smets. Dynamic Logics of Belief Change. In H. van Ditmarsch, J. Y. Halpern, W. van der Hoek, and B. Kooi, editors, *Handbook of Logics for Knowledge and Belief*, pages 299–368. College Publications, 2015.
15. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
16. T. Bolander and M. Birkegaard. Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1):9–34, 2011.
17. T. Bolander, M. H. Jensen, and F. Schwarzentruber. Complexity Results in Epistemic Planning. In *Proceedings of IJCAI 2015 (24th International Joint Conference on Artificial Intelligence)*, 2015.
18. C. Dégremont. *The Temporal Mind: Observations on the logic of belief change in interactive systems*. PhD thesis, University of Amsterdam, 2010.
19. H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2008.

20. H. van Ditmarsch and B. Kooi. Semantic Results for Ontic and Epistemic Change. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 87–117. Amsterdam University Press, 2008.
21. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge.* The MIT Press, 1995.
22. D. Fernández-Duque. A sound and complete axiomatization for Dynamic Topological Logic. *Journal of Symbolic Logic*, 77(3):1–26, 2012.
23. D. Fernández-Duque. Dynamic topological logic of metric spaces. *Journal of Symbolic Logic*, 77(1):308–328, 2012.
24. D. Fernández-Duque. Non-Finite axiomatizability of dynamic topological logic. *ACM Transactions on Computational Logic*, 15(1):1–18, 2014.
25. P. Galeazzi and E. Lorini. Epistemic logic meets epistemic game theory: a comparison between multi-agent Kripke models and type spaces. *Synthese*, 193(7):2097–2127, 2016.
26. J. Gerbrandy and W. Groeneveld. Reasoning about Information Change. *Journal of Logic, Language and Information*, 6(2):147–169, 1997.
27. V. Goranko and M. Otto. Model Theory of Modal Logic. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier, 2008.
28. J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions.* College Publications, 2nd, 2005 edition, 1962.
29. T. Hoshi. *Epistemic Dynamics and Protocol Information.* PhD thesis, ILLC, Universiteit van Amsterdam, 2009.
30. T. Hoshi. Merging DEL and ETL. *Journal of Logic, Language and Information*, 19(4):413–430, 2010.
31. D. Klein and R. K. Rendsvig. Convergence, Continuity and Recurrence in Dynamic Epistemic Logic. In A. Baltag and J. Seligman, editors, *Logic and Rational Interaction (LORI VI)*, Lecture Notes in Computer Science. Springer, 2017.
32. P. Kremer and G. Mints. Dynamical Topological Logic. *Bulleting of Symbolic Logic*, 3:371–372, 1997.
33. P. Kremer and G. Mints. Dynamic Topological Logic. In M. Aiello, I. Pratt-Hartmann, and J. van Benthem, editors, *Handbook of Spatial Logics*, chapter 10, pages 565–606. 2007.
34. R. van der Meyden. Constructing finite state implementations of knowledge-based programs with perfect recall. In *Intelligent Agent Systems Theoretical and Practical Issues*, pages 135–151. Springer, 1997.
35. M. J. Osbourne and A. Rubinstein. *A Course in Game theory.* The MIT Press, 1994.
36. R. Parikh and R. Ramanujam. A knowledge based semantics of messages. *Journal of Logic, Language and Information*, 12(4):453–467, Sep 2003.
37. J. A. Plaza. Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.
38. R. K. Rendsvig. Aggregated Beliefs and Informational Cascades. In D. Grossi and O. Roy, editors, *Logic, Rationality, and Interaction*, Lecture Notes in Computer Science, pages 337–341. Springer, 2013.
39. R. K. Rendsvig. Pluralistic ignorance in the bystander effect: Informational dynamics of unresponsive witnesses in situations calling for intervention. *Synthese*, 191(11):2471–2498, 2014.
40. R. K. Rendsvig. Model Transformers for Dynamical Systems of Dynamic Epistemic Logic. In W. van der Hoek, W. H. Holliday, and W. F. Wang, editors, *Logic, Rationality, and Interaction (LORI 2015, Taipei)*, LNCS, pages 316–327. Springer, 2015.
41. B. Rodenhäuser. A logic for extensional protocols. *Journal of Applied Non-Classical Logics*, 21(3-4):477–502, 2011.
42. D. Sarenac. Modal Logic for Qualitative Dynamics. In O. Roy, P. Girard, and M. Marion, editors, *Dynamic Formal Epistemology*, volume 351 of *Synthese Library*, pages 75–101. Springer, 2011.
43. Y. Wang. *Epistemic Modelling and Protocol Dynamics.* Doctoral thesis, Universiteit van Amsterdam, 2010.
44. M. Wooldridge and A. Lomuscio. Reasoning about visibility, perception, and knowledge. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 1–12. Springer, 1999.

## Appendix: Proofs

**Proposition 1.** Let $(X, f, x)$ be a pointed DEL dynamical system given by multi-pointed action model $\Sigma\Gamma$. Then there exists a singleton uniform DEL protocol that produces the orbit of $f$ from $x$.

*Proof* For each $k \in \mathbb{N}$, let $\sigma_k \in \Gamma$ be such that $f^k(x) \models \mathsf{pre}(\sigma_k)$. As $\Sigma\Gamma$ is $X$-deterministic, for each $k$ there is at most one such $\sigma_k$. Define a uniform DEL protocol $\mathsf{p}$ as the smallest protocol for which $\mathsf{p}_k(s) = \Sigma\sigma_k$ (for all $s \in x$) whenever $\sigma_k$ exists. Then when $\mathsf{p}$ is sequentially applied to $x$ using product update, it produces the sequence $\langle f^k(x) \rangle_{k \in \mathbb{N}}$ of pointed Kripke models (up to the deletion of redundant states not connected to the designated points, cf. Sec. 2). $\square$

**Proposition 2.** If saturated ETL model $(\mathcal{H}, \underline{H})$ is generated by a pointed DEL dynamical system, then $(\mathcal{H}, \underline{H})$ satisfies seven of the eight properties of Def. 6, namely Synchronicity, Connected Time-Steps, Perfect Recall, Local No Miracles, Precondition Describable, Postcondition Describable, and Point Bisimulation Invariance.

*Proof* Let $(X, f, x)$ be a pointed DEL dynamical system with orbit $(f^k(x))_{k \in \mathbb{N}}$. The **length** of a state $s$ in $f^k(x)$ is $len(s) := k + 1$.

Let $(\mathcal{H}, \underline{H})$ be the saturated ETL model generated by $(X, f, x)$. Given the construction of $\gamma$ in Def. 7, there exists a family of isomorphisms $\{g_k\}_{k \in \mathbb{N}}$ with each $g_k$ mapping $\llbracket f^k(x) \rrbracket$ to $H_k := \{h \in H : len(h) = k\}$ satisfying $g_1(s) = e_s$ and $g_{k+1}((s, \sigma)) = g_k(s)e_\sigma$. Using this family, it is shown that $(\mathcal{H}, \underline{H})$ satisfies the listed properties in order:

**Synchronicity.** Assume for arbitrary $h, h' \in H$ that $h\mathsf{R}_i h'$. Then by the construction of the generated $\mathsf{R}_i$ (Def. 7), $\exists k \in \mathbb{N} : g_k^{-1}(h)R_i g_k^{-1}(h')$. Hence $g_k^{-1}(h), g_k^{-1}(h') \in \llbracket f^k(x) \rrbracket$. Thus, $len(g_k^{-1}(h)) = len(g_k^{-1}(h'))$. Hence, by the construction of $g$, $len(h) = len(h')$.

**Perfect Recall.** Assume for arbitrary $he, h'e' \in H$ that $he\mathsf{R}_i h'e'$. Then $\exists k \in \mathbb{N} : g_k^{-1}(he)R_i g_k^{-1}(h'e')$. By construction of $f$, $f^k(x) = C(f^{k-1}(x) \otimes \Sigma\Gamma)s'$ for $\Sigma\Gamma$ the multi pointed action model. As $g_k^{-1}(he)R_i g_k^{-1}(h'e')$, by definition of $\otimes$ and clean maps, $g_{k-1}^{-1}(h)R_i g_{k-1}^{-1}(h')$. Hence, by definition of $\mathsf{R}_i$, it follows that $h\mathsf{R}_i h'$.

**Local No Miracles.** Assume that 1) $h\mathsf{R}_i h'$, 2) $h_1 e\mathsf{R}_i h_2 e'$ and 3) $h\mathsf{R}^* h_1$ for arbitrary $he, h'e', h_1 e, h_2 e' \in H$. 1) implies that 1*) $g_k^{-1}(h)R_i g_k^{-1}(h')$ for $k = len(h)$. 3) implies that $len(h) = len(h_1)$ by Synchronicity. In conjunction with 2), this implies that 2*) $g_{k+1}^{-1}(h_1 e)R_i g_{k+1}^{-1}(h_2 e')$.

By construction of $f$, $f^{k+1}(x) = C(f^k(x) \otimes \Sigma\Gamma)s'$. By 2*) and the definition of $\otimes$ and clean maps, there must be 4) $\sigma_e, \sigma_{e'} \in \Sigma\Gamma$ such that $\sigma_e R_i \sigma_{e'}$, for $\sigma_e$ the $\sigma$ such that $g_{k+1}((s, \sigma)) = h_1 e$, for some $s \in \llbracket f^k(x) \rrbracket$, and $\sigma_{e'}$ the $\sigma'$ such that $g_{k+1}((t, \sigma')) = h_2 e'$, for some $t \in \llbracket f^k(x) \rrbracket$.

Now assume that $(g_k^{-1}(h), \sigma_e), (g_k^{-1}(h'), \sigma_{e'}) \in [\![f^{k+1}(x)]\!]$. Then 1\*), 4) and Def. $\otimes$ jointly imply that $(g_{k+1}^{-1}(h), \sigma_e)R_i(g_{k+1}^{-1}(h'), \sigma_{e'})$. By the definition of the generated $R_i$, it thus follows that $heR_ih'e'$.

**Connected Time-Steps.** For arbitrary $h, h' \in H$ assume $len(h) = len(h')$. Let $k \in \mathbb{N}$ such that $h, h' \in H_k$. Then $g_k^{-1}(h), g_k^{-1}(h') \in [\![f^k(x)]\!]$. By definition of product update $\otimes$, clean maps and the fact that $x$ is connected, $g_k^{-1}(h)R^*g_k^{-1}(h')$. By definition of $(\mathcal{H}, \underline{H})$ it follows that $hR^*h'$.

**Precondition Describable.** For arbitrary $e \in E$, let $\delta_e = \mathsf{pre}(\sigma_e)$. Recall that by definition of $\otimes$, $\forall k \in \mathbb{N}$, $(g_k^{-1}(h), \sigma_e) \in [\![f^{k+1}(x)]\!]$ iff $g_k^{-1}(h) \models \mathsf{pre}(\sigma_e)$ and $\sigma_e \in \Sigma^{k+1}$ ($*$). Assume $\exists h' \in H : h'e \in H$ and let $h \in H$ such that $h'R^*h$.

$\Rightarrow$: Assume for some $k \in \mathbb{N}$ that $(H_k, h) \models \delta_e$. Then $g_k^{-1}(h) \models \mathsf{pre}(\sigma_e)$. By assumption, $\sigma_e \in \Sigma_{k+1}$. By ($*$), thus $(g_k^{-1}(h), \sigma_e) \in [\![f^{k+1}(x)]\!]$. Therefore, $he \in H$.

$\Leftarrow$: Assume $he \in H$. Then for some $k \in \mathbb{N}$, $g_{k+1}^{-1}(he) = (g_k^{-1}(h), \sigma_e) \in [\![f^{k+1}(x)]\!]$. By ($*$), $g_k^{-1}(h) \models \mathsf{pre}(\sigma_e)$. And thus $h \models \delta_e$.

**Postcondition Describable.** For arbitrary $he \in H$ (in specific for some $k \in \mathbb{N}$, $he \in H_{k+1}$) let $\delta_{D_e} = \mathsf{post}(\sigma_e)$ where $D_e = D_1 \cup D_2$, for $p, q \in \Phi$ such that $D_1 = \{p : h \notin V(p), he \in V(p)\}$ and $D_2 = \{\neg q : h \in V(q), he \notin V(q)\}$.

Consider an arbitrary $p \in D_1$. Then by definition of the generated ETL model $(\mathcal{H}, \underline{H})$, $g_k^{-1}(h) \notin [\![p]\!]_k$ and $g_{k+1}^{-1}(he) \in [\![p]\!]_{k+1}$ ($*$). By construction, $g_{k+1}^{-1}(he) = (g_k^{-1}(h), \sigma)$ for some $\sigma \in [\![\Sigma]\!]_{k+1}$ ($**$). By definitions of $(\mathcal{H}, \underline{H})$ and $\otimes$, $(g_k^{-1}(h), \sigma) \in [\![p]\!]_{k+1}$ iff $\mathsf{post}(\sigma) \models p$. Then, by ($*$) and ($**$), $\mathsf{post}(\sigma) \models p$. As $p \in D_1$ was arbitrary, $\mathsf{post}(\sigma) \models D_1$.

The argument for $\mathsf{post}(\sigma) \models D_2$ is identical. Conclude that $\mathsf{post}(\sigma) \models D_e$ and thus $\delta_{D_e} \models D_e$.

**Point Bisimulation Invariance.** Let be given arbitrary $C(\mathcal{H}h)\underline{h} = (H_k, \underline{h})$ and $C(\mathcal{H}h')\underline{h'} = (H_l, \underline{h'})$ such that $(H_k, \underline{h}) \leftrightarrow (H_l, \underline{h'})$. Then $f^k(x) \leftrightarrow f^l(x)$ ($*$).

Further, assume for arbitrary $h \in H_k$ and $h' \in H_l$ that $(H_k, h) \leftrightarrow (H_l, h')$.

$\Rightarrow$: Assume $he \in H$. By construction of $g$ and definition of clean maps, both $H_k$ and $H_l$ are connected components, i.e, $\forall h, h' \in H_k : hR^*h'$ and idem for $H_l$. By the Hennessy-Milner Theorem (see Section 6), it follows that $h$ and $h'$ satisfy exactly the same modal formulas. Hence, by construction of $g$ and the definition of $\mathcal{H}$, $g_k^{-1}(h)$ and $g_l^{-1}(h')$ satisfy exactly the same modal formulas as well. Now as $he \in H$, $g_{k+1}^{-1}(he) \in [\![f^{k+1}(x)]\!]$ and thus $g_k^{-1}(h) \models \mathsf{pre}(\sigma_e)$. Hence $g_l^{-1}(h') \models \mathsf{pre}(\sigma_e)$ ($**$). By ($*$) and ($**$), it follows that $(g_l^{-1}(h'), \sigma_e) \in [\![f^{l+1}(x)]\!]$. Hence $h'e \in H$.

$\Leftarrow$: By the same argument, $h'e \in H$ implies $he \in H$.

This concludes the proof of Proposition 2. $\quad\square$

**Property 3.** Not all saturated ETL models generated by DEL dynamical systems are Component Collection Describable.

*Proof* Let $Ms$ and $f$ be as in Fig. 5 (cf. [40]) and $X$ be the orbit of $f$ from $Ms$. Then the ETL model generated by the DEL dynamical system $(X, f)$ from initial model $Ms$ is not Component Collection Describable.
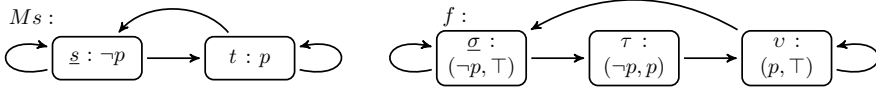


**Fig. 5** Initial Kripke model $Ms$ and pointed action model. The orbit of $f$ from $Ms$ produces non-bisimilar models forever: the unique state not satisfying $p$ will split, inserting a new $p$-state as it's child with $\tau$; any other state gets exactly one child:
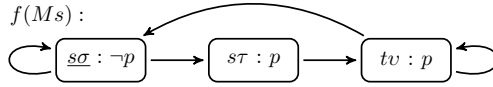


Consider $A = \{f^n(Ms) : n$ is even$\}$. There does not exist a $\varphi$ such that for all $x \in X$, $x \models \varphi$ iff $x \in A$: Assume the modal depth of $\varphi$ is $k$. Let $m > k$. Then $f^m(Ms) \models \varphi$ iff $f^{m+1}(Ms) \models \varphi$ as such two models will not differ in the first $m + 1$ relational steps from the point. Hence the ETL model generated by $(X, f)$ from $Ms$ will not be Component Collection Describable.  □

**Proposition 4.** If $(\mathcal{H}, \underline{H})$ is a saturated ETL model that satisfies all eight properties of Def. 6, then there exists a pointed DEL dynamical system that generates $(\mathcal{H}, \underline{H})$.

*Proof* Proposition 4 is shown by constructing a DEL dynamical system $(X, f)$ with $f$ the clean map of a $X$-deterministic multi-pointed action model $\Sigma\Gamma$ and an initial Kripke model $x \in X$ such that the saturated ETL model $(\mathcal{H}', \underline{H}') = (E', H', \mathsf{R}', V', \underline{H}')$ generated by $(X, f)$ from $x$ is ETL isomorphic to $(\mathcal{H}, \underline{H})$. The latter is shown by induction on $len(h)$ of $h \in \mathcal{H}$ for a map $\gamma^* : E \longrightarrow E'$. As in Def. 5, for $h = e_0...e_n$, write $\gamma^*(h) := \gamma^*(e_0)...\gamma^*(e_n)$.

As $(\mathcal{H}, \underline{H})$ satisfies property Connected Time-Steps, the ETL model is a saturated terminal component branch. To emphasize this, in this proof $(\mathcal{H}, \underline{H})$ is written $(\mathcal{H}_{\mathbf{b}}, \underline{H})$.

**1. Initial Kripke model**
To obtain a practical and consistent naming of states, the initial Kripke model $x = (\llbracket x \rrbracket, R, \llbracket \cdot \rrbracket, s)$ is set to be a re-naming of the initial component of $\mathbf{b}$: Let $\llbracket x \rrbracket = \{\sigma_e : e \in \mathbf{b}_0\underline{h}\}$. For the relations and valuation of the initial model, simply copy over the relations and valuation from the initial component of $\mathbf{b}$: For all $i \in \mathcal{A}$, let $\sigma_e R_i \sigma_{e'}$ iff $e\mathsf{R}_i e'$, and for all $p \in \Phi$, let $\sigma_e \in \llbracket p \rrbracket$ iff $e \in V(p)$. Finally, let the point of $x$ be the copy of the point of $\mathbf{b}_0\underline{h}$: Let $s = \sigma_{\underline{h}}$.

**2. Constructing $(X, f)$**
To define the DEL dynamical system $(X, f)$, first construct a multi-pointed action model $\Sigma\Gamma = (\llbracket \Sigma \rrbracket, \mathsf{R}, pre, post, \Gamma)$. Construct $\Sigma\Gamma$ so that for each time-step $\mathbf{b}_k$ of the component branch $\mathbf{b}$, it will contain a designated action $\underline{\sigma}_k$

and associated component that will from the Kripke model-equivalent of $\mathbf{b}_k\underline{h}$ produce the equivalent if $\mathbf{b}_{k+1}\underline{h}$. In the precondition of $\underline{\sigma}_k$, include a formula $\delta_{\mathbf{b}_k\underline{h}}$ characterizing $\mathbf{b}_k\underline{h}$. As $(\mathcal{H}_\mathbf{b}, \underline{H})$ is Component Collection Describable by assumption, such a formula exists.

Construct $\Sigma\Gamma$ piecewise as follows: Let $\mathbf{b}_k\underline{h}$ and $\mathbf{b}_{k+1}\underline{h}$ be given. $\Sigma_{k+1}\sigma$ is constructed such that $C(f^k(x) \otimes \Sigma_k\sigma)s'$ mirrors the structure of $\mathbf{b}_{k+1}\underline{h}$:

Let the single-pointed action model $\Sigma_{k+1}\underline{\sigma}_{k+1}$ be $(\llbracket\Sigma_{k+1}\rrbracket, \mathsf{R}_{k+1}, \mathsf{pre}_{k+1}, \mathsf{post}_{k+1}, \underline{\sigma}_{k+1})$, given by

$\llbracket\Sigma_{k+1}\rrbracket = \{\sigma_e : he \in \mathbf{b}_{k+1}\}$ with $\underline{\sigma}_{k+1} = \sigma_e$ such that $\exists h : he \in \mathbf{b}_{k+1} \cap \underline{H}$.

$(\sigma_e, \sigma_{e'}) \in \mathsf{R}_i$ iff $\exists he, h'e' \in \mathbf{b}_{k+1} : (he, h'e') \in \mathsf{R}_i$.

$$pre(\sigma_e) = \begin{cases} \delta_e & \text{if } \sigma_e \neq \underline{\sigma}_{k+1} \\ \delta_e \wedge \delta_{\mathbf{b}_k\underline{h}} & \text{if } \sigma_e = \underline{\sigma}_{k+1} \end{cases}$$

with $\delta_e$ and $\delta_{\mathbf{b}_k\underline{h}}$ given by Precondition Describable and Component Collection Describable, respectively.

$\mathsf{post}(\sigma_e) = \delta_{D_e}$ as given Postcondition Describable.

Let the multi-pointed action model $\Sigma\Gamma = (\llbracket\Sigma\rrbracket, \mathsf{R}, pre, post, \Gamma)$ be given by, for $\ast \in \{\llbracket\Sigma\rrbracket, \mathsf{R}, pre, post, \}$, $\ast = \bigcup_{k:\mathbf{b}_k\in\mathbf{b}} \ast_k$ and $\Gamma = \bigcup_{k:\mathbf{b}_k\in\mathbf{b}} \{\underline{\sigma}_k\}$. This is well-defined: For $pre$ and $post$, this follows from Precondition Describable and Postcondition Describable.

Let $X$ be the closure of $\{x\}$ under the operation $\otimes\Sigma\Gamma$. On $X$, $\Sigma\Gamma$ is guaranteed deterministic as any two characteristic formulas $\delta_{\mathbf{b}_k\underline{h}}$ and $\delta'_{\mathbf{b}_{k'}\underline{h'}}$ will be mutually excluding. Finally, let $f$ be the clean map of $\Sigma\Gamma$ on $X$. Then $(X, f)$ is a DEL dynamical system with $x \in X$.

### 3. Constructing the Isomorphism

Let $(\mathcal{H}', \underline{H}') = (E', H', \mathsf{R}', V', \underline{H}')$ be the saturated ETL model generated by $(X, f)$ from $x$. Define the two mappings: $\gamma^\dagger : E \longrightarrow \llbracket\Sigma\rrbracket$ with $\gamma^\dagger(e) = \sigma_e$ and $\gamma : \llbracket\Sigma\rrbracket \longrightarrow E'$ for $\gamma(\sigma) = e_\sigma$ cf. Def. 7. From these, define $\gamma^* : E \longrightarrow E'$ as $\gamma^* := \gamma \circ \gamma^\dagger$.

Define subsets of $E$ based on history length: For all $k \in \mathbb{N}$ let $E_0 = \{e : e \in \mathbf{b}_0\}$ and $E_{k+1\geq1} = \{e : h \in \mathbf{b}_k, he \in \mathbf{b}_{k+1}\} \subseteq E$. Let $E'_k, k \in \mathbb{N}$, be given *mutatis mutandis*. Let $\gamma^\dagger_k : E_k \longrightarrow \llbracket\Sigma_k\rrbracket$, $\gamma_k : \llbracket\Sigma_k\rrbracket \longrightarrow E'_k$ and $\gamma^*_k : E_k \longrightarrow E'_k$ be the restrictions of $\gamma^\dagger, \gamma$ and $\gamma^*$ to $E_k \times \llbracket\Sigma_k\rrbracket$, $\llbracket\Sigma_k\rrbracket \times E'_k$ and $E_k \times E'_k$, respectively. Then, of course, $\gamma^* = \bigcup_{k\in\mathbb{N}} \gamma^*_k$, whereby $\gamma^*(e) = e'_{\sigma_e}$. By induction on $len(h)$, $h \in \mathcal{H}$, it is now shown that $\gamma^*$ is an ETL isomorphism.

**Claim: The map $\gamma^*$ is a bijection.** By the construction of $\gamma^\dagger$ and $\gamma$, each $\gamma^*_k$ is an injection: if $e \neq e'$, then $\gamma^*_k(e) \neq \gamma^*_k(e')$. By construction, it is also guaranteed that $\gamma^*_k$ is a surjection: $\forall e' \in E' \exists e \in E : \gamma^*_k(e) = e'$. Hence for each $k \in \mathbb{N}$, $\gamma^*_k$ is a bijection.

Furthermore, $\gamma^*$ is a total map: if $e \in E_k \cap E_m$, then $\gamma_k^\dagger(e) = \gamma_m^\dagger(e)$ (i.e., $\sigma_e \in [\![\Sigma_k]\!] \cap [\![\Sigma_m]\!]$) and $\gamma_k(\sigma_e) = \gamma_m(\sigma_e)$. Thus $\gamma \circ \gamma^\dagger(e)$ is well-defined and in $E'_k \cap E'_m$.

Finally, $\gamma^*$ inherits injectivity and surjectivity from its restrictions. Hence, the map $\gamma^*$ is a bijection.

**Claim: The map $\gamma^*$ is an ETL isomorphism.** The claim is shown by 4 inductive sub-proofs.

*1) Domain and Temporal Structure.*
**Base.** Let $h \in H_0$. This is the case iff $\gamma^\dagger(h) \in f^0(x)$ (by construction of initial Kripke model) iff $\gamma \circ \gamma^\dagger(h) \in H'_0$ (by Def. 7).
**Step.** It is shown that $he \in H_{k+1}$ iff $\gamma^*(he) \in H'_{k+1}$.

$\Rightarrow$: Assume $he \in H_{k+1}$. Then $h \in H_k$. By the induction hypothesis, $\gamma^\dagger(h) \in [\![f^k(x)]\!]$. By construction of $\Sigma_{k+1}$, $\gamma^\dagger(e) \in [\![\Sigma_{k+1}]\!]$. By the same construction and Precondition Describable, $\gamma^\dagger(h) \models \mathsf{pre}(\sigma_e)$. Hence $(\gamma^\dagger(h), \gamma^\dagger(e)) \in [\![f^{k+1}(x)]\!]$. By Def. 7, in particular the construction of $H_{k+1}$, $\gamma((\gamma^\dagger(h), \gamma^\dagger(e))) \in H'_{k+1}$.

$\Leftarrow$: Assume $\gamma((\gamma^\dagger(h), \gamma^\dagger(e))) \in H'_{k+1}$. Then $(\gamma^\dagger(h), \gamma^\dagger(e)) \in [\![f^{k+1}(x)]\!]$ by Def. 7, so $\gamma^\dagger(h) \in [\![f^k(x)]\!]$ and $\gamma^\dagger(e) \in [\![\Sigma_{k+1}]\!]$. By the induction hypothesis, $h \in H_k$. If $he \notin H_{k+1}$, a contradiction is reached: $\mathsf{pre}(\gamma^\dagger(e))$ is satisfied by exactly those $\gamma^\dagger(h) \in [\![f^k(x)]\!]$ such that $(\gamma^\dagger(h), \gamma^\dagger(e)) \in [\![f^{k+1}(x)]\!]$ – by the construction of action models in this proof and as $(\mathcal{H}_\mathbf{b}, \underline{H})$ is Precondition Describable. So $he \in H_{k+1}$.

*2) Epistemic relations.*
**Base.** It follows by construction of initial Kripke model and Def. 7.
**Step.** It is shown that $\forall he, h'e' \in \mathbf{b}_{k+1}$, $he\mathtt{R}_i h'e'$ iff $\gamma^*(he)\mathtt{R}'_i \gamma^*(h'e')$.

$\Rightarrow$: Assume that $he\mathtt{R}_i h'e'$. By Perfect Recall, $h\mathtt{R}_i h'$. By the induction hypothesis, $\gamma^\dagger(h)R_i\gamma^\dagger(h')$. By construction of $\Sigma_{k+1}$, $\gamma^\dagger(e)\mathtt{R}_i\gamma^\dagger(e')$. By definition of $\otimes$, $(\gamma^\dagger(h), \gamma^\dagger(e))R_i(\gamma^\dagger(h'), \gamma^\dagger(e'))$. By Def. 7, $\gamma((\gamma^\dagger(h), \gamma^\dagger(e)))\mathtt{R}_i\gamma((\gamma^\dagger(h'), \gamma^\dagger(e')))$.

$\Leftarrow$: Assume $\gamma((\gamma^\dagger(h), \gamma^\dagger(e)))\mathtt{R}_i\gamma((\gamma^\dagger(h'), \gamma^\dagger(e')))$. By Def. 7, $(\gamma^\dagger(h), \gamma^\dagger(e))R_i(\gamma^\dagger(h'), \gamma^\dagger(e'))$. By definition of $\otimes$, both $\gamma^\dagger(h)R_i\gamma^\dagger(h')$ and $\gamma^\dagger(e)\mathtt{R}_i\gamma^\dagger(e')$. So by construction of $\Sigma_{k+1}$, $\exists h_1 e, h_2 e' \in \mathbf{b}_{k+1} : h_1 e\mathtt{R}_i h_2 e'$. By the induction hypothesis, $h\mathtt{R}_i h'$. Further, note that $\forall h, h' \in \mathbf{b}_k : h\mathtt{R}^* h'$. Hence, by Local No Miracles, $he\mathtt{R}_i h', e'$.

*3) Valuation.*
**Base.** It follows by construction of initial Kripke model and Def. 7.
**Step.** It is shown that $\forall he \in \mathbf{b}_{k+1}$, $he \in V(p)$ iff $\gamma(he) \in V'(p)$.

$\Rightarrow$: Assume $he \in V(p)$. Either i) $h \in V(p)$ or ii) $h \notin V(p)$. If i), then by the induction hypothesis, $\gamma \circ \gamma^\dagger(h) \in V'(p)$. By construction of $\Sigma_{k+1}$, $\mathsf{post}(\gamma^\dagger(e)) \not\models \neg p$. Hence $(\gamma^\dagger(h), \gamma^\dagger(e)) \in [\![p]\!]_{k+1}$. By Def. 7, $\gamma((\gamma^\dagger(h), \gamma^\dagger(e))) \in V'(p)$. If ii), then by the induction hypothesis, $\gamma^\dagger(h) \notin [\![p]\!]_k$. As $(\mathcal{H}_\mathbf{b}, \underline{H})$ is Postcondition Describable, by construction of $\Sigma_{k+1}$, $\mathsf{post}(\gamma^\dagger(e)) \models p$. Thus $(\gamma^\dagger(h), \gamma^\dagger(e)) \in [\![p]\!]_{k+1}$. By Def. 7, $\gamma((\gamma^\dagger(h), \gamma^\dagger(e))) \in V'(p)$.

$\Leftarrow$: Assume $\gamma((\gamma^\dagger(h), \gamma^\dagger(e))) \in V'(p)$. Then $(\gamma^\dagger(h), \gamma^\dagger(e)) \in [\![p]\!]_{k+1}$ by Def. 7. Again, either i) $\gamma^\dagger(h) \in [\![p]\!]_k$ or ii) $\gamma^\dagger(h) \notin [\![p]\!]_k$. If i), then by the induction hypothesis, $h \in V(p)$. For a contradiction, suppose $he \notin V(p)$. Then $\mathsf{post}(\gamma^\dagger(e)) \models \neg p$. But by construction of $\Sigma_{k+1}$, $\mathsf{post}(\gamma^\dagger(e)) \not\models \neg p$. This is a contradiction. Hence $he \in V(p)$. If ii), then by the definition of $\otimes$, $\mathsf{post}(\gamma^\dagger(e)) \models p$. By the induction hypothesis, $h \notin V(p)$. If it was the case that $he \notin V(p)$, then $\mathsf{post}(\gamma^\dagger(e)) \not\models p$. Contradiction. Thus $he \in V(p)$.

*4) Points.*
**Base.** It follows by construction of initial Kripke model and Def. 7.
**Step.** It is shown that $\underline{he} \in \underline{H}_{k+1}$ iff $\gamma^*(\underline{he}) \in \underline{H}'_{k+1}$. Let $f^k(x) = Nt$ and $f^{k+1}(x) = Ms$.
$\Rightarrow$: Assume $\underline{he} \in \underline{H}_{k+1}$. By the induction hypothesis, $\gamma^\dagger(\underline{h}) = t$. By saturation of $\mathbf{b}_k$, $\exists e \in E : \underline{he} \in \mathbf{b}_{k+1} \cap \underline{H}$. By construction of $\Sigma_{k+1}$, $\gamma^\dagger(e) \in [\![\Sigma_{k+1}]\!]$ and, as $(\mathcal{H}_\mathbf{b}, \underline{H})$ is Precondition Describable, $\gamma^\dagger(\underline{h}) \models \mathsf{pre}(\gamma^\dagger(e))$. By construction of $f$, $f^{k+1}(x) = C(Nt \otimes (\Sigma_{k+1}, \gamma^\dagger(e))) = C(N \otimes \Sigma_{k+1}, (t, \gamma^\dagger(e))) = Ms$. By Def. 7, $\gamma(s) = \gamma(\gamma^\dagger(\underline{he})) \in \underline{H}'_{k+1}$.
$\Leftarrow$: Assume $\gamma(\gamma^\dagger(\underline{he})) \in \underline{H}'_{k+1}$. By Def. 7, $f^{k+1}(x) = (C(N \otimes \Sigma_{k+1}), (t, \gamma^\dagger(\underline{e})))$. By induction hypothesis, $\gamma^{\dagger-1}(t) = h \in \underline{H}_k$. By construction of the Action Model, $\gamma^\dagger(\underline{e}) = \gamma^{-1}(e)$ for $e$ such that $\exists h' : h'e \in \underline{H}_{k+1}$ and $h' \in \underline{H}_k$. As points in $H_m$ are unique for all $m$ by Def. 8, $h'$ must be $h$. Thus $he \in \underline{H}_{k+1}$.

This concludes the proof of Proposition 4.[21]  $\square$

**Lemma 1.** *If there exists an image-finite and concluding pointed DEL dynamical system that generates $(\mathcal{H}, \underline{H})$, then $(\mathcal{H}, \underline{H})$ is image-finite and concluding.*

*Proof* As the DEL dynamical system is image-finite, the construction of the generated ETL model (see Definition 7) ensures that all $H_k$ are image-finite. Hence $(\mathcal{H}, \underline{H})$ is image-finite.

If the DEL dynamical system terminates, then, by construction, the ETL model is finite and thus it concludes. If the DEL dynamical system is periodic, then the ETL model is, by construction, repeating and thus it concludes.  $\square$

**Lemma 2.** *If a saturated ETL model $(\mathcal{H}, \underline{H})$ is image-finite, concluding and satisfies Connected Time-Steps, then $(\mathcal{H}, \underline{H})$ is Component Collection Describable.*

*Proof* Let $(\mathcal{H}, \underline{H})$ be an image-finite and concluding saturated ETL model. Set $\mathcal{B} := \{C(\mathcal{H}h) : h \in \underline{H}\}$, the set of all connected components in $(\mathcal{H}, \underline{H})$. As all $C(\mathcal{H}h) \in \mathcal{B}$ are image-finite, by the Hennessy-Milner Theorem (see Sec. 6), for

---

[21] When constructing an action model that produces a pointed Kripke model isomorphic to a particular level in the to-be-generated ETL model, Proposition 3.2 of [20] (which states, roughly, that for almost any two pointed Kripke models, there exists an action model with postconditions that will produce one from the other using product update) is not applicable. The proposition is not applicable as the transforming action model allows only the designated point to survive, from which the desired Kripke model is unfolded. The resulting generated ETL models would therefore (most often) not be ETL isomorphic to the original ETL model, as ETL isomorphisms require that the temporal structure is preserved.

each pair $\underline{h}_1, \underline{h}_2 \in \underline{H}$ if $\underline{h}_1 \not\leftrightarrow \underline{h}_2$, there exists a formula $\varphi_{\underline{h}_1, \underline{h}_2}$ distinguishing between $\underline{h}_1$ and $\underline{h}_2$: $\underline{h}_1 \models \varphi_{\underline{h}_1, \underline{h}_2}$ while $\underline{h}_2 \not\models \varphi_{\underline{h}_1, \underline{h}_2}$.

Let $[C(\mathcal{H}\underline{h})\underline{h}]_{\leftrightarrow}$ be the equivalence class $\{C(\mathcal{H}\underline{h}')\underline{h}' \colon C(\mathcal{H}\underline{h}') \in \mathcal{B}$ and $\underline{h}' \leftrightarrow \underline{h}\}$. Then, since $(\mathcal{H}, \underline{H})$ is concluding, the set $\mathcal{B}_{\leftrightarrow} := \{[C(\mathcal{H}\underline{h})\underline{h}]_{\leftrightarrow} \colon \underline{h} \in \underline{H}, C(\mathcal{H}\underline{h}) \in \mathcal{B}\}$ is finite. Therefore, the conjunction $\bigwedge_{\underline{h}_2 \in \{\underline{h} \in \underline{H} \colon \underline{h} \neq \underline{h}_1\}} \varphi_{\underline{h}_1, \underline{h}_2}$ is well-defined for any $\underline{h}_1 \in \underline{H}$. This conjunction distinguishes $\underline{h}_1$ from any point in $\underline{H}$ that is not bisimilar to $\underline{h}_1$. Denote this formula $\varphi_{\underline{h}_1}$.

Moreover, as $\mathcal{B}_{\leftrightarrow}$ is finite, all sets $A \subseteq \underline{H}$ for which $\underline{h} \in A$ and $\underline{h}' \leftrightarrow \underline{h}$ implies $\underline{h}' \in A$ are finite. Hence, for any such $A$, the disjunction $\bigvee_{\underline{h}_1 \in A} \varphi_{\underline{h}_1}$ is well-defined. This disjunction distinguishes the connected components in $A$ from those not in $A$.  $\square$

**Lemma 3.** Let $\{(X_k, f_k)\}_{k \in K}$ be minimal in generating $\mathcal{H}$ and let $(X_k, f_k)$ from $x_k$ generate the saturated ETL model $(\mathcal{H}_k, \underline{H}_k)$. Then $(\mathcal{H}_k, \underline{H}_k)$ is the component branch sub-model for some terminal component branch $\mathbf{b}$ of $\mathcal{H}$.

*Proof* As $\mathcal{H}$ is obtained from $\{(\mathcal{H}_k, \underline{H}_k)\}_{k \in K}$, each $\mathcal{H}_k$ is a sub-model of $\mathcal{H}$. Hence $(\mathcal{H}_k, \underline{H}_k)$ is the component branch sub-model for a component branch $\mathbf{b}$ of $\mathcal{H}$, with $\mathbf{b} = (C(\mathcal{H}h))_{h \in \underline{H}_k}$. That $\mathbf{b}$ is terminal in $\mathcal{H}$ follows from $\{(X_k, f_k)\}_{k \in K}$ being minimal in generating $\mathcal{H}$: If $(\mathcal{H}_k, \underline{H}_k)$ was not terminal, there would be some $(\mathcal{H}_j, \underline{H}_j)$ extending it, generated by some $(X_j, f_k)$. But then $(X_k, f_k)$ would be redundant in generating $\mathcal{H}$ and $\{(X_k, f_k)\}_{k \in K}$ thus not minimal.  $\square$

**Theorem 2.** Let an image-finite and concluding ETL model $\mathcal{H}$ be given. $\mathcal{H}$ is generatable up to ETL isomorphism by a family of image-finite, concluding pointed DEL dynamical systems, if, and only if, there exists a saturation of each terminal component branch $\mathbf{b}$ of $\mathcal{H}$ such that $(\mathcal{H}_{\mathbf{b}}, \underline{H})$ satisfies all eight properties of Def. 6.

*Proof  Left-to-right:* Suppose $\mathcal{H}$ is generatable by a family of image-finite, concluding DEL dynamical systems $\{(X_k, f_k)\}_{k \in K}$. Let $(\mathcal{H}_k, \underline{H}_k)$ be the saturated ETL model generated by $(X_k, f_k)$ (cf. Def. 7). By Lemma 3, $(\mathcal{H}_k, \underline{H}_k)$ is a component branch sub-model of $\mathcal{H}$ for some terminal component branch of $\mathcal{H}$. As $(\mathcal{H}_k, \underline{H}_k)$ was generated by an image-finite and concluding DEL dynamical system, the saturation $\underline{H}_k$ of $\mathcal{H}_k$ makes $(\mathcal{H}_k, \underline{H}_k)$ satisfy all 8 properties of Def. 6 by Proposition 2 and Lemma 2.

*Right-to-left:* Let $\mathbf{B}$ be the set of terminal component branches of $\mathcal{H}$ and assume that for each $\mathbf{b} \in \mathbf{B}$, there exists a saturation such that $(\mathcal{H}_{\mathbf{b}}, \underline{H}_{\mathbf{b}})$ satisfies all eight properties of Def. 6. As $\mathcal{H}$ is image-finite and concluding, also each $(\mathcal{H}_{\mathbf{b}}, \underline{H}_{\mathbf{b}})$ is image-finite and concluding. By the constructions in the proof of Proposition 4, it follows that each $(\mathcal{H}_{\mathbf{b}}, \underline{H}_{\mathbf{b}})$ is generatable up to ETL isomorphism by an image-finite and concluding DEL dynamical system.

Let $(X_{\mathbf{b}}, f_{\mathbf{b}})$, $x_{\mathbf{b}}$ be the DEL dynamical system and initial Kripke model that generates $(\mathcal{H}_{\mathbf{b}}, \underline{H}_{\mathbf{b}})$ up to isomorphism, as given by the construction in the proof of Proposition 4. Let $(\mathcal{H}'_{\mathbf{b}}, \underline{H}'_{\mathbf{b}})$ be the specific ETL model generated by $(X_{\mathbf{b}}, f_{\mathbf{b}})$ from $x_{\mathbf{b}}$, as given by Def. 7. It is shown that the union

structure $U_{\mathbf{B}} = (E_{\mathbf{B}}, H_{\mathbf{B}}, \text{R}_{\mathbf{B},a}, V_{\mathbf{B}})_{a \in \mathcal{A}}$ of $\{(\mathcal{H}'_{\mathbf{b}}, \underline{H}'_{\mathbf{b}})\}_{\mathbf{b} \in \mathbf{B}}$ is isomorphic to $\mathcal{H} = (E, H, \text{R}_a, V)_{a \in \mathcal{A}}$. To this end, the existence of a bijection $g : H \longrightarrow H_{\mathbf{B}}$ is shown. That $g$ is also the sought ETL isomorphism follows from the ETL isomorphism of $(\mathcal{H}_{\mathbf{b}}, \underline{H}_{\mathbf{b}})$ and $(\mathcal{H}'_{\mathbf{b}}, \underline{H}'_{\mathbf{b}})$, for all $\mathbf{b}$.

Let $g_{\mathbf{b}}$ be the isomorphism between $(\mathcal{H}_{\mathbf{b}}, \underline{H}_{\mathbf{b}})$ and $(\mathcal{H}'_{\mathbf{b}}, \underline{H}'_{\mathbf{b}})$, guaranteed to exist by Proposition 4 and specifically given by 1) the construction of a DEL dynamical system from a component branch of the proof of Proposition 4 and 2) the construction for generating an ETL model from a DEL dynamical system of Def. 7. Combining the state-history naming schemes used in these two constructions yield $g_{\mathbf{b}}$ given by $g_{\mathbf{b}}(h) = h$ for $len(h) = 1$ in $\mathcal{H}_{\mathbf{b}}$, and $g_{\mathbf{b}}(he) = g_{\mathbf{b}}(h)e'_{\sigma_e}$ for $len(h) = k$. Notice that the history names from $\mathcal{H}$ are carried over, either directly or as indices, to the histories of $U_{\mathbf{B}}$.

Define a mapping $g : H \longrightarrow H_{\mathbf{B}}$ by $g(h) = g_{\mathbf{b}}(h)$ for $h \in \mathcal{H}_{\mathbf{b}}$. This mapping is well-defined as either $i)$ for exactly one $\mathbf{b} \in \mathbf{B}$, $h \in \mathcal{H}_{\mathbf{b}}$ (in which case $g(h)$ is well-defined), or $ii)$ if $h \in \mathcal{H}_{\mathbf{b}}$ and $h \in \mathcal{H}_{\mathbf{b}'}$, then $g_{\mathbf{b}}(h) = g_{\mathbf{b}'}(h)$. The latter is ensured as history names are carried over. This and that each $g_{\mathbf{b}}$ is an isomorphism implies that $g$ is a bijection. This completes the proof. □

**Lemma 4.** The saturated ETL model properties Synchronicity, Perfect Recall and Postcondition Describable persist under ETL model union.

I.e.: Let $\{(\mathcal{H}_k, \underline{H}_k)\}_{k \in K}$ be a countable set of saturated ETL models. If all $(\mathcal{H}_k, \underline{H}_k)$ satisfy either of the mentioned properties, then the (unsaturated) union structure $U_K$ satisfies that property.

*Proof* Synchronicity, Perfect Recall and Postcondition Describable persist because they are defined on histories which occur uniquely in the ETL forest, which makes them local by nature. Hence these properties are evaluated locally within a branch to ensure that there will be no conflicts when taking the union of different ETL sub-models. □

**Property 5.** Local No Miracles, Precondition Describable, Point Bisimulation Invariance and Connected Time-Steps do not persist under union.

*Proof* Local No Miracles does not persist under union: Consider a family of two DEL dynamical systems (that each individually satisfy property Local No Miracles.) with multi-pointed action models $f$ and $g$ with equal initial Kripke model $x = \{h, h', h_1, h_2\}$ where $h\text{R}_i h'$ and $h\text{R}^* h_1$ (by default, as initial models are connected), but disjunct Kripke models at the next level: $f^1(x) = \{he, h'e'\}$ and $g^1(x) = \{h_1e, h_2e'\}$ where $h_1e\text{R}_i h_2e'$ while not $he\text{R}_i h'e'$. In this example, Local No Miracles fails because not $he\text{R}^* h_1e$.

Precondition Describable does not persist under union: Consider a family of two DEL dynamical systems (that each individually satisfy property Precondition Describable) with multi-pointed action models $f$ and $g$ with equal initial Kripke model $x = \{h, h'\}$ with $h\text{R}^* h'$ (by default as initial models are connected), but disjunct Kripke models at the next level: $f^1(x) = \{he\}$ and $g^1(x) = \{h'e\}$. Then it is possible that $h' \models \delta_e$ while $h \not\models \delta_e$, which breaks

property Precondition Describable. It is left as an open question whether a suitably weakened version of Precondition Describable exists.

That Point Bisimulation Invariance is not preserved under union follows as the property is stated based on a saturation, but the union structure is unpointed, and hence unsaturated. If the union structure would be pointed with the set of points chosen as the union of the sets of points from the united ETL models, then the resulting set of points need not be a saturation, as the united ETL models may overlap, but have distinct points. In that case, the union structure would be "oversaturated".

Connected Time-Steps does not persist under union as histories within the same time-step are no longer necessarily epistemically connected in a union of disconnected ETL sub-models. □

**Proposition 5.** If an ETL model $\mathcal{H}$ is generated by a family of pointed DEL dynamical systems (possibly neither image-finite nor concluding), then $\mathcal{H}$ satisfies Synchronicity, Perfect Recall, Postcondition Describable and Very Local No Miracles and Local Bisimulation Invariance.

*Proof* That $\mathcal{H}$ satisfies Synchronicity, Perfect Recall and Postcondition Describable follows from Proposition 2 on page 17 and Lemma 4.

That $\mathcal{H}$ satisfies Very Local No Miracles follows directly by the additional requirement compared to Local No Miracles that $he\mathtt{R}^*h_1e$.

That $\mathcal{H}$ satisfies Local Bisimulation Invariance follows as 1) any ETL model that satisfies Point Bisimulation Invariance also satisfies Local Bisimulation Invariance, 2) Local Bisimulation Invariance is a property local to a connected component, and 3) connected components remain untouched under union. □

# Paper IV

# Convergence, Continuity and Recurrence in Dynamic Epistemic Logic

Dominik Klein[1,2]([⊠]) and Rasmus K. Rendsvig[3,4]([⊠])

[1] Department of Philosophy, Bayreuth University, Bayreuth, Germany
dominik.klein@uni-bayreuth.de
[2] Department of Political Science, University of Bamberg, Bamberg, Germany
[3] Theoretical Philosophy, Lund University, Lund, Sweden
rendsvig@gmail.com
[4] Center for Information and Bubble Studies, University of Copenhagen, Copenhagen, Denmark

**Abstract.** The paper analyzes dynamic epistemic logic from a topological perspective. The main contribution consists of a framework in which dynamic epistemic logic satisfies the requirements for being a topological dynamical system thus interfacing discrete dynamic logics with continuous mappings of dynamical systems. The setting is based on a notion of logical convergence, demonstratively equivalent with convergence in Stone topology. Presented is a flexible, parametrized family of metrics inducing the latter, used as an analytical aid. We show maps induced by action model transformations continuous with respect to the Stone topology and present results on the recurrent behavior of said maps.

**Keywords:** Dynamic epistemic logic · Limit behavior · Convergence · Recurrence · Dynamical systems · Metric spaces · General topology · Modal logic

## 1  Introduction

Dynamic epistemic logic is a framework for modeling information dynamics. In it, systematic change of Kripke models are punctiliously investigated through model transformers mapping Kripke models to Kripke models. The iterated application of such a map may constitute a model of information dynamics, or be may be analyzed purely for its mathematical properties [6, 8, 10, 11, 13, 16, 18, 40–43].

Dynamical systems theory is a mathematical field studying the long-term behavior of spaces under the action of a continuous function. In case of discrete time, this amounts to investigating the space under the iterations of a continuous map. The field is rich in concepts, methodologies and results developed with the aim of understanding general dynamics.

The two fields find common ground in the iterated application of maps. With dynamic epistemic logic analyzing very specific map types, the hope is that general results from dynamical systems theory may shed light on properties

of the former. There is, however, a chasm between the two: Dynamical systems theory revolves around spaces imbued with metrical or topological structure with respect to which maps are continuous. No such structure is found in dynamic epistemic logic. This chasm has not gone unappreciated: In his 2011 *Logical Dynamics of Information and Interaction* [10], van Benthem writes

> **From discrete dynamic logics to continuous dynamical systems**
> "We conclude with what we see as a major challenge. Van Benthem [7,8] pointed out how update evolution suggests a long-term perspective that is like the evolutionary dynamics found in dynamical systems. [...] Interfacing current dynamic and temporal logics with the continuous realm is a major issue, also for logic in general."        [10, Sect. 4.8. Emph. is org. heading]

This paper takes on the challenge and attempts to bridge this chasm.

We proceed as follows. Section 2 presents what we consider natural spaces when working with modal logic, namely sets of pointed Kripke models *modulo* logical equivalence. These are referred to as *modal spaces*. A natural notion of "logical convergence" on modal spaces is provided. Section 3 seeks a topology on modal spaces for which topological convergence coincides with logical convergence. We consider a metric topology based on $n$-bisimulation and prove it insufficient, but show an adapted Stone topology satisfactory. Saddened by the loss of a useful aid, the metric inducing the $n$-bisimulation topology, a family of metrics is introduced that all induce the Stone topology, yet allow a variety of subtle modelling choices. Sets of pointed Kripke models are thus equipped with a structure of compact metric spaces. Section 4 considers maps on modal spaces based on multi-pointed action models using product update. Restrictions are imposed to ensure totality, and the resulting *clean maps* are shown continuous with respect to the Stone topology. With that, we present our main contribution: A modal space under the action of a clean map satisfies the standard requirements for being a topological dynamical system. Section 5 applies the now-suited terminology from dynamical systems theory, and present some initial results pertaining to the recurrent behavior of clean maps on modal spaces. Section 6 concludes the paper by pointing out a variety of future research venues. Throughout, we situate our work in the literature.

*Remark 1.* To make explicit what may be apparent, note that the primary concern is the *semantics* of dynamic epistemic logic, i.e., its models and model transformation. Syntactical considerations are briefly touched upon in Sect. 6.

*Remark 2.* The paper is not self-contained. For notions from modal logic that remain undefined here, refer to e.g. [14,27]. For topological notions, refer to e.g. [37]. For more on dynamic and epistemic logic than the bare minimum of standard notions and notations rehearsed, see e.g. [2–5,10,20,22,30,38,39]. Finally, a background document containing generalizations and omitted proofs is our [31].

## 2    Modal Spaces and Logical Convergence

Let there be given a countable set $\Phi$ of **atoms** and a finite set $I$ of **agents**. Where $p \in \Phi$ and $i \in I$, define the **language** $\mathcal{L}$ by

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i\varphi.$$

Modal logics may be formulated in $\mathcal{L}$. By a **logic** $\Lambda$ we refer only to extensions of the **minimal normal modal logic** $K$ over the language $\mathcal{L}$. With $\Lambda$ given by context, let $\boldsymbol{\varphi}$ be the set of formulas $\Lambda$-provably equivalent to $\varphi$. Denote the resulting partition $\{\boldsymbol{\varphi} : \varphi \in \mathcal{L}\}$ of $\mathcal{L}$ by $\boldsymbol{\mathcal{L}}_\Lambda$.[1] Call $\boldsymbol{\mathcal{L}}_\Lambda$'s elements $\Lambda$-**propositions**.

We use relational semantics to evaluate formulas. A **Kripke model** for $\mathcal{L}$ is a tuple $M = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket)$ where $\llbracket M \rrbracket$ is a countable, non-empty set of **states**, $R : I \longrightarrow \mathcal{P}(\llbracket M \rrbracket \times \llbracket M \rrbracket)$ assigns to each $i \in I$ an **accessibility relation** $R_i$, and $\llbracket \cdot \rrbracket : \Phi \longrightarrow \mathcal{P}(\llbracket M \rrbracket)$ is a **valuation**, assigning to each atom a set of states. With $s \in \llbracket M \rrbracket$, call $Ms = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket, s)$ a **pointed Kripke model**. The used semantics are standard, including the modal clause:

$$Ms \vDash \Box_i\varphi \text{ iff for all } t : sR_it \text{ implies } Mt \vDash \varphi.$$

Throughout, we work with pointed Kripke models. Working with modal logics, we find it natural to identify pointed Kripke models that are considered equivalent by the logic used. The domains of interest are thus the following type of quotient spaces:

**Definition 1.** *The* $\boldsymbol{\mathcal{L}}_\Lambda$ *modal space of a set of pointed Kripke models $X$ is the set $\boldsymbol{X} = \{\boldsymbol{x} : x \in X\}$ for $\boldsymbol{x} = \{y \in X : y \vDash \varphi \text{ iff } x \vDash \varphi \text{ for all } \boldsymbol{\varphi} \in \boldsymbol{\mathcal{L}}_\Lambda\}$.*

Working with an $\boldsymbol{\mathcal{L}}_\Lambda$ modal space portrays that we only are interested in differences between pointed Kripke models insofar as these are modally expressible and are considered differences by $\Lambda$.

In a modal space, how may we conceptualize that a sequence $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots$ converges to some point $\boldsymbol{x}$? Focusing on the concept from which we derive the notion of identity in modal spaces, namely $\Lambda$-propositions, we find it natural to think of $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots$ as converging to $\boldsymbol{x}$ just in case $\boldsymbol{x_n}$ moves towards satisfying all the same $\Lambda$-propositions as $\boldsymbol{x}$ as $n$ goes to infinity. We thus offer the following definition:

**Definition 2.** *A sequence of points $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots$ in an $\boldsymbol{\mathcal{L}}_\Lambda$ modal space $\boldsymbol{X}$ is said to **logically converge** to the point $\boldsymbol{x}$ in $\boldsymbol{X}$ iff for every $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{L}}_\Lambda$ for which $x \vDash \varphi$, there is an $N \in \mathbb{N}$ such that $x_n \vDash \varphi$ for all $n \geq N$.*

To avoid re-proving useful results concerning this notion of convergence, we next turn to seeking a topology for which logical convergence coincides with topological convergence. Recall that for a topology $\mathcal{T}$ on a set $X$, a sequence of points $x_1, x_2, \dots$ is said to **converge** to $x$ in the **topological space** $(X, \mathcal{T})$ iff for every open set $U \in \mathcal{T}$ containing $x$, there is an $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$.

---

[1] $\boldsymbol{\mathcal{L}}_\Lambda$ is isomorphic to the domain of the **Lindenbaum algebra** of $\Lambda$.

## 3    Topologies on Modal Spaces

One way of obtaining a **topology** on a space is to define a **metric** for said space. Several metrics have been suggested for sets of pointed Kripke models [1,17]. These metrics are only defined for finite pointed Kripke models, but incorporating ideas from the metrics of [36] on **shift spaces** and [26] on sets of **first-order logical theories** allows us to simultaneously generalize and simplify the $n$-**Bisimulation-based Distance** of [17] to the degree of applicability:

Let $\boldsymbol{X}$ be a modal space for which modal equivalence and **bisimilarity** coincide[2] and let $\leftrightarrow_n$ relate $x, y \in X$ iff $x$ and $y$ are $n$-**bisimilar**. Then proving

$$d_B(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } x \leftrightarrow_n y \text{ for all } n \\ \frac{1}{2^n} & \text{if } n \text{ is the least intenger such that } x \not\leftrightarrow_n y \end{cases}$$

a metric on $\boldsymbol{X}$ is trivial. We refer to $d_B$ as the $n$-**bisimulation metric**, and to the induced **metric topology** as the $n$-**bisimulation topology**, denoted $\mathcal{T}_B$. A basis of the topology $\mathcal{T}_B$ is given by the set of elements $B_{\boldsymbol{x}n} = \{\boldsymbol{y} \in \boldsymbol{X} : y \leftrightarrow_n x\}$.

Considering the intimate link between modal logic and bisimulation, we consider both $n$-bisimulation metric and topology highly natural.[3] Alas, logical convergence does not:

**Proposition 1.** *Logical convergence in arbitrary modal space $\boldsymbol{X}$ does not imply convergence in the topological space $(\boldsymbol{X}, \mathcal{T}_B)$.*

*Proof.* Let $\boldsymbol{X}$ be an $\mathcal{L}_\Lambda$ modal space with $\mathcal{L}$ based on the atoms $\Phi = \{p_k : k \in \mathbb{N}\}$. Let $x \in X$ satisfy $\square\bot$ and $p_k$ for all $k \in \mathbb{N}$. Let $\boldsymbol{x_1}, \boldsymbol{x_2}, ...$ be a sequence in $\boldsymbol{X}$ such that for all $k \in \mathbb{N}$, $x_k$ satisfies $\square\bot$, $p_m$ for all $m \leq k$, and $\neg p_l$ for all $l > k$. Then for all $\boldsymbol{\varphi} \in \mathcal{L}_\Lambda$ for which $x \vDash \varphi$, there is an $N$ such that $x_n \vDash \varphi$ for all $n \geq N$, hence the sequence $\boldsymbol{x_1}, \boldsymbol{x_2}, ...$ converges to $\boldsymbol{x}$. There does not, however, exist any $N'$ such that $\boldsymbol{x_{n'}} \in B_{\boldsymbol{x}0}$ for all $n' \geq N'$. Hence $\boldsymbol{x_1}, \boldsymbol{x_2}, ...$ does not converge to $\boldsymbol{x}$ in $\mathcal{T}_B$. □

Proposition 1 implies that the $n$-bisimulation topology may not straightforwardly be used to establish negative results concerning logical convergence. That it may be used for positive cases is a corollary to Propositions 2 and 6 below. On the upside, logical convergence coincides with convergence in the $n$-bisimulation topology – i.e. Proposition 1 fails – when $\mathcal{L}$ has finite atoms. This is a corollary to Proposition 5.

An alternative to a metric-based approach to topologies is to construct the set of all open sets directly. Comparing the definition of logical convergence with that of convergence in topological spaces is highly suggestive: Replacing every occurrence of the formula $\varphi$ with an open set $U$ while replacing satisfaction $\vDash$ with inclusion $\in$ transforms the former definition into the latter. Hence the

---

[2] That all models in $X$ are **image-finite** is a sufficient condition, cf. the Hennessy-Milner Theorem. See e.g. [14] or [27].

[3] Space does not allow for a discussion of the remaining metrics of [1,17], but see [31].

collection of sets $U_{\boldsymbol{\varphi}} = \{\boldsymbol{x} \in \boldsymbol{X} : x \vDash \varphi\}$, $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{L}}_\Lambda$, seems a reasonable candidate for a topology. Alas, this collection is not closed under arbitrary unions, as all formulas are finite. Hence it is not a topology. It does however constitute the basis for a topology, in fact the somewhat influential **Stone topology**, $\mathcal{T}_S$.

The Stone topology is traditionally defined on the collection of complete theories for some propositional, first-order or modal logic, but is straightforwardly applicable to modal spaces. Moreover, it satisfies our *desideratum*:

**Proposition 2.** *For any $\boldsymbol{\mathcal{L}}_\Lambda$ modal space $\boldsymbol{X}$, a sequence $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots$ logically converges to the point $\boldsymbol{x}$ if, and only if, it converges to $\boldsymbol{x}$ in $(\boldsymbol{X}, \mathcal{T}_S)$.*

*Proof.* Assume $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots$ logically converges to $\boldsymbol{x}$ in $\boldsymbol{X}$ and that $U$ containing $\boldsymbol{x}$ is open in $\mathcal{T}_S$. Then there is a basis element $U_{\boldsymbol{\varphi}} \subseteq U$ with $\boldsymbol{x} \in U_{\boldsymbol{\varphi}}$. So $x \vDash \varphi$. By assumption, there exists an $N$ such that $x_n \vDash \varphi$ for all $n \geq N$. Hence $\boldsymbol{x_n} \in U_{\boldsymbol{\varphi}} \subseteq U$ for all $n \geq N$.

Assume $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots$ converges to $\boldsymbol{x}$ in $(\boldsymbol{X}, \mathcal{T}_S)$ and let $x \vDash \varphi$. Then $\boldsymbol{x} \in U_{\boldsymbol{\varphi}}$, which is open. As the sequence converges, there exists an $N$ such that $\boldsymbol{x_n} \in U_{\boldsymbol{\varphi}}$ for all $n \geq N$. Hence $x_n \vDash \varphi$ for all $n \geq N$. $\qquad\square$

Apart from its attractive characteristic concerning convergence, working on the basis of a logic, the Stone topology imposes a natural structure. As is evident from its basis, every subset of $\boldsymbol{X}$ characterizable by a single $\Lambda$-proposition $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{L}}_\Lambda$ is clopen. If the logic $\Lambda$ is compact and $\boldsymbol{X}$ saturated (see footnote 7), also the converse is true: every clopen set is of the form $U_\varphi$ for some $\varphi$. We refer to [31] for proofs and a precise characterization result. In this case, a subset is open, but not closed, iff it is characterizable only by an infinitary disjunction of $\Lambda$-propositions, and a subset if closed, but not open, iff it is characterizable only by an infinitary conjunction of $\Lambda$-propositions. The Stone topology thus transparently reflects the properties of logic, language and topology. Moreover, it enjoys practical topological properties:

**Proposition 3.** *For any $\boldsymbol{\mathcal{L}}_\Lambda$ modal space $\boldsymbol{X}$, $(\boldsymbol{X}, \mathcal{T}_S)$ is **Hausdorff** and **totally disconnected**. If $\Lambda$ is (logically) **compact**[4] and $\boldsymbol{X}$ is **saturated**[5], then $(\boldsymbol{X}, \mathcal{T}_S)$ is also (topologically) **compact**.*

*Proof.* These properties are well-known for the Stone topology applied to complete theories. For the topology applied to modal spaces, we defer to [31]. $\qquad\square$

One may interject that, as having a metric may facilitate obtaining results, it may cause a loss of tools to move away from the $n$-bisimulation topology. The Stone topology, however, is **metrizable**. A family of metrics inducing it, generalizing the Hamming distance to infinite strings by using weighted sums, was introduced in [31]. We here present a sub-family, suited for modal spaces:

---

[4] A logic $\Lambda$ is logically compact if any arbitrary set $A$ of formulas is $\Lambda$-consistent iff every finite subset of $A$ is $\Lambda$-consistent.

[5] An $\boldsymbol{\mathcal{L}}_\Lambda$ modal space $\boldsymbol{X}$ is saturated iff for each $\Lambda$-consistent set of formulas $A$, there is an $\boldsymbol{x} \in \boldsymbol{X}$ such that $x \vDash A$. Saturation relates to the notion of **strong completeness**, cf. e.g. [14, Proposition 4.12]. See [31] for its use in a more general context.

**Definition 3.** *Let $D \subseteq \boldsymbol{\mathcal{L}}_\Lambda$ contain for every $\psi \in \boldsymbol{\mathcal{L}}_\Lambda$ some $\{\varphi_i\}_{i \in I}$ that $\Lambda$-entails either $\psi$ or $\neg\psi$, and let $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, ...$ be an enumeration of $D$.*

*Let $\boldsymbol{X}$ be an $\boldsymbol{\mathcal{L}}_\Lambda$ modal space. For all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$, for all $k \in \mathbb{N}$, let*

$$d_k(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } x \vDash \varphi \text{ iff } y \vDash \varphi \text{ for } \varphi \in \boldsymbol{\varphi}_k \\ 1 & \text{else} \end{cases}$$

*Let $w : D \longrightarrow \mathbb{R}_{>0}$ assign strictly positive **weight** to each $\boldsymbol{\varphi}_k$ in $D$ such that $(w(\boldsymbol{\varphi}_n))$ forms a convergent series. Define the function $d_w : \boldsymbol{X}^2 \longrightarrow \mathbb{R}$ by*

$$d_w(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=0}^{\infty} w(\boldsymbol{\varphi}_k) d_k(\boldsymbol{x}, \boldsymbol{y})$$

*for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. The set of these functions is denoted $D_{\boldsymbol{X}}$. Let $\mathcal{D}_{D,\boldsymbol{X}} = \cup_{D \subseteq \boldsymbol{\mathcal{L}}_\Lambda} D_{\boldsymbol{X}}$.*

We refer to [31] for the proof establishing the following proposition:

**Proposition 4.** *Let $\boldsymbol{X}$ be an $\boldsymbol{\mathcal{L}}_\Lambda$ modal space and $d_w$ belong to $\mathcal{D}_{\boldsymbol{X}}$. Then $d_w$ is a metric on $\boldsymbol{X}$ and the metric topology $\mathcal{T}_w$ induced by $d_w$ on $\boldsymbol{X}$ is the Stone topology of $\Lambda$.*

For a metric space $(\boldsymbol{X}, d)$, we will also write $\boldsymbol{X}_d$.

With variable parameters $D$ and $w$, $\mathcal{D}_{\boldsymbol{X}}$ allows one to vary the choice of metric with the problem under consideration. E.g., if the $n$-bisimulation metric seems apt, one could choose that, with one restriction:

**Proposition 5.** *If $\boldsymbol{X}$ is an $\boldsymbol{\mathcal{L}}_\Lambda$ modal space with $\mathcal{L}$ based on a finite atom set, then $\mathcal{D}_{\boldsymbol{X}}$ contains a topological equivalent to the $n$-bisimulation metric.*

*Proof (sketch).* As $\mathcal{L}$ is based on a finite set of atoms, for each $\boldsymbol{x} \in \boldsymbol{X}, n \in \mathbb{N}_0$, there exists a characteristic formula $\varphi_{x,n}$ such that $y \vDash \varphi_{x,n}$ iff $y \leftrightarrows_n x$, cf. [27]. Let $D_n = \{\boldsymbol{\varphi}_{\boldsymbol{x},\boldsymbol{n}} : \boldsymbol{x} \in \boldsymbol{X}\}$ and $D = \cup_{n \in \mathbb{N}_0} D_n$. Then each $D_n$ is finite and $D$ satisfies Definition 3. Finally, let $w(\boldsymbol{\varphi}) = \frac{1}{|D_n|} \cdot \frac{1}{2^{n+1}}$ for $\boldsymbol{\varphi} \in D_n$. Then $d_w \in \mathcal{D}_{\boldsymbol{X}}$ and is equivalent to the $n$-bisimulation metric $d_b$. $\square$

As a corollary to Proposition 5, it follows that, for finite atom languages, the $n$-bisimulation topology is the Stone topology. This is not true in general, as witnessed by Proposition 1 and the following:

**Proposition 6.** *If $\boldsymbol{X}$ is an $\boldsymbol{\mathcal{L}}_\Lambda$ modal space with $\mathcal{L}$ based on a countably infinite atom set, then the $n$-bisimulation metric topology on $\boldsymbol{X}$ is strictly finer than the Stone topology on $\boldsymbol{X}$.*

*Proof (sketch).* We refer to [31] for details, but for $\mathcal{T}_B \nsubseteq \mathcal{T}_S$, note that the set $B_{\boldsymbol{x}0}$ used in the proof of Proposition 1, is open in $\mathcal{T}_B$, but not in $\mathcal{T}_S$. $\square$

With this comparison, we end our exposition of topologies on modal spaces.

## 4   Clean Maps on Modal Spaces

We focus on a class of maps induced by action models applied using product update. Action models are a popular and widely applicable class of model transformers, generalizing important constructions such as public announcements. An especially general version of action models is *multi-pointed* action models with *postconditions*. Postconditions allow action states in an action model to change the valuation of atoms [12,19], thereby also allowing the representation of information dynamics concerning situations that are not factually static. Permitting multiple points allows the actual action states executed to depend on the pointed Kripke model to be transformed, thus generalizing single-pointed action models.[6]

A **multi-pointed action model** is a tuple $\Sigma_\Gamma = (\llbracket \Sigma \rrbracket, \mathsf{R}, pre, post, \Gamma)$ where $\llbracket \Sigma \rrbracket$ is a countable, non-empty set of **actions**. The map $\mathsf{R} : I \to \mathcal{P}(\llbracket \Sigma \rrbracket \times \llbracket \Sigma \rrbracket)$ assigns an **accessibility relation** $\mathsf{R}_i$ on $\llbracket \Sigma \rrbracket$ to each agent $i \in I$. The map $pre : \llbracket \Sigma \rrbracket \to \mathcal{L}$ assigns to each action a **precondition**, and the map $post : \llbracket \Sigma \rrbracket \to \mathcal{L}$ assigns to each action a **postcondition**,[7] which must be $\top$ or a conjunctive clause[8] over $\Phi$. Finally, $\emptyset \neq \Gamma \subseteq \llbracket \Sigma \rrbracket$ is the set of designated actions.

To obtain well-behaved total maps on a modal spaces, we must invoke a set of mild, but non-standard, requirements: Let $X$ be a set of pointed Kripke models. Call $\Sigma_\Gamma$ **precondition finite** if the set $\{pre(\sigma) \in \mathcal{L}_\Lambda : \sigma \in \llbracket \Sigma \rrbracket\}$ is finite. This is needed for our proof of continuity. Call $\Sigma_\Gamma$ **exhaustive over** $X$ if for all $x \in X$, there is a $\sigma \in \Gamma$ such that $x \vDash pre(\sigma)$. This conditions ensures that the action model $\Sigma_\Gamma$ is universally applicable on $X$. Finally, call $\Sigma_\Gamma$ **deterministic over** $X$ if $X \vDash pre(\sigma) \wedge pre(\sigma') \to \bot$ for each $\sigma \neq \sigma' \in \Gamma$. Together with exhaustivity, this condition ensures that the product of $\Sigma_\Gamma$ and any $Ms \in X$ is a (single-)pointed Kripke model, i.e., that the actual state after the updates is well-defined and unique.

Let $\Sigma_\Gamma$ be exhaustive and deterministic over $X$ and let $Ms \in X$. Then the **product update** of $Ms$ with $\Sigma_\Gamma$, denoted $Ms \otimes \Sigma_\Gamma$, is the pointed Kripke model $(\llbracket M\Sigma \rrbracket, R', \llbracket \cdot \rrbracket', s')$ with

$$\llbracket M\Sigma \rrbracket = \{(s,\sigma) \in \llbracket M \rrbracket \times \llbracket \Sigma \rrbracket : (M,s) \vDash pre(\sigma)\}$$
$$R' = \{((s,\sigma),(t,\tau)) : (s,t) \in R_i \text{ and } (\sigma,\tau) \in \mathsf{R}_i\}, \text{ for all } i \in N$$
$$\llbracket p \rrbracket' = \{(s,\sigma) : s \in \llbracket p \rrbracket, post(\sigma) \nvDash \neg p\} \cup \{(s,\sigma) : post(\sigma) \vDash p\}, \text{ for all } p \in \Phi$$
$$s' = (s,\sigma) : \sigma \in \Gamma \text{ and } Ms \vDash pre(\sigma)$$

Call $\Sigma_\Gamma$ **closing over** $X$ if for all $x \in X$, $x \otimes \Sigma_\Gamma \in X$. With $\Sigma_\Gamma$ exhaustive and deterministic, $\Sigma_\Gamma$ and $\otimes$ induce a well-defined total map on $X$.

The class of maps of interest in the present is then the following:

---

[6] Multi-pointed action models are also referred to as *epistemic programs* in [2], and allow encodings akin to *knowledge-based programs* [22] of interpreted systems, cf. [42].

[7] The precondition of $\sigma$ specify the conditions under which $\sigma$ is executable, while its postcondition may dictate the posterior values of a finite, possibly empty, set of atoms.

[8] I.e. a conjuction of literals, where a literal is an atom or a negated atom.

**Definition 4.** *Let $X$ be an $\mathcal{L}_\Lambda$ modal space. A map $f : X \to X$ is called **clean** if there exists a precondition finite, multi-pointed action model $\Sigma\Gamma$ closing, deterministic and exhaustive over $X$ such that $f(x) = y$ iff $x \otimes \Sigma\Gamma \in y$ for all $x \in X$.*

Clean maps are total by the assumptions of being closing and exhaustive. They are well-defined as $f(x)$ is independent of the choice of representative for $x$: If $x' \in x$, then $x' \otimes \Sigma\Gamma$ and $x \otimes \Sigma\Gamma$ are modally equivalent and hence define the same point in $X$. The latter follows as multi-pointed action models applied using product update preserve bisimulation [2], which implies modal equivalence. Clean maps moreover play nicely with the Stone topology:

**Proposition 7.** *Let $f$ be a clean map on an $\mathcal{L}_\Lambda$ modal space $X$. Then $f$ is continuous with respect to the Stone topology of $\Lambda$.*

*Proof (sketch).* We defer to [31] for details, but offer a sketch: The map $f$ is shown uniformly continuous using the $\varepsilon$-$\delta$ formulation of continuity. The proof relies on a lemma stating that for every $d_w \in \mathcal{D}_X$ and every $\epsilon > 0$, there are formulas $\chi_1, \ldots, \chi_l \in \mathcal{L}$ such that every $x \in X$ satisfies some $\chi_i$ and whenever $y \vDash \chi_i$ and $z \vDash \chi_i$ for some $i \leq l$, then $d_w(y, z) < \epsilon$. The main part of the proof establishes the claim that there is a function $\delta : \mathcal{L} \to (0, \infty)$ such that for any $\varphi \in \mathcal{L}$, if $f(x) \vDash \varphi$ and $d_a(x, y) < \delta(\varphi)$, then $f(y) \vDash \varphi$. Setting $\delta = \min\{\delta(\chi_i) \colon i \leq l\}$ then yields a $\delta$ with the desired property.     □

With Proposition 7, we are positioned to state our main theorem:

**Theorem 1.** *Let $f$ be a clean map on a saturated $\mathcal{L}_\Lambda$ modal space $X$ with $\Lambda$ compact and let $d \in \mathcal{D}_X$. Then $(X_d, f)$ is a **topological dynamical system**.*

*Proof.* Propositions 2, 3, 4 and 7 jointly imply that $X_d$ is a compact metric space on which $f$ is continuous, thus satisfying the requirements of e.g. [21, 29, 44].     □

With Theorem 1, we have, in what we consider a natural manner, situated dynamic epistemic logic in the mathematical discipline of dynamical systems. A core topic in this discipline is to understand the long-term, qualitative behavior of maps on spaces. Central to this endeavor is the concept of *recurrence*, i.e., understanding when a system returns to previous states as time goes to infinity.

## 5   Recurrence in the Limit Behavior of Clean Maps

We represent results concerning the limit behavior of clean maps on modal spaces. In establishing the required terminology, we follow [29]: Let $f$ be a continuous map on a metric space $X_d$ and $x \in X_d$. A point $y \in X$ is a **limit point**[9] for $x$ under $f$ if there is a strictly increasing sequence $n_1, n_2, \ldots$ such that the subsequence $f^{n_1}(x), f^{n_2}(x), \ldots$ of $(f^n(x))_{n \in \mathbb{N}_0}$ converges to $y$. The **limit set** of $x$ under $f$ is the set of all limit points for $x$, denoted $\omega_f(x)$. Notably, $\omega_f(x)$ is closed under $f$: For $y \in \omega_f(x)$ also $f(y) \in \omega_f(x)$. We immediately obtain that any modal system satisfying Theorem 1 has a nonempty limit set:

---

[9] Or $\omega$-limit point. The $\omega$ is everywhere omitted as time here only moves forward.

**Proposition 8.** *Let $(\boldsymbol{X}_d, \boldsymbol{f})$ be as in Theorem 1. For any point $\boldsymbol{x} \in \boldsymbol{X}$, the limit set of $\boldsymbol{x}$ under $\boldsymbol{f}$ is non-empty.*

*Proof.* Since $\boldsymbol{X}$ is is compact, every sequence in $\boldsymbol{X}$ has a convergent subsequence, cf. e.g. [37, Theorem 28.2].

Proposition 8 does not inform us of the *structure* of said limit set. In the study of dynamical systems, such structure is often sought through classifying the possible repetitive behavior of a system, i.e., through the system's *recurrence* properties. For such studies, a point $x$ is called (positively) **recurrent** if $x \in \omega_f(x)$, i.e., if it is a limit point of itself.

The simplest structural form of recurrence is *periodicity*: For a point $x \in X$, call the set $\mathcal{O}_f(x) = \{f^n(x) \colon n \in \mathbb{N}_0\}$ its **orbit**. The orbit $\mathcal{O}_f(x)$ is **periodic** if $f^{n+k}(x) = f^n(x)$ for some $n \geq 0, k > 0$; the least such $k$ is the **period** of $\mathcal{O}_f(x)$. Periodicity is thus equivalent to $\mathcal{O}_f(x)$ being finite. Related is the notion of a **limit cycle**: a periodic orbit $\mathcal{O}_f(x)$ is a limit cycle if it is the limit set of some $y$ not in the period, i.e., if $\mathcal{O}_f(x) = \omega_f(y)$ for some $y \notin \mathcal{O}_f(x)$.

It was conjectured by van Benthem that certain clean maps—those based on finite action models and without postconditions—would, whenever applied to a finite $x$, have a periodic orbit $\mathcal{O}_f(x)$. I.e., after finite iterations, the map would oscillate between a finite number of states. This was the content of van Benthem's "*Finite Evolution Conjecture*" [8]. The conjecture was refuted using a counterexample by Sadzik in his 2006 paper, [43].[10] The example provided by Sadzik (his Example 33) uses an action model with only Boolean preconditions. Interestingly, the orbit of the corresponding clean map terminates in a limit cycle. This is a corollary to Proposition 9 below.

Before we can state the proposition, we need to introduce some terminology. Call a multi-pointed action model $\Sigma\Gamma$ **finite** if $[\![\Sigma]\!]$ is finite, **Boolean** if $pre(\sigma)$ is a Boolean formula for all $\sigma \in [\![\Sigma]\!]$, and **static** if $post(\sigma) = \top$ for all $\sigma \in [\![\Sigma]\!]$. We apply the same terms to a clean map $\boldsymbol{f}$ based on $\Sigma\Gamma$. In this terminology, Sadzik showed that for any finite, Boolean, and static clean map $\boldsymbol{f} \colon \boldsymbol{X} \to \boldsymbol{X}$, if the orbit $\mathcal{O}_{\boldsymbol{f}}(\boldsymbol{x})$ is periodic, then it has period 1.[11] This insightful result immediates the following:

**Proposition 9.** *Let $(\boldsymbol{X}_d, \boldsymbol{f})$ be as in Theorem 1 with $\boldsymbol{f}$ finite, Boolean, and static. For all $\boldsymbol{x} \in \boldsymbol{X}$, the orbit $\mathcal{O}_{\boldsymbol{f}}(\boldsymbol{x})$ is periodic with period 1.*

*Proof.* By Proposition 8, the limit set $\omega_{\boldsymbol{f}}(\boldsymbol{x})$ of $\boldsymbol{x}$ under $\boldsymbol{f}$ is non-empty. Sadzik's result shows that it contains a single point. Hence $(\boldsymbol{f}^n(\boldsymbol{x}))_{n \in \mathbb{N}_0}$ converges to this point. As the limit set $\omega_{\boldsymbol{f}}(\boldsymbol{x})$ is closed under $\boldsymbol{f}$, its unique point is a fix-point. □

Proposition 9 may be seen as a partial vindication of van Benthem's conjecture: Forgoing the requirement of reaching the limit set in finite time and the possibility of modal preconditions, the conjecture holds, even if the initial state has an infinite set of worlds $[\![x]\!]$. This simple recurrent behavior is, however, not

---

[10] We paraphrase van Benthem and Sadzik using the terminology introduced.

[11] See [16] for an elegant and generalizing exposition.

the general case. More complex clean maps may exhibit **nontrivial recurrence**, i.e., produce non-periodic orbits with recurrent points:

**Proposition 10.** *There exist finite, static, but non-Boolean, clean maps that exhibit nontrivial recurrence.*

**Proposition 11.** *There exist finite, Boolean, but non-static, clean maps that exhibit nontrivial recurrence.*

We show these propositions below, building a clean map which, from a selected initial state, has uncountably many limit points, despite the orbit being only countable. Moreover, said orbit also contains infinitely many recurrent points. In fact, every element of the orbit is recurrent. We rely on Lemma 1 in the proof. A proof of Lemma 1.1 may be found in [32], a proof of Lemma 1.2 in [15].

**Lemma 1.** *Any Turing machine can be emulated using a set $X$ of $S5$ pointed Kripke models for finite atoms and a finite multi-pointed action model $\Sigma\Gamma$ deterministic over $X$. Moreover, $\Sigma\Gamma$ may be chosen 1. static, but non-Boolean, or 2. Boolean, but non-static.*

*Proof (of Propositions 10 and 11).* For both propositions, we use a Turing machine *ad infinitum* iterating the successor function on the natural numbers. Numbers are represented in mirrored base-2, i.e., with the *leftmost* digit the *lowest*. Such a machine may be build with alphabet $\{\triangleright, 0, 1, \sqcup\}$, where the symbol $\triangleright$ is used to mark the starting cell and $\sqcup$ is the *blank* symbol. We omit the exact description of the machine here. Of importance is the content of the tape: Omitting blank ($\sqcup$) cells, natural numbers are represented as illustrated in Fig. 1.

$$
\begin{array}{llllll}
0: \boxed{\triangleright}\boxed{0} & 2: \boxed{\triangleright}\boxed{0}\boxed{1} & 4: \boxed{\triangleright}\boxed{0}\boxed{0}\boxed{1} & 6: \boxed{\triangleright}\boxed{0}\boxed{1}\boxed{1} & 8: \boxed{\triangleright}\boxed{0}\boxed{0}\boxed{0}\boxed{1} \\
1: \boxed{\triangleright}\boxed{1} & 3: \boxed{\triangleright}\boxed{1}\boxed{1} & 5: \boxed{\triangleright}\boxed{1}\boxed{0}\boxed{1} & 7: \boxed{\triangleright}\boxed{1}\boxed{1}\boxed{1} & 9: \boxed{\triangleright}\boxed{1}\boxed{0}\boxed{0}\boxed{1}
\end{array}
$$
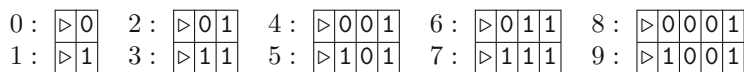
**Fig. 1.** Mirrored base-2 Turing tape representation of $0, .., 9 \in \mathbb{N}_0$, blank cells omitted. Notice that the mirrored notation causes perpetual change close to the start cell, $\triangleright$.

Initiated with its read-write head on the cell with the start symbol $\triangleright$ of a tape with content $n$, the machine will go through a number of configurations before returning the read-write head to the start cell with the tape now having content $n+1$. Auto-iterating, the machine will thus, over time, produce a tape that will have contained every natural number in order.

This Turing machine may be emulated by a finite $\Sigma\Gamma$ on a set $X$ cf. Lemma 1. Omitting the details[12], the idea is that the Turing tape, or a finite fragment,

---

[12] The details differ depending on whether $\Sigma\Gamma$ must be static, but non-Boolean for Proposition 10, or Boolean, but non-static for Proposition 11. See resp. [15,32].
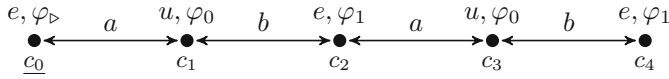
**Fig. 2.** A pointed Kripke model emulating the configuration of the Turing machine with cell content representing the number 10. The designated state is the underlined $c_0$. Each state is labeled with a formula $\varphi_\triangleright$, $\varphi_0$ or $\varphi_1$ expressing its content. Relations $a$ and $b$ allow expressing distance of cells: That $c_0$ satisfies $\Diamond_a(u \wedge \Diamond_b(e \wedge \varphi_1))$ exactly expresses that cell $c_2$ contains a 1. Omitted are reflexive loops for relations, and the additional structure marking cell content and read-write head position.

thereof may be encoded as a pointed Kripke model: Each cell of the tape corresponds to a state, with the cell's content encoded by additional structure,[13] which is modally expressible. By structuring the cell states with two equivalence relations and atoms $u$ and $e$ true at cells with odd (even) index respectively, (cf. Fig. 2), also the position of a cell is expressible. The designated state corresponds to the start cell, marked $\triangleright$.

Let $(c_n)_{n \in \mathbb{N}_0}$ be the sequence of configurations of the machine when initiated on a tape with content 0. Each $c_n$ may be represented by a pointed Kripke model, obtaining a sequence $(x_n)_{n \in \mathbb{N}_0}$. By Lemma 1, there thus exists a $\Sigma\Gamma$ such that for all $n$, $x_n \otimes \Sigma\Gamma = x_{n+1}$. Hence, moving to the full modal space $\boldsymbol{X}$ for the language used, a clean map $\boldsymbol{f} \colon \boldsymbol{X} \to \boldsymbol{X}$ based on $\Sigma\Gamma$ will satisfy $\boldsymbol{f}(\boldsymbol{x}_n) = \boldsymbol{x}_{n+1}$ for all $n$. The Turing machine's run is thus emulated by $(\boldsymbol{f}^k(\boldsymbol{x}_0))_{k \in \mathbb{N}_0}$.

Let $(c'_n)_{n \in \mathbb{N}_0}$ be the subsequence of $(c_n)_{n \in \mathbb{N}_0}$ where the machine has finished the successor operation and returned its read write head to its starting position $\triangleright$, ready to embark on the next successor step. The tape of the first 9 of these $c'_n$ are depicted in Fig. 1. Let $(\boldsymbol{x}'_n)_{n \in \mathbb{N}_0}$ be the corresponding subsequence of $(\boldsymbol{f}^k(\boldsymbol{x}_0))_{k \in \mathbb{N}_0}$. We show that $(\boldsymbol{x}'_n)_{n \in \mathbb{N}_0}$ has uncountably many limit points:

For each subset $Z$ of $\mathbb{N}$, let $c^Z$ be a tape with content 1 on cell $i$ iff $i \in Z$ and 0 else. On the Kripke model side, let the corresponding $x^Z \in X$ be a model structurally identical to those of $(x'_n)_{n \in \mathbb{N}_0}$, but satisfying $\varphi_1$ on all "cell states" distance $i \in Z$ from the designated "$\triangleright$" state, and $\varphi_0$ on all other.[14] The set $\{x^Z : Z \subseteq \mathbb{N}\}$ is uncountable, and each $\boldsymbol{x^Z}$ is a limit point of $\overline{\boldsymbol{x}}$: For each $Z \subseteq \mathbb{N}$ and $n \in \mathbb{N}$, there are infinitely many $k$ for which $x_k \vDash \varphi$ iff $x^Z \vDash \varphi$ for all $\varphi$ of modal depth at most $n$. Hence, for every $n$, the set $\{\boldsymbol{x}_k : d_b(\boldsymbol{x}_k, \boldsymbol{x^Z}) < 2^{-n}\}$ is infinite, with $d_b$ the equivalent of the $n$-bisimulation metric, cf. Proposition 5. Hence, for each of the uncountably many $Z \subseteq \mathbb{N}$, $\boldsymbol{x^Z}$ is a limit point of the sequence $\overline{\boldsymbol{x}}$.

Finally, every $\boldsymbol{x}'_k \in (\boldsymbol{x}'_n)_{n \in \mathbb{N}_0}$ is recurrent: That $\boldsymbol{x}'_k \in \omega_{\boldsymbol{f}}(\boldsymbol{x}'_k)$ follows from $\boldsymbol{x}'_k$ being a limit point of $(\boldsymbol{x}'_n)_{n \in \mathbb{N}_0}$, which it is as $\boldsymbol{x}'_k = \boldsymbol{x^Z}$ for some $Z \subseteq \mathbb{N}$.[15] As the set of recurrent points is thus infinite, it cannot be periodic.  $\square$

---

[13] For Proposition 11, tape cell content may be encoded using atomic propositions, changeable through postconditions, cf. [15]; for Proposition 10, cell content is written by adding and removing additional states, cf. [32].

[14] The exact form is straightforward from the constructions used in [15,32].

[15] A similar argument shows that all $x^Z$ with $Z \subseteq \mathbb{N}$ co-infinite are recurrent points. Hence $\omega_f(\boldsymbol{x}'_k)$ for any $\boldsymbol{x}'_k \in (\boldsymbol{x}'_n)_{n \in \mathbb{N}_0}$ contains uncountably many recurrent points.

As a final result on the orbits of clean maps, we answer an open question: After having exemplified a period 2 system, Sadzik [43] notes that it is unknown whether finite, static, but non-Boolean, clean maps exhibiting longer periods exist. They do:

**Proposition 12.** *For any $n \in \mathbb{N}$, there exists finite, static, but non-Boolean clean maps with periodic orbits of period $n$. This is also true for finite Boolean, but non-static, clean maps.*

*Proof.* For the given $n$, find a Turing machine that, from some configuration, loops with period $n$. From here, Lemma 1 does the job.  □

Finally, we note that brute force determination of a clean map's orbit properties is not in general a feasible option:

**Proposition 13.** *The problems of determining whether a Boolean and non-static, or a static and non-Boolean, clean map, a) has a periodic orbit or not, and b) contains a limit cycle or not, are both undecidable.*

*Proof.* The constructions from the proofs of Lemma 1 allows encoding the halting problem into either question.  □

# 6 Discussion and Future Venues

We consider Theorem 1 our main contribution. With it, an interface between the discrete semantics of dynamic epistemic logic with dynamical systems have been provided; thus the former has been situated in the mathematical field of the latter. This paves the way for the application of results from dynamical systems theory and related fields to the information dynamics of dynamic epistemic logic.

The term *nontrivial recurrence* is adopted from Hasselblatt and Katok, [29]. They remark that "[nontrivial recurrence] is the first indication of complicated asymptotic behavior." Propositions 10 and 11 indicate that the dynamics of action models and product update may not be an easy landscape to map. Hasselblatt and Katok continue: "In certain low-dimensional situations [...] it is possible to give a comprehensive description of the nontrivial recurrence that can appear." [29, p. 24]. That the Stone topology is zero-dimensional fuels the hope that general topology and dynamical systems theory yet has perspectives to offer on dynamic epistemic logic. One possible direction is seeking a finer parametrization of clean maps combined with results specific to zero-dimensional spaces, as found, e.g., in the field of symbolic dynamics [36]. But also other venues are possible: The introduction of [29] is an inspiration.

The approach presented furthermore applies to model transformations beyond multi-pointed action models and product update. Given the equivalence shown in [33] between single-pointed action model product update and *general arrow updates*, we see no reason to suspect that "clean maps" based on the latter should not be continuous on modal spaces. A further conjecture is that the *action-priority update* of [5] on plausibility models[16] yields "clean maps"

---

[16] Hence also the multi-agent belief revision policies *lexicographic upgrade* and *elite change*, also known as *radical* and *conservative upgrade*, introduced in [9], cf. [5].

continuous w.r.t. the suited Stone topology, and that this may be shown using a variant of our proof of the continuity of clean maps. A more difficult case is the *PDL-transformations* of *General Dynamic Dynamic Logic* [25] given the signature change the operation involves.

There is a possible clinch between the suggested approach and epistemic logic with common knowledge. The state space of a dynamical system is compact. The Stone topology for languages including a common knowledge operator is non-compact. Hence, it cannot constitute the space of a dynamical system— but its *one-point compactification* may. We are currently working on this clinch, the consequences of compactification, and relations to the problem of attaining common knowledge, cf. [28].

Questions also arise concerning the *dynamic logic* of dynamic epistemic logic. Propositions 10 and 11 indicate that there is more to the semantic dynamics of dynamic epistemic logic than is representable by finite compositional dynamic modalities—even when including a Kleene star. An open question still stands on how to reason about limit behavior. One interesting venue stems from van Benthem [10]. He notes[17] that the reduction axioms of dynamic epistemic logic could possibly be viewed on par with differential equations of quantitative dynamical systems. As modal spaces are zero-dimensional, they are imbeddable in $\mathbb{R}$ cf. [37, Theorem 50.5], turning clean maps into functions from $\mathbb{R}$ to $\mathbb{R}$, possibly representable as discrete-time difference equations.

An alternative approach is possible given by consulting Theorem 1. With Theorem 1, a connection arises between dynamic epistemic logic and *dynamic topological logic* (see e.g. [23,24,34,35]): Each system $(\boldsymbol{X}_d, \boldsymbol{f})$ may be considered a dynamic topological model with atom set $\mathcal{L}_\Lambda$ and the 'next' operator's semantics given by an application of $\boldsymbol{f}$, equivalent to a $\langle \boldsymbol{f} \rangle$ dynamic modality of DEL. The topological 'interior' operator has yet no DEL parallel. A 'henceforth' operator allows for a limited characterization of recurrence [35]. We are wondering about and wandering around the connections between a limit set operator with semantics $x \vDash [\omega_{\boldsymbol{f}}]\varphi$ iff $y \vDash \varphi$ for all $\boldsymbol{y} \in \omega_{\boldsymbol{f}}(\boldsymbol{x})$, dynamic topological logic and the study of oscillations suggested by van Benthem [11].

With the focal point on pointed Kripke models and action model transformations, we have only considered a special case of logical dynamics. It is our firm belief that much of the methodology here suggested is generalizable: With structures described logically using a countable language, the notion of logical convergence will coincide with topological convergence in the Stone topology on the quotient space *modulo* logical equivalence, and the metrics introduced will, *mutatis mutandis*, be applicable to said space [31]. The continuity of maps and compactness of course depends on what the specifics of the chosen model transformations and the compactness of the logic amount to.

---

[17] In the omitted part of the quotation from the introduction.

# References

1. Aucher, G.: Generalizing AGM to a multi-agent setting. Logic J. IGPL **18**(4), 530–558 (2010)
2. Baltag, A., Moss, L.S.: Logics for epistemic programs. Synthese **139**(2), 165–224 (2004)
3. Baltag, A., Moss, L.S., Solecki, S.: The logic of public announcements, common knowledge, and private suspicions. In: TARK 1998. Morgan Kaufmann (1998)
4. Baltag, A., Renne, B.: Dynamic epistemic logic. In: The Stanford Encyclopedia of Philosophy (2016). Fall 2016th Edition
5. Baltag, A., Smets, S.: A qualitative theory of dynamic interactive belief revision. In: Proceedings of LOFT 7. Amsterdam University Press (2008)
6. Baltag, A., Smets, S.: Group belief dynamics under iterated revision: fixed points and cycles of joint upgrades. In: TARK 2009. ACM (2009)
7. van Benthem, J.: Games in dynamic-epistemic logic. Bull. Econ. Res. **53**(1), 219–249 (2001)
8. van Benthem, J.: "One is a Lonely Number": logic and communication. In: Logic Colloquium 2002. Lecture Notes in Logic, vol. 27. Association for Symbolic Logic (2006)
9. van Benthem, J.: Dynamic logic for belief revision. J. Appl. Non-Class. Logics **17**(2), 129–155 (2007)
10. van Benthem, J.: Logical Dynamics of Information and Interaction. Cambridge University Press, Cambridge (2011)
11. van Benthem, J.: Oscillations, logic, and dynamical systems. In: Ghosh, S., Szymanik, J. (eds.) The Facts Matter. College Publications, London (2016)
12. van Benthem, J., van Eijck, J., Kooi, B.: Logics of communication and change. Inf. Comput. **204**(11), 1620–1662 (2006)
13. van Benthem, J., Gerbrandy, J., Hoshi, T., Pacuit, E.: Merging frameworks for interaction. J. Philos. Logic **38**(5), 491–526 (2009)
14. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press, Cambridge (2001)
15. Bolander, T., Birkegaard, M.: Epistemic planning for single- and multi-agent systems. J. Appl. Non-Class. Logics **21**(1), 9–34 (2011)
16. Bolander, T., Jensen, M., Schwarzentruber, F.: Complexity results in epistemic planning. In: Proceedings of IJCAI 2015. AAAI Press (2015)
17. Caridroit, T., Konieczny, S., de Lima, T., Marquis, P.: On distances between KD45n Kripke models and their use for belief revision. In: ECAI 2016. IOS Press (2016)
18. Dégremont, C.: The temporal mind: observations on the logic of belief change in interactive systems. Ph.D. thesis, University of Amsterdam (2010)
19. van Ditmarsch, H., Kooi, B.: Semantic results for ontic and epistemic change. In: Logic and the Foundations of Game and Decision Theory (LOFT 7). Texts in Logic and Games, vol. 3. Amsterdam University Press (2008)
20. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Springer, Dordrecht (2008). doi:10.1007/978-1-4020-5839-4

21. Eisner, T., Farkas, B., Haase, M., Nagel, R.: Operator Theoretic Aspects of Ergodic Theory. Springer, Heidelberg (2015). doi:10.1007/978-3-319-16898-2
22. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning About Knowledge. The MIT Press, Cambridge (1995)
23. Fernández-Duque, D.: A sound and complete axiomatization for dynamic topological logic. J. Symb. Logic **77**(3), 1–26 (2012)
24. Fernández-Duque, D.: Dynamic topological logic of metric spaces. J. Symb. Logic **77**(1), 308–328 (2012)
25. Girard, P., Seligman, J., Liu, F.: General dynamic dynamic logic. In: Bolander, T., Brauner, T., Ghilardi, S., Moss, L. (eds.) Advances in modal logics, vol. 9. College Publications, London (2012)
26. Goranko, V.: Logical topologies and semantic completeness. In: Logic Colloquium 1999. Lecture Notes in Logic, vol. 17. AK Peters (2004)
27. Goranko, V., Otto, M.: Model theory of modal logic. In: Blackburn, P., van Benthem, J., Wolter, F. (eds.) Handbook of Modal Logic. Elsevier, Amsterdam (2007)
28. Halpern, Y.J., Moses, Y.: Knowledge and common knowledge in a distributed environment. J. ACM **37**(3), 549–587 (1990)
29. Hasselblatt, B., Katok, A.: Principal structures. In: Hasselblatt, B., Katok, A. (eds.) Handbook of Dynamical Systems, vol. 1A. Elsevier, Amsterdam (2002)
30. Hintikka, J.: Knowledge and Belief: An Introduction to the Logic of the Two Notions. College Publications, London (1962). 2nd, 2005 Edition
31. Klein, D., Rendsvig, R.K.: Metrics for Formal Structures, with an Application to Kripke Models and their Dynamics. arXiv:1704.00977 (2017)
32. Klein, D., Rendsvig, R.K.: Turing Completeness of Finite, Epistemic Programs. arXiv:1706.06845 (2017)
33. Kooi, B., Renne, B.: Generalized arrow update logic. In: TARK 2011. ACM, New York (2011)
34. Kremer, P., Mints, G.: Dynamical topological logic. Bull. Symb. Logic **3**, 371–372 (1997)
35. Kremer, P., Mints, G.: Dynamic Topological Logic. In: Aiello, M., Pratt-Hartmann, I., Van Benthem, J. (eds.) Handbook of Spatial Logics. Springer, Dordrecht (2007). doi:10.1007/978-1-4020-5587-4_10
36. Lind, D., Marcus, B.: An Introduction to Symbolic Dynamics and Coding. Cambridge University Press, Cambridge (1995)
37. Munkres, J.R.: Topology, 2nd edn. Prentice-Hall, Englewood Cliffs (2000)
38. Plaza, J.A.: Logics of public communications. In: Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems (1989)
39. Rendsvig, R.K.: Towards a theory of semantic competence. Master's thesis, Department of Philosophy and Science Studies and Department of Mathematics, Roskilde University (2011)
40. Rendsvig, R.K.: Diffusion, influence and best-response dynamics in networks: an action model approach. In: Proceedings of ESSLLI 2014 Student Session arXiv:1708.01477 (2014)
41. Rendsvig, R.K.: Pluralistic ignorance in the bystander effect: Informational dynamics of unresponsive witnesses in situations calling for intervention. Synthese **191**(11), 2471–2498 (2014)
42. Rendsvig, R.K.: Model transformers for dynamical systems of dynamic epistemic logic. In: Hoek, W., Holliday, W.H., Wang, W. (eds.) LORI 2015. LNCS, vol. 9394, pp. 316–327. Springer, Heidelberg (2015). doi:10.1007/978-3-662-48561-3_26
43. Sadzik, T.: Exploring the iterated update universe. ILLC PP-2006-26 (2006)
44. de Vries, J.: Topological Dynamical Systems. de Gruyter, Berlin (2014)

# Appendix to Paper IV

# Turing Completeness of
# Finite Epistemic Programs

Dominik Klein⋆ and Rasmus K. Rendsvig⋆⋆

In this note, we present the proof of Lemma 1.1 of [8], namely that the class of **epistemic programs** [1] is **Turing complete**. Following preliminary definitions in Section 1, Section 2 states and proves the theorem.

## 1 Definitions

Let there be given a countable set $\Phi$ of **atoms** and a finite set $I$ of **agents**. Where $p \in \Phi$ and $i \in I$, define the **language** $\mathcal{L}$ by

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i\varphi.$$

We use relational semantics to evaluate formulas. A **Kripke model** for $\mathcal{L}$ is a tuple $M = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket)$ where $\llbracket M \rrbracket$ is a countable, non-empty set of **states**, $R : I \longrightarrow \mathcal{P}(\llbracket M \rrbracket \times \llbracket M \rrbracket)$ assigns to each $i \in I$ an **accessibility relation** $R_i$, and $\llbracket \cdot \rrbracket : \Phi \longrightarrow \mathcal{P}(\llbracket M \rrbracket)$ is a **valuation**, assigning to each atom a set of states. With $s \in \llbracket M \rrbracket$, call $Ms = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket, s)$ a **pointed Kripke model**. The used semantics are standard (see e.g. [4,7]), including the modal clause:

$$Ms \models \Box_i\varphi \text{ iff for all } t : sR_it \text{ implies } Mt \models \varphi.$$

Pointed Kripke models may be updated using **action models** and **product update** [1–3,5,6]. We here invoke a set of mild, but non-standard, requirements to fit the framework of [8].

A **multi-pointed action model** is a tuple $\Sigma\Gamma = (\llbracket \Sigma \rrbracket, \mathsf{R}, pre, \Gamma)$ where $\llbracket \Sigma \rrbracket$ is a countable, non-empty set of **actions**. The map $\mathsf{R} : I \rightarrow \mathcal{P}(\llbracket \Sigma \rrbracket \times \llbracket \Sigma \rrbracket)$ assigns an **accessibility relation** $\mathsf{R}(i)$ on $\Sigma$ to each agent $i \in I$. The map $pre : \llbracket \Sigma \rrbracket \rightarrow \mathcal{L}$ assigns to each action a **precondition**. Finally, $\emptyset \neq \Gamma \subseteq \llbracket \Sigma \rrbracket$ is the set of designated actions.

Where $X$ is a set of pointed Kripke models, call $\Sigma\Gamma$ **deterministic** if $\models$ $pre(\sigma) \wedge pre(\sigma') \rightarrow \bot$ for each $\sigma \neq \sigma' \in \Gamma$.

Let $\Sigma\Gamma$ be deterministic over $X$ and let $Ms \in X$. Then the **product update** of $Ms$ with $\Sigma\Gamma$, denoted $Ms \otimes \Sigma\Gamma$, is the pointed Kripke model $(\llbracket M\Sigma \rrbracket, R', \llbracket \cdot \rrbracket', s')$

⋆ Department of Philosophy, Bayreuth University, and Department of Political Science, University of Bamberg. dominik.klein@uni-bayreuth.de

⋆⋆ Theoretical Philosophy, Lund University, and Center for Information and Bubble Studies, University of Copenhagen. rendsvig@gmail.com

with

$$\llbracket M\Sigma \rrbracket = \{(s, \sigma) \in \llbracket M \rrbracket \times \llbracket \Sigma \rrbracket : (M, s) \models pre(\sigma)\}$$
$$R' = \{((s, \sigma), (t, \tau)) : (s, t) \in R_i \text{ and } (\sigma, \tau) \in \mathsf{R}_i\}, \text{ for all } i \in N$$
$$\llbracket p \rrbracket' = \{(s, \sigma) : s \in \llbracket p \rrbracket\}, \text{ for all } p \in \Phi$$
$$s' = (s, \sigma) : \sigma \in \Gamma \text{ and } Ms \models pre(\sigma)$$

As $\Sigma\Gamma$ is assumed deterministic over $X$ at most one suitable $s'$ exists. If $Ms \models \neg pre(\sigma)$ for all $\sigma \in \Gamma$, $Ms \otimes \Sigma\Gamma$ is undefined.

## 2 Theorem and Proof

Call a finite, deterministic multi-pointed action an **epistemic program**.[1] We then show:

**Theorem 1.** *The set of epistemic programs is Turing complete.*

*Remark 1.* The proof uses a strict sub-class of the mentioned action models, all with only equivalence relations as suited for multi-agent $S5$ logics, and requires only the use of finite, $S5$ pointed Kripke models. □

**Preliminaries.** Define a **Turing machine** as a 7-tuple

$$\mathsf{M} = (Q, q_0, q_h, \Gamma, b, \Sigma, \delta)$$

where $Q$ is a finite set of **states** with $q_0 \in Q$ the **start state** and $q_h \in Q$ the **halt state**, $\Gamma$ a finite set of **tape symbols** with $b \in \Gamma$ the **blank symbol** and $\Sigma = \Gamma \backslash \{b\}$ the set of **input symbols**, and $\delta$ a partial function

$$\delta : Q \times \Gamma \to Q \times \Gamma \times \{l, h, r\}$$

with $\delta(q_h, \gamma)$ undefined for all $\gamma \in \Gamma$, called the **transition function**. If $\delta(q, \gamma)$ is undefined, the machine will halt.

A Turing machine acts on a bi-infinite **tape** with cells indexed by $\mathbb{Z}$ and labeled with $\Gamma$ such that only $b$ occurs on the tape infinitely often. With the machine in state $q \in Q$ and reading label $\gamma \in \Gamma$, the transition function determines a possibly new state of the machine $q' \in Q$, a symbol $s'$ to replace $s$ at the current position on the tape, and a movement of the metaphorical "read/write head": Either one cell to the left $(l)$, none (stay here, $h$), or one cell to the right $(r)$.

A **configuration** of a machine is fully given by *i)* the current labeling of the tape, *ii)* the position of the $r/w$-head on the tape, and *iii)* the state of the machine. The space of possible configurations of a machine $\mathsf{M}$ is thus $\mathfrak{C} = \mathfrak{T} \times \mathbb{Z} \times Q$, where $\mathfrak{T}$ is the set of bi-infinite strings $\mathsf{t} = (\ldots, \gamma_{-2}, \gamma_{-1}, \gamma_0, \gamma_1, \gamma_2, \ldots)$

---
[1] The term stems from the seminal [1] in which postconditions where not included.

over $\Gamma$ such that only $b$ occurs infinitely often in $\mathfrak{t}$. The transition function $\delta$ of $\mathsf{M}$ may thus be recast as a partial function $\delta : \mathfrak{C} \to \mathfrak{C}$.

We want to recast $\delta$ in a slightly different manner. Each tape $\mathfrak{t}$ has infinite head and tail consisting solely of $b$s. Ignoring all but a finite segment of these yields a finite non-unique representation of the tape. Formally, for a string $\mathfrak{t} = (\ldots, \gamma_{-2}, \gamma_{-1}, \gamma_0, \gamma_1, \gamma_2, \ldots)$ and $k < k'$ let $\mathfrak{t}_{\restriction [k, k']}$ be the substring $(\gamma_k, \ldots, \gamma_{k'})$. The set of all such finite representations of $\mathfrak{T}$ is then given by $T = \{t = (\gamma_k, \ldots, \gamma_{k'}) \colon \exists \mathfrak{t} \in \mathfrak{T} \text{ s.t. } t = \mathfrak{t}_{\restriction [k, k']} \text{ and } \forall j < k, \forall j' > k', \mathfrak{t}_j = \mathfrak{t}_{j'} = b\}$. Each $t \in T$ corresponds to a unique $\mathfrak{t} \in \mathfrak{T}$. Conversely, each configuration $\mathfrak{c} = (\mathfrak{t}, i, q) \in \mathfrak{C}$ may be represented by the equivalence class $\{(\mathfrak{t}_{\restriction [k, k']}, i, q) \colon k < k'\}$ of its finite approximations. In each such equivalence class, there exists representatives for which the position $i$ of the read-write head is "on the tape", i.e., satisfies that $\gamma_i \in t$. We impose this as a requirement and define a restricted equivalence class for each $\mathfrak{c} = (\mathfrak{t}, i, q) \in \mathfrak{C}$ by $[\mathfrak{c}] = \{(\mathfrak{t}_{\restriction [k, k']}, i, q) \colon k \le i \le k'\}$. With $\mathsf{C} = \{[\mathfrak{c}] \colon \mathfrak{c} \in \mathfrak{C}\}$, i.e., the set of equivalence classes of finite representations of configurations for which the read-write head is on the finite tape, the transition function may finally be recast as a partial function $\delta : \mathsf{C} \to \mathsf{C}$.

*Remark 2.* The class of Turing machines with $\Gamma = \{0, 1\}, b = 0$, is Turing complete. Henceforth, we restrict attention to this sub-class. □

## Proof

To prove Theorem 1, it must be shown that any Turing machine can be simulated by an epistemic program. We show that as follows: First, we define an invertible operator $\mathsf{K}$ that for any finite representation of a configuration $\mathsf{c} \in [\mathfrak{c}] \in \mathsf{C}$ produces a pointed Kripke model $\mathsf{K}(\mathsf{c})$. Second, we define an epistemic program $\Sigma_\Gamma$ which satisfies that

$$\mathsf{K}^{-1}(\mathsf{K}(\mathsf{c}) \otimes \Sigma_\Gamma) \in \delta([\mathfrak{c}]), \tag{1}$$

for any $[\mathfrak{c}] \in \mathsf{C}$. Hence $P$ may be used to calculate the trajectory of $\delta$.

**Machine, Language and Logic.** Fix a Turing machine $\mathsf{M}$ with states $Q$, and fix from this a set of relation indices $Q' = Q \cup \{a, b, 1\}$. Let the modal language $\mathcal{L}$ be based on the single atom $p$ and operators $\Box_i, i \in Q'$.

**Configuration Space.** Let $\mathsf{C} = \{[\mathfrak{c}] \colon \mathfrak{c} \in \mathfrak{C}\}$ be the set of equivalence classes of finite representations of configurations for which the read-write head is on the finite tape for $\mathsf{M}$ and let $\mathsf{c} = (t, i, q) \in \mathsf{C}$. We construct a pointed Kripke model $\mathsf{K}(\mathsf{c})$ representing $(t, i, q)$. We exemplify the construction to be in Fig. 1.

First, in three steps, we construct the set of worlds: *i)* Construct slightly too many "tape cells": Let $\lceil u \rceil = \max\{|k|, |k'|\}$ if this is even, else let $\lceil u \rceil = \max\{|k|, |k'|\} + 1$ and take a set of worlds $C = \{c_j \colon -(\lceil u \rceil + 5) \le j \le \lceil u \rceil + 5\}$. *ii)* Represent the content of a cell: Add worlds $S = \{s_j \colon \gamma_j = 1\}$ to indicate
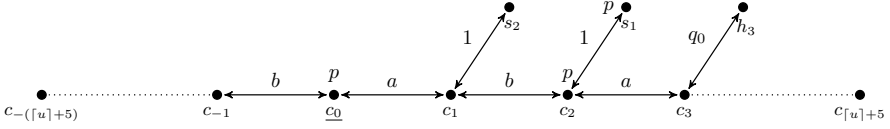
**Fig. 1.** An emulation of a Turing machine in state $q_0$ with the read/write head in position 3. Cells 1 and 2 are marked with 1 (or $A$), cells -1, 0 and 3 are not.

"cells" with the unique non-blank "symbol" 1. Let $iii$) Add a "read/write head": Let $H = \{h_j : j = r/w\}$. Finally, we define the set of worlds as $W = C \cup S \cup H$.

Second, we add relations between the worlds, also in three steps. In the following let $R^*$ denote the reflexive, symmetric, and transitive closure of the relation $R$ on a given base set, here $W$. In particular $(w, w) \in R^*$ for all $w \in W$. $i$) We structure the cells $c_i$ into a tape using relations $R_a$ and $R_b$: $R_a = \{(c_j, c_{j+1}) : j \text{ is even}\}^*$, $R_b = \{(c_j, c_{j+1}) : j \text{ is odd}\}^*$. $ii$) We attach the non-blank symbols to the appropriate cells: Let $R_1 = \{(c_j, s_j) : s_j \in S\}^*$. $iii$) We mount the read/write head at the correct position and in the correct state, $q$: Let $R_q = \{(c_j, h_j) : h_j \in H\}^*$. For the remaining states $q' \in Q \backslash \{q\}$, let $R_{q'} = \{\}^*$. Finally, let $[\![p]\!] = \{c_j, s_j, h_j \in C \cup S \cup H : j \text{ is even}\}$ and the actual world be $c_0$.

We thus obtain a pointed Kripke model $\mathsf{K}(\mathsf{c}) = (W, \{R_i\}_{i \in Q'}, [\![\cdot]\!], c_0)$ for the finite configuration representation $\mathsf{c}$ of Turing machine $\mathsf{M}$. Figure 1 illustrates this, depicting the model $\mathsf{K}(\mathsf{c})$ for configuration $\mathsf{c} = (t, 3, q_0)$. Given $\mathsf{K}(\mathsf{c})$, we may clearly invert the construction process and re-obtain an element of $[\mathsf{c}]$. Finally let $\mathcal{C} = \{\mathsf{K}(\mathsf{c}) : \mathsf{c} \in \mathsf{C}\}$.

**Expressible Properties.** To construct an epistemic program that simulates $\delta : \mathsf{C} \rightarrow \mathsf{C}$, i.e., satisfies Eq. (1), we take advantage of the fact that various properties of configurations are modally expressible. Hence, we can use these as preconditions. The relevant properties and formulas are summarized in Table 1.

**Epistemic Program.** We construct $\Sigma_\Gamma = (\Sigma, \{R_j\}_{j \in Q'}, pre, \Gamma)$, an epistemic program that simulates $\delta : \mathsf{C} \rightarrow \mathsf{C}$, cf. Eq. (1). An example of such an epistemic program is illustrated in Fig. 2. We argue for the adequacy of the epistemic program in parallel with its construction. In the following, the precondition of action $\sigma_\varphi$ is the formula $\varphi$.

**Actual Actions, Halting, and Tape Enlargement.** Let the set of actual actions be given by $\Gamma = \{\gamma_\varphi : \varphi \in \Phi\}$ with $\Phi = \{R, L, 2_{AM}\} \cup \{h_{qi}, l_{qi}, r_{qi} : q \in Q, i \in \{0, 1\}\}$, cf. Table 1.

Then, for any $\mathsf{K}(\mathsf{c}) \in \mathcal{C}$, for every cell state $c_j \in C$ of $\mathsf{K}(\mathsf{c})$, $c_j$ will satisfy exactly one of the formulas in $\Phi$. $\Sigma_\Gamma$ is thus deterministic over $\mathcal{C}$, and the actual world of $\mathsf{K}(\mathsf{c}) \otimes \Sigma_\Gamma$ is a cell. Finally, formulas from $\Phi$ are only satisfied at cell

| Property | Formula |
|---|---|
| Being a cell[†] | $c := (\Diamond_a p \vee \Diamond_b p) \wedge \Diamond_a \neg p$ |
| Being a 1 symbol | $s := \neg c \wedge \Diamond_1 c$ |
| Being a cell with symbol 1 | $1 := c \wedge \Diamond_1 \neg c$ |
| Being a cell with symbol 0 | $0 := c \wedge \neg \Diamond_1 \neg c$ |
| Being the cell of the $r/w$-head is while the machine is in state $q$ | $h_j := c \wedge \Diamond_q \neg c$ |
| Being the cell immediately left of the $r/w$-head while the machine is in state $q$ | $l_q := c \wedge \neg h_q \wedge ((p \wedge \Diamond_a h_q) \vee (\neg p \wedge \Diamond_b h_q))$ |
| Being the cell immediately right of the $r/w$-head while the machine is in state $q$ | $r_j := c \wedge \neg h_q \wedge ((p \wedge \Diamond_b h_q) \vee (\neg p \wedge \Diamond_a h_q))$ |
| Being the cell of/im. left of/im. right of the $r/w$-head while the machine is in state $j$ and the cell of the $r/w$-head contains a 1/0 | $h_{q1}/l_{q1}/r_{q1}/h_{q0}/l_{q0}/r_{q0}$ : Replace $c$ in $h_q/l_q/r_q$ with formula 1/0. |
| Being a cell at least two cells away from the $r/w$-head | $h_{\geq 2} := c \wedge \bigwedge_{q \in Q} (\neg h_q \wedge \neg l_q \wedge \neg r_q)$ |
| Being the rightmost[†] cell | $R := c \wedge \Box_b \neg p$ |
| Being the leftmost[†] cell | $L := c \wedge \Box_a \neg p$ |
| Being the penultimate cell to the right | $PR = c \wedge \neg R \wedge \Diamond_a R$ |
| Being the penultimate cell to the right | $PL = c \wedge \neg L \wedge \Diamond_b L$ |
| Being at least two steps away from the $r/w$-head and not being the left- or rightmost cell | $2_{AM} := h_{\geq 2} \wedge \neg R \wedge \neg L$ |

**Table 1.** Expressible properties used as preconditions. **Notes.** †: Recall that the extreme states of $C$ are $c_{-(\lceil u \rceil + 5)}$ and $c_{\lceil u \rceil + 5}$ with $\lceil u \rceil$ even.

states of $\mathsf{K}(\mathsf{c})$. Jointly, this implies that $\Gamma$ "copies" the set of tape cells from $\mathsf{K}(\mathsf{c})$ to $\mathsf{K}(\mathsf{c}) \otimes \Sigma \Gamma$.

The copied over tape may not be long enough for future operations, so we include a set of actions to preemptively enlarge it.[2] To this end, let $\Upsilon = \{v_L, v_{PL}, v_R, v_{PR}\}$. The precondition $\varphi$ of each $v_\varphi \in \Upsilon$ is satisfied by exactly one state $c_j$ of $\mathsf{K}(\mathsf{c})$ which is a cell state. These cell state will thus have two successors in $\mathsf{K}(\mathsf{c}) \otimes \Sigma \Gamma$: $(c_j, \gamma_\varphi)$ defined before and $(c_j, v_\varphi)$. We thus gain four new cell states. Setting

$$R_a = \{(\gamma_\varphi, \gamma_\psi) \colon \varphi, \psi \in \Phi \backslash \{L\}\}^* \cup \{(v_{PR}, v_R), (\gamma_L, v_{PL})\}^*$$
$$R_b = \{(\gamma_\varphi, \gamma_\psi) \colon \varphi, \psi \in \Phi \backslash \{R\}\}^* \cup \{(\gamma_R, v_{PR}), (v_{PL}, v_L)\}^*$$

---

[2] To save tape, this could be done in a more economical manner, only creating extra cells where actually needed.
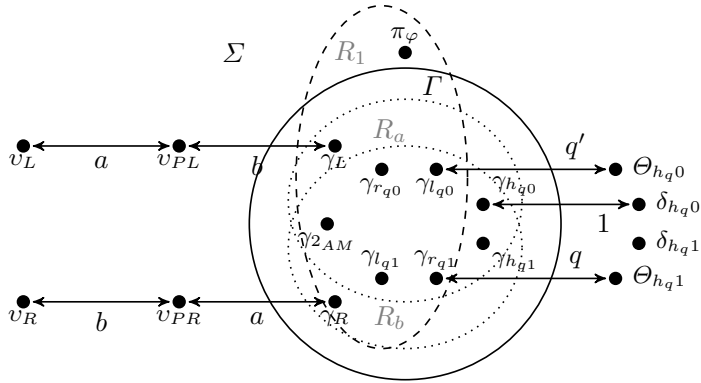
**Fig. 2.** An illustration of the epistemic program $(\Sigma, \Gamma)$ for a Turing machine with t $\delta(q,0) = (q',1,l)$ and $\delta(q,1) = (q,0,r)$. That $\Theta_{h_q0}R_{q'}\gamma_{l_{q0}}$ ensures that on input $(q,0)$ the $r/w$-head moves to the left and the machine is set to state $q'$ and the relation $\gamma_{h_q0}R_1\delta_{h_q0}$ ensures that the content of the current cell is set to 1. Similarly that $\Theta_{h_q1}R_q\gamma_{r_{q1}}$ ensures that on input $(q,1)$ the $r/w$-head moves to the right, the machine remains in state $q$ and the absence of relation $\gamma_{h_q1}R_1\delta_{h_q1}$ ensures that the content of the current cell is set to 0.

copies over the tape structure and suitably extends it to the new cell states, which are as the left most, penultimate left, penultimate right, and right most tape cells. Fig. 3 illustrates.
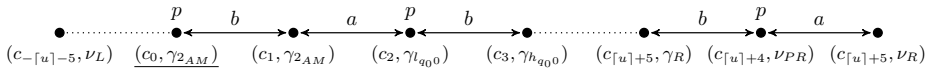


**Fig. 3.** Illustration of the extended tape resulting from applying $P$ to the model in Figure 1.

**Symbol Transfer.** We copy all symbols from the old tape to the new, safe for the symbol at the current position of the $r/w$-head. To this end, add an action $\pi_\varphi$ with $\varphi = s \wedge \neg \Diamond_1(\bigvee_{\Diamond q \in Q} h_q)$. The formula $\varphi$ is then satisfied in $\mathsf{K}(\mathsf{c})$ exactly at the symbols states $s_j \in S$ on which the $r/w$-head is not. Let $\Gamma' = \{\gamma_\varphi \colon \varphi \in \Phi\}$ with $\Phi = \{R, L, 2_{AM}\} \cup \{l_{qi}, r_{qi} \colon q \in Q, i \in \{0,1\}\}$. Requiring that $(\Gamma' \times \{\pi_\varphi\})^* \subseteq R_1$ ensures that the symbol states copied over to $\mathsf{K}(\mathsf{c}) \otimes \Sigma\Gamma$ are connected to the correct cell world. We give the precise definition of $R_1$ below.

**Symbol Writing.** We implement the symbol writing part of the transition function $\delta$. Define a new set of actions by

$$\Delta = \{\delta_{h_q i} \colon q \in Q, i \in \{0,1\} \text{ and } \delta(i,q) \text{ is defined}\}.$$

At most one action from $\Delta$ will have its precondition satisfied at any $\mathsf{K}(\mathsf{c})$ and just in case $\delta(\mathsf{c})$ is defined. The world satisfying this precondition is a cell world, $c_j$, which will have two successors in $\mathsf{K}(\mathsf{c}) \otimes P$:[3] a cell world successor $(c_j, \gamma_{h_q i})$ defined above and a symbol world successor $(c_j, \delta_{h_q i})$ defined here. We ensure that the emulation writes the correct symbol by connecting $(c_j, \delta_{h_q i})$ to $(c_j, \gamma_{h_q i})$ by $R_1$ or not: Let

$$R_{tmp} = \{\{(\delta_{h_q i}, \gamma_{h_q i}) \colon \gamma \in \varGamma\} \mid \delta(i, q) = (\cdot, 1, \cdot)\}$$

and let $R_1 = ((\varGamma' \times \{\pi_\varphi\}) \cup R_{tmp})^*$. This and the above ensures that the emulation produces a correctly labeled tape.

**State Change and Head Repositioning.** We finally implement the state change and head repositioning encoded by $\delta$. To this end, define a set of events

$$\varTheta = \{\theta_{h_q i} \colon q \in Q, i \in \{0, 1\} \text{ and } \delta(i, q) \text{ is defined}\}.$$

Again, at most one action from $\varTheta$ will have its precondition satisfied at any $\mathsf{K}(\mathsf{c})$ and just in case $\delta(\mathsf{c})$ is defined. The world satisfying this precondition is a cell world, $c_j$, which will hence have two successors in $\mathsf{K}(\mathsf{c}) \otimes \varSigma \varGamma$:[4] a cell world successor $(c_j, \gamma_\varphi)$ defined above and a $r/w$-head world successor $(c_j, \theta_{h_q i})$ defined here. We "mount" the $r/w$-head world at the correct position and in the correct state using the relations $\{R_{q'}\}_{q' \in Q}$: For all $q' \in Q$, let

$$R_{q'} = \{(\gamma_{xq'}, \theta_{h_q i}) \colon \delta(q, i) = (q', \cdot, x), i \in \{0, 1\}, q \in Q\}^*.$$

The definition of $\{R_q\}_{q \in Q}$ ensures that the $r/w$-head is moved and changes state appropriately, whenever $\delta(i, q)$ is defined. When $\delta(i, q)$ is not defined, the $r/w$-head world $(c_j, \theta_{h_q i})$ will be disconnected from the tape cell worlds. In that case, $\mathsf{K}(\mathsf{c}) \otimes P$ will not be in $\mathcal{C}$, and the emulation is said to halt. This concludes the construction and proof. QED

*Remark 3.* The proof generalizes to $k$-tape Turing machines or bigger input symbol sets by replacing modality $\Box_1$ with $\Box_1, \Box_k$ and the corresponding formula 1 with $1, .., k$. □

## References

1. A. Baltag and L. S. Moss. Logics for Epistemic Programs. *Synthese*, 139(2):165–224, 2004.
2. A. Baltag, L. S. Moss, and S. Solecki. The Logic of Public Announcements, Common Knowledge, and Private Suspicions (extended abstract). In *TARK 1998*, pages 43–56. Morgan Kaufmann Publishers, 1998.

---

[3] Possibly three, see below.
[4] Possibly three, cf. the above.

3. J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, 2006.

4. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

5. H. van Ditmarsch and B. Kooi. Semantic Results for Ontic and Epistemic Change. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 87–117. Amsterdam University Press, 2008.

6. H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2008.

7. V. Goranko and M. Otto. Model Theory of Modal Logic. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier, 2008.

8. D. Klein and R. K. Rendsvig. Convergence, Continuity and Recurrence in Dynamic Epistemic Logic. *Accepted for LORI 2017, Sapporo, Japan*, 2017.

# Paper v

# METRICS FOR FORMAL STRUCTURES, WITH AN APPLICATION TO KRIPKE MODELS AND THEIR DYNAMICS

DOMINIK KLEIN AND RASMUS K. RENDSVIG

**Abstract.** This paper introduces a broad family of metrics applicable to finite and countably infinite strings, or, by extension, to formal structures serving as semantics for countable languages. We introduce the metrics in a general setting, and then focus on their application to sets of pointed Kripke models, a semantics for modal logics. We study the topological properties resulting from equipping sets of Kripke models with said metrics: We classify which metrics give rise to the same topological spaces, provide sufficient conditions for compactness, characterize clopen sets, characterize convergence by a logical convergence concept, and characterize isolated points. Moreover, we relate our metrical approach to concepts known from dynamic epistemic logic. In our main result, we show that a widely used type of model transformations, product updates with action models, gives rise to continuous maps in the induced topology, allowing to interpret iterated updates as discrete time dynamical systems.

**Keywords:** metric space, general topology, modal logic, Kripke model, model transformation, dynamic epistemic logic.

**§1. Introduction.** This paper introduces and investigates a family of metrics applicable to finite and countably infinite strings and, by extension, formal structures described by a countable language. This family of metrics is a weighted generalization of the Hamming distance [40]. On formal structures, each such metric corresponds to assigning positive weights to a chosen subset of some language describing the structure. The distance between two structure, then, is the sum of the weights of formulas on which the two structures differ in valuation.

While the approach is generally applicable, our main target is metrics on sets of pointed Kripke models, the most widely used semantic structures for modal logic. Apart from mathematical interest, there are several motivations for having a metric between pointed Kripke models. Among these are applications in iterated multi-agent belief revision [3, 4, 23, 25, 27, 46], logical meta-theory [37], and the application of dynamical systems theory to information dynamics modeled using dynamic epistemic logic [11–14, 42, 55, 56]. The latter is our main interest. In a nutshell, this paper contains a theoretical foundation for considering the logical dynamics of dynamic epistemic logic as discrete time **dynamical systems**: Compact metric spaces (of pointed Kripke models) together with continuous transition functions acting on them.

The paper lays this foundation by the following progression: In Section 2, we introduce a weighted generalization of the Hamming distance. In Section 3, we present a general case for applying the metrics to arbitrary sets of structures, given that the structures are abstractly described by a countable set of formulas within a possibly multi-valued semantics. In Section 4, we start focusing on pointed Kripke models: We show how the metrics may be applied to these and how the family of metrics defined here allows to represent various metrics that are natural from a modal logical point of view. Section 5 is on topological properties of the resulting spaces: We show that two metrics are topologically equivalent whenever they agree on which formulas of the modal language should receive strictly positive weight. The resulting topologies are a generalization of the *Stone topology*. We refer to these as *Stone-like topologies*, and show that each such is Hausdorff, totally disconnected and, under certain additional assumptions, compact. In the same section, we characterize the open, closed and clopen sets of Stone-like topologies and relate Stone-like topologies to a natural semantic topology, the *n-bisimulation topology*. In Section 6, we turn to convergence and limit points: We characterize convergence in Stone-like topologies by a logical analogue thereof, strengthening a result of [42]. Moreover, we show a characterization theorem for isolated points and exemplify perfect, imperfect and discrete spaces. In Section 7, we turn to mappings. In particular, we investigate the widely used family of mappings defined through *product updates with multi-pointed action models*, a particular graph product, well-known from the literature on dynamic epistemic logic [6–8, 13, 30]. As final our result, we show such product updates continuous with respect to Stone-like topologies, thus establishing the desired connection between dynamic epistemic logic and discrete time dynamical systems.

The approach and results presented may be applied in several domains, but our main goal is to bridge the fields of discrete time dynamical systems and dynamic epistemic logic. The aim is to render tools and theories from the field of dynamical systems applicable to logical dynamics, and in particular to the form of dynamics found in dynamic epistemic logic.[1] With this purpose in mind, it may seem a limited result to show specifically the mentioned family of maps continuous. However, we hold that the result and its proof constitute a significant step forward towards the general goal: Product update is a natural and flexible framework for modeling information dynamics, generalizing *truthful public announcements* [7, 36, 50], *untruthful announcements* [1] and *partially observable announcements* [7, 36]. Moreover, its flexibility allows it to mimic other natural update forms, such as arrow-deletion updates [44]. Such features have caused product update to be widely used and studied (see e.g. [2, 13, 21, 53, 61, 63] for usage and [6, 7, 16, 17, 21, 22, 29–32, 43, 45, 55, 56] for theoretical studies). Moreover, product update has inspired a host of similar update forms (see e.g. [5, 9, 10, 18, 19, 24, 26, 52, 54, 60]), to which we believe our framework is applicable *mutatis mutandis*.

---

[1]As to why this in turn should be a worthwhile pursuit, we refer the reader to [11–14, 55, 56] and in particular [42], which rests on results shown here.

REMARK. This paper is not self-contained. We only include definitions for selected standard terms and do so mainly to fix notation. For modal logic notions that remain undefined here, refer to e.g. [20, 38]. For topological notions, refer to e.g. [49]. For more than the bare minimum of standard notions and notations of dynamic and epistemic logic rehearsed here, see e.g. [6–9, 13, 16, 29, 30, 33, 41, 50, 51].

**1.1. Related Work.** Metrics for formal structures have been considered elsewhere. In particular, Caridroit *et al.* [23] present metrics on pointed Kripke models for the purpose of belief revision. In Example 10, we show their six semantic metrics, defined on finite sets of finite Kripke models, special cases of our syntactic approach. The authors also consider a semantic similarity measure of Aucher [3, 4] from which they define a distance between finite pointed Kripke models. The exact weights in this distance are somewhat involved and we do not attempt a quantitative comparison. As to a qualitative analysis, then neither Caridroit *et al.* nor Aucher offer any form of topological analysis of the metric, making comparison non-straightforward. However, as the fundamental measuring component in Aucher's distance is based on degree of *n-bisimilarity*, we conjecture that the topology on the spaces of Kripke models generated by this distance is the *n-bisimulation topology*, defined in Section 5.4. The same section offers a comparison between this topology and the Stone topology. The *n-bisimulation metric* (see Example 8) inducing the *n*-bisimulation topology is inspired by the metric introduced by Goranko in [37] on sets of theories of first-order structures. Finally, Sokolsky *et al.* [58] introduce a quantitative bisimulation distance for finite, labeled transition systems and consider its computation. Again, we conjecture the induced topology is the *n*-bisimulation topology.

As Sokolsky *et al.* remark, distances on rooted graphs (or pointed Kripke models) may fall into one of two categories: Either *extremal*, where the first met difference between graphs fully determines the distance, or *aggregate*, where all differences are summarized. The distance of Sokolsky *et al.* is aggregate, as is that of Aucher and some of the distances of Caridroit *et al.* The remaining are extremal, together with the distance of Goranko and the *n*-bisimulation distance. The syntactic approach we introduce allows for either option.

**§2. Generalizing the Hamming Distance.** The method we propose for defining distances between pointed Kripke models is a particular instance of a more general approach. This general approach concerns distances between finite or infinite strings of letters from some given set, $V$. In a logical context, the set $V$ may be thought of as containing the possible truth values for some logic. For classical logics such as normal modal logic, $V$ would be binary, and the resulting strings be made of, e.g., 0s and 1s. We think of pointed Kripke models as being represented by such countably infinite strings: Given some enumeration of the corresponding modal language, a string will have a 1 on place $k$ just in case the model satisfies the $k$th formula, 0 else. See Section 4 for details.

A distance on sets of finite strings of a fixed length has been known since 1950, when it was introduce by R.W. Hamming [40]. Informally, the **Hamming distance** between two such strings is the number of places on which the two

strings differ. Clearly, this distance is, in general, not well-defined on sets of infinite strings.

However, for faithfully representing pointed Kripke models as strings of formula truth values, we need to work with infinite strings. This is the case as the modal language is infinite and, *a fortiori*, there are, in general, infinitely many modally expressible mutually non-equivalent properties of pointed Kripke models. We return to this below. To accommodate infinite strings, we generalize the Hamming distance:

DEFINITION. Let $S$ be a set of strings over a set $V$ such that either $S \subseteq V^n$ for some $n \in \mathbb{N}$, or $S \subseteq V^\omega$. For any $s, s' \in S$, any $k \in \mathbb{N}$, let

$$d_k(s, s') = \begin{cases} 0 & \text{if } s_k = s'_k, \text{ or if } s \in V^n \text{ and } k > n \\ 1 & \text{else} \end{cases}$$

Let $w : \mathbb{N} \to \mathbb{R}_{>0}$ assign a strictly positive **weight** to each natural number such that $(w(k))_{k \in \mathbb{N}}$ forms a convergent series, i.e., $\sum_{k=1}^{\infty} w(k) < \infty$.

The function $d_w : S \times S \to \mathbb{R}$ is then defined by, for each $s, s' \in S$

$$d_w(s, s') = \sum_{k=1}^{\infty} w(k) d_k(s, s').$$

PROPOSITION 1. *Let $S$ and $d_w$ be as above. Then $d_w$ is a metric on $S$.*

PROOF. Each $d_w$ is a metric on $S$ as it satisfies for all $s, s', s'' \in X$
*Positivity*, $d_w(s, s') \geq 0$: The sum defining $d_w$ contains only non-negative terms.
*Identity of indiscernibles*, $d_w(s, s') = 0$ iff $s = s'$: $d_w(s, s') = 0$ iff $d_k(s, s') = 0$ for all $k$ iff $s_k = s'_k$ for all $k$ iff $s = s'$.
*Symmetry*, $d_w(s, s') = d_w(s', s)$: As $d_k(s, s') = d_k(s', s)$ for all $k$, we get $d_w(s, s') = d_w(s', s)$.
*Triangular inequality*, $d_w(s, s'') \leq d_w(s, s') + d_w(s', s'')$: If $s$ and $s''$ agree on any position $k$, we have $d_k(s, s'') = 0$ and hence $w(k)d_k(s, s'') \leq w(k)d_k(s, s') + w(k)d_k(s', s'')$. If $s$ and $s''$ differ on $k$, then either $s$ and $s'$ or $s'$ and $s''$ have to differ on the same position. Hence $w(k)d_k(s, s'') \leq w(k)d_k(s, s') + w(k)d_k(s', s'')$, for each $k$. Together, this establishes the triangular equality:
$\sum_{k=1}^{\infty} w(k)d_k(s, s'') \leq \sum_{k=1}^{\infty} w(k)d_k(s, s') + \sum_{k=1}^{\infty} w(k)d_k(s', s'').$                ⊣

REMARK. The Hamming distance is a special case of the family defined here. For $S \subseteq \mathbb{R}^n$, the Hamming distance $d_H$ is defined, cf. [28], by $d_H(s, s') = |\{i : 1 \leq i \leq n, s_i \neq s'_i\}|$. This function is a member of the above family, namely $d_h$ with strictly positive weights $h(k) = 1$ for $1 \leq k \leq n$, $h(k) = 2^{-k}$ for $k > n$.

**§3. Metrics for Formal Structures.** The metrics defined above may be indirectly applied to any set of structures that serves as semantics for a countable language. In essence, what is required is simply an assignment of suitable weights to formulas of the language. To illustrate the generality of the approach, we initially take the following inclusive view on semantic valuation:

DEFINITION. Let a **valuation** be any map $\nu : X \times D \longrightarrow V$ where $X$ and $V$ are arbitrary sets, and $D$ is countable. Refer to elements of $X$ as **structures**, to $D$ as the **descriptor**, and to elements of $V$ as **values**.

A valuation $\nu$ assigns a value from $V$ to every pair $(x, \varphi) \in X \times D$. Jointly, $\nu$ and $X$ thus constitute a $V$-valued semantics for the descriptor $D$.

REMARK. The term *descriptor* is used here and below to emphasize the potential *lack of grammar* in the set $D$. The descriptor *may* be a formal language, but it is not required to be. In particular, the descriptor may be a *strict subset* of a formal language, containing only some formulas that are judged to be of special interest. This is exemplified in Section 4.5.

Two structures in $X$ may be considered equivalent by $\nu$, i.e., be assigned identical values for all $\varphi \in D$. To avoid that two non-identical, but semantically equivalent, structures receive a distance of zero (and thus violate the *identity of indiscernibles* requirement of a metric), metrics are defined over suitable quotients:

DEFINITION. Given a valuation $\nu : X \times D \longrightarrow V$ and a subset $D'$ of $D$, denote by $\boldsymbol{X}_{D'}$ the **quotient of $X$ under $D'$ equivalence**, i.e., $\boldsymbol{X}_{D'} = \{\boldsymbol{x}_{D'} : x \in X\}$ with $\boldsymbol{x}_{D'} = \{y \in X : \nu(y, \varphi) = \nu(x, \varphi) \text{ for all } \varphi \in D'\}$.

When the descriptor $D$ is clear from context, we write $\boldsymbol{x}$ for elements of $\boldsymbol{X}_D$. We also write $\nu(\boldsymbol{x}, \varphi)$ for $\nu(x, \varphi)$ when $\varphi \in D$.

REMARK. Quotients are defined for subsets $D'$ of $D$ in accordance with the comment concerning the term *descriptor* above: For some structures, it may be natural to define a semantics for a complete formal language, $\mathcal{L}$. However, if only a subset $D' \subseteq \mathcal{L}$ is deemed relevant in determining distance, it is natural to focus on structures under $D'$ equivalence. The terminological usage is consistent as the subset $D'$ is itself a descriptor for the restricted map $\nu_{|X \times D'}$.

Finally, we obtain a family of metrics on a quotient $\boldsymbol{X}_D$ in the following manner:

DEFINITION. Let $\nu : X \times D \longrightarrow V$ be a valuation and $\varphi_1, \varphi_2, ...$ an enumeration of $D$. For all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_D$ and all $k \in \mathbb{N}$, let

$$d_k(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } \nu(\boldsymbol{x}, \varphi_k) = \nu(\boldsymbol{y}, \varphi_k), \text{ or if } k > |D| \\ 1 & \text{else} \end{cases}$$

Call $w : D \longrightarrow \mathbb{R}_{>0}$ a **weight function** if it assigns a strictly positive **weight** to each $\varphi \in D$ such that $\sum_{\varphi \in D}(w(\varphi)) < \infty$.

The function $d_w : \boldsymbol{X}_D \times \boldsymbol{X}_D \longrightarrow \mathbb{R}$ is then defined by, for each $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_D$

$$d_w(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{|D|} w(\varphi_k) d_k(\boldsymbol{x}, \boldsymbol{y}).$$

The set of such maps $d_w$ is denoted $\mathcal{D}_{(X, \nu, D)}$.

PROPOSITION 2. *Every $d_w \in \mathcal{D}_{(X, \nu, D)}$ is a metric on $\boldsymbol{X}_D$.*

PROOF. That $d_w$ is a metric on $\boldsymbol{X}_D$ is argued using Proposition 1: For each $\boldsymbol{x} \in \boldsymbol{X}_D$ we define a string $s_{\boldsymbol{x}}$ of length $|D|$ by $s_{\boldsymbol{x},i} = \nu(\boldsymbol{x}, \varphi_i)$. Let $S = \{s_{\boldsymbol{x}} : \boldsymbol{x} \in \boldsymbol{X}_D\}$. Then the map $f : \boldsymbol{X}_D \to S$ given by $f(\boldsymbol{x}) = s_{\boldsymbol{x}}$ is a bijection. Let $w' : \mathbb{N} \to \mathbb{R}_{>0}$ be given by $w'(k) = w(\varphi_k)$ for all $k \leq |D|$ and $w'(k) = \frac{1}{2^k}$

else, and let $d_{w'}$ be the metric on $S$ given by $w'$ cf. Proposition 1. Then $d_w(\boldsymbol{x}, \boldsymbol{y}) = d_{w'}(s_{\boldsymbol{x}}, s_{\boldsymbol{y}})$ for all $x, y \in X$. Hence, $d_w$ is a metric on $\boldsymbol{X}_D$. ⊣

REMARK. The choice of descriptor affect both the coarseness of the space $\boldsymbol{X}_D$ as well as the metrics definable. We return to this point several times below.

**§4. The Application to Pointed Kripke Models.** To apply the metrics to pointed Kripke models, we follow the above approach. The set $X$ will be a set of pointed Kripke models and $D$ a set of modal logical formulas. Interpreting the latter over the former using standard modal logical semantics gives rise to a binary set of values, $V$, and a valuation function $\nu : X \times D \to V$ equal to the classic interpretation of modal formulas on pointed Kripke models. In the following, we consequently omit references to $\nu$, writing $\mathcal{D}_{(X,D)}$ for the family of metrics $\mathcal{D}_{(X,\nu,D)}$.

**4.1. Pointed Kripke Models, their Language and Logics.** Let be given a **signature** consisting of a countable, non-empty set of propositional **atoms** $\Phi$ and a countable, non-empty set of **operator indices**, $\mathcal{I}$. Call the signature **finite** when both $\Phi$ and $\mathcal{I}$ are finite. The **modal language** $\mathcal{L}$ for $\Phi$ and $\mathcal{I}$ is given by the BNF

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box_i \varphi$$

with $p \in \Phi$ and $i \in \mathcal{I}$. The language $\mathcal{L}$ is countable.

A **Kripke model** for $\Phi$ and $\mathcal{I}$ is a tuple $M = (\llbracket M \rrbracket, R, \llbracket \cdot \rrbracket)$ where

- $\llbracket M \rrbracket$ is a non-empty set of **states**;
- $R : \mathcal{I} \to \mathcal{P}(\llbracket M \rrbracket^2)$ assigns to each $i \in \mathcal{I}$ an **accessibility relation** $R(i)$;
- $\llbracket \cdot \rrbracket : \Phi \to \mathcal{P}(\llbracket M \rrbracket)$ is a **valuation**, assigning to each atom a set of states.

A pair $(M, s)$ with $s \in \llbracket M \rrbracket$ is a **pointed Kripke model**. For the pointed Kripke model $(M, s)$, the shorter notation $Ms$ is used. For $R(i)$, we write $R_i$.

The modal language is evaluated over pointed Kripke models with standard semantics:

$$
\begin{array}{lll}
Ms \vDash p & \text{iff} & s \in \llbracket p \rrbracket, \text{ for } p \in \Phi \\
Ms \vDash \neg\varphi & \text{iff} & Ms \nvDash \varphi \\
Ms \vDash \varphi \wedge \psi & \text{iff} & Ms \vDash \varphi \text{ and } Ms \vDash \psi \\
Ms \vDash \Box_i \varphi & \text{iff} & \text{for all } t, sR_i t \text{ implies } Mt \vDash \varphi
\end{array}
$$

Throughout, when referring to a modal language $\mathcal{L}$ alongside a sets of pointed Kripke models $X$, we tacitly assume that all models in $X$ share the signature of $\mathcal{L}$. As usual, modal logics may be formulated over the language $\mathcal{L}$. In this article, we only make use of *normal modal logics* $\Lambda$ over $\mathcal{L}$,[2] additionally assuming that every $\Lambda$-consistent formula $\varphi$ belongs to some maximally $\Lambda$-consistent set.

**4.2. Descriptors for Pointed Kripke Models.** As descriptors for pointed Kripke models, we use sets of $\mathcal{L}$-formulas. The contribution to the distance between two models given by disagreeing on the truth value of some formula $\varphi \in \mathcal{L}$ will simply be $w(\varphi)$. To avoid unnecessary double counting, we will

---

[2]A modal logic is **normal** if it contains all propositional tautologies and the $K$ axiom $\Box_i(\varphi \to \psi) \to (\Box_i \varphi \to \Box_i \psi)$, and is closed under *modus ponens*, *uniform substitution*, and *necessitation (generalization)*.

usually pick our descriptors such that they do not contain logically equivalent formulas.

DEFINITION. Let $\mathcal{L}$ be a modal language. A **descriptor** is any set $D \subseteq \mathcal{L}$.

The choice of descriptor determines which $\mathcal{L}$-formulas are given non-zero weight in the metric. Choosing, e.g,. the set of atomic propositions as descriptor will result in a rather coarse perspective. In most of this paper, we will occupy ourselves with descriptors that are not too coarse. We will be particularly interested in descriptors that are rich enough to reflect all $\mathcal{L}$-expressible differences between models from some set of interest:

DEFINITION. Let $X$ be a set of pointed Kripke models. The descriptor $D \subseteq \mathcal{L}$ is $\mathcal{L}$-**representative over** $X$ if, for every $\varphi \in \mathcal{L}$, there is a set $\{\psi_i\}_{i \in I} \subseteq D$ such that any valuation of $\{\psi_i\}_{i \in I}$ will semantically entail either $\varphi$ or $\neg\varphi$ over $X$. I.e.: For all $\varphi \in \mathcal{L}$ and all $S \subseteq I$, at most one of the sets $\{\psi_i : i \in I\} \cup \{\neg\psi_i : \psi_i \in I \setminus S\} \cup \{\varphi\}$ and $\{\psi_i : i \in I\} \cup \{\neg\psi_i : \psi_i \in I \setminus S\} \cup \{\neg\varphi\}$ has a model in $X$.

If the set $\{\psi_i\}_{i \in I}$ can always be chosen finite, we call $D$ **finitely $\mathcal{L}$-representative over** $X$. Finally, for a logic $\Lambda$, call $D$ $\Lambda$-**representative** if it is $\mathcal{L}$-representative over some space $X$ of pointed $\Lambda$-models in which every $\Lambda$-consistent set is satisfied in some $x \in X$.

The main implication of a descriptor being representative is that $\boldsymbol{X}_D$ is identical to $\boldsymbol{X}_{\mathcal{L}}$. This is stated formally in Lemma 3 below. When $D$ is $\Lambda$-representative, $D$ thus offers the most fine-grained perspective possible on any set of models $X$ for which $\Lambda$ is sound. Both representative and non-representative descriptors are exemplified in Section 4.5.

**4.3. Modal Spaces.** As stated in Section 3, we construct metrics on sets of structures *modulo* some modal equivalence. The choice to use a syntactic over a semantic quotient is motivated by general applicability: The notion of language equivalence of structures is conceptually uniform across all possible languages. Concepts of semantic equivalence, on the other hand, characterizing structural identity relative to the language in question, may be highly variable across frameworks.[3] In our parlance, we follow [42] in referring to *modal spaces*:

DEFINITION. With $X$ a set of pointed Kripke models and $D$ a descriptor, the **modal space $\boldsymbol{X}_D$** is the set $\{\boldsymbol{x}_D : x \in X\}$ with $\boldsymbol{x}_D = \{y \in X : \forall \varphi \in D, y \vDash \varphi \text{ iff } x \vDash \varphi\}$.

The subscript of $\boldsymbol{x}_D$ is omitted when the descriptor is clear from context. In this case boldface symbols $\boldsymbol{x}$ refer to members of $\boldsymbol{X}_D$ while lightface $x$ refers to members of $X$.

The choice of descriptor influences the resulting modal space: $\boldsymbol{X}_D$ may be a more or less coarse partition of $X$, with two extremes: If the descriptor is $\mathcal{L}$ itself (or, more economically, some $\mathcal{L}$-representative set $D$, to avoid double counting) the finest partition is achieved: $\boldsymbol{X}_{\mathcal{L}}$, the quotient of $X$ under $\mathcal{L}$-equivalence. For the coarsest partition, choose $\{\top\}$ as descriptor: $\boldsymbol{X}_{\{\top\}}$ is simply $\{\{X\}\}$.

---

[3]Compare e.g. isomorphism as an identity concept for first-order languages with bisimulation suited for standard modal languages and again with the many specialized versions of bisimulation suited to non-standard modal languages. See also Example 8.

We are mainly interested in modal spaces that retain the structure of $X$ as seen by a language $\mathcal{L}$, i.e., $\boldsymbol{X}_{\mathcal{L}}$. This does not entail that $\mathcal{L}$ is the only descriptor of interest. Others are sufficient:

LEMMA 3. *If $D \subseteq \mathcal{L}$ is a $\mathcal{L}$-representative descriptor for $X$, then $\boldsymbol{X}_D$ is identical to $\boldsymbol{X}_{\mathcal{L}}$, i.e., for all $x, y \in X$, $y \in \boldsymbol{x}_D$ iff $y \in \boldsymbol{x}_{\mathcal{L}}$.*

PROOF. We first show that $y \in \boldsymbol{x}_D$ entails $y \in \boldsymbol{x}_{\mathcal{L}}$. Assume $y \in \boldsymbol{x}_D$. To show that $y \in \boldsymbol{x}_{\mathcal{L}}$, we need to prove that for all $\varphi \in \mathcal{L}$ it holds that $x \vDash \varphi \Leftrightarrow y \vDash \varphi$. We only show the left-to-right implication, the other direction being similar. Assume $x \vDash \varphi$. Let $S = \{\psi \in D : x \vDash \psi\}$. By representativity, there is no $x' \in X$ satisfying $S \cup \{\neg\psi : \psi \in D \setminus S\} \cup \{\neg\varphi\}$. Since $y \in \boldsymbol{x}_D$ it satisfies $S \cup \{\neg\psi : \psi \in D \setminus S\}$ and hence also $\varphi$, i.e. $y \vDash \varphi$.

Next we show that that $y \in \boldsymbol{x}_{\mathcal{L}}$ entails $y \in \boldsymbol{x}_D$ . Assume $y \in \boldsymbol{x}_{\mathcal{L}} \in \boldsymbol{X}_{\mathcal{L}}$ It hence holds that $x \vDash \varphi \Leftrightarrow y \vDash \varphi$ for all $\varphi \in \mathcal{L}$. In particular, $x \vDash \varphi \Leftrightarrow y \vDash \varphi$ for all $\varphi$ with $\varphi \in D$ which implies that $y \in \boldsymbol{x}_D$. ⊣

REMARK. We do *not* generally assume a descriptor representative. When we do, we state so. For several of our results, the assumption is not necessary.

DEFINITION. Given the modal space $\boldsymbol{X}_D$, for a descriptor $D \subseteq \mathcal{L}$, the **truth set** of $\varphi \in \mathcal{L}$ in $\boldsymbol{X}_D$ is $[\varphi]_D = \{\boldsymbol{x} \in \boldsymbol{X}_D : \forall x \in \boldsymbol{x}, x \vDash \varphi\}$.

The subscript of $[\varphi]_D$ is again omitted when the descriptor is clear from context.

While the truth set is well-defined for any $\varphi \in \mathcal{L}$ and any $\boldsymbol{X}_D$ and we always have that $[\varphi]_D \cap [\neg\varphi]_D = \emptyset$, there are degenerate cases where $[\varphi]_D \cup [\neg\varphi]_D \neq \boldsymbol{X}_D$. This can only occur if $\varphi, \neg\varphi \notin D$ and if there are $x, y \in X$ with $\boldsymbol{x}_D = \boldsymbol{y}_D$ but $x \vDash \varphi$, $y \vDash \neg\varphi$. Note that when $\varphi \in D$ it always holds that $[\varphi]_D \cup [\neg\varphi]_D = \boldsymbol{X}_D$. Moreover, if $D$ is $\mathcal{L}$-representative over $X$, no degenerate cases occur: Then $[\varphi]_D \cup [\neg\varphi]_D = \boldsymbol{X}_D$ for all $\varphi \in \mathcal{L}$. We write $\boldsymbol{x}_D \vDash \varphi$ when $\boldsymbol{x}_D \in [\varphi]_D$.

**4.4. Metrics on Modal Spaces.** Finally, we obtain the family $\mathcal{D}_{(X,D)}$ of metrics on the $D$-modal space of a set of pointed Kripke models $X$:

PROPOSITION 4. *Let $D$ be a descriptor and $X$ a set of pointed Kripke models. Let $\nu : \boldsymbol{X}_D \times D \to \{0,1\}$ be a valuation given by $\nu(\boldsymbol{x}, \varphi) = 1$ iff $\boldsymbol{x} \in [\varphi]_D$. Let $w : D \to \mathbb{R}_{>0}$ be a weight function. Then $d_w$ as defined on page 5 is a metric on $\boldsymbol{X}_D$.*

PROOF. This follows immediately from Proposition 2 as $\nu$ is well-defined. ⊣

COROLLARY 5. *For any descriptor $D$, $\mathcal{D}_{(X,D)}$ is a family of metrics on $\boldsymbol{X}_D$.*

**4.5. Examples.** In constructing a metric $d_w \in \mathcal{D}_{(X,D)}$ for a modal space $\boldsymbol{X}_D$, two parameters must be fixed: The descriptor and the weight function. Jointly, these two parameters allow much freedom in picking a metric according to desired properties. In this section, we provide three classes of examples: One of non-representative descriptors, one of representative descriptors, and one of representative descriptors on finite sets. For the latter, we show Proposition 11 proving the metrics of Caridroit *et al.* on pointed Kripke models special cases of our approach.

**4.5.1.** *Non-Representative Descriptors.*

EXAMPLE 6 (Hamming Distance on Partial Atom Valuation). Let $\mathcal{L}$ be a modal language and $X$ a set of pointed Kripke models for $\mathcal{L}$. Let $p_1, p_2, ...$ be an enumeration of the atoms of $\mathcal{L}$. Pick as descriptor $D = \{p_1, ..., p_n\} \subseteq \mathcal{L}$ and weight function $w$ satisfying $w(p_k) = 1$ for all $p_k \in D$. Then $d_w$ is a metric on $\boldsymbol{X}_D$ cf. Proposition 4. If $X$ contains a model for each possible valuation of $p_1, ..., p_n$, then the metric space $(\boldsymbol{X}_D, d_w)$ is isometric to the metric space of strings of length $n$ under the Hamming distance. For $\boldsymbol{X}_D$, pointed Kripke models are compared only by their valuation of the first $n$ atoms. The space and the underlying metric reflects no modal structure.

If the set of atoms $\Phi$ of $\mathcal{L}$ is countably infinite, then we *cannot* assign all atoms equal weight: The sequence $(w'(p_n))_{n \in \mathbb{N}}$ would not give rise to a convergent series, so $w'$ is not a weight function. Partitioning $\Phi$ into cells $P_1, P_2, ...$ with each $P_k, k \in \mathbb{N}$, finite but arbitrarily large, and assigning $w''(p) = \frac{a_k}{|P_k|}$ for all $p \in P_k$ with $a_k$ the $k$th term of some convergent series does, however, give rise to a weight function.

EXAMPLE 7 (World Views and Situation Similarity). Consider an agent, $a$, who only cares about her beliefs about some atom $p$ and her beliefs about the beliefs of another agent, $b$, about the same. Working in a doxastic logic with operators $B_a$ and $B_b$, agent $a$'s set of interest may be described by $D = \{B_a\varphi, B_a\neg\varphi, \neg B_a\varphi \wedge \neg B_a\neg\varphi\}$ with $\varphi \in \{p, B_bp, B_b\neg p\}$. Similarities in situations (pointed Kripke models) from the viewpoint of $a$ may then be represented by using weight functions and their distances. E.g.: If $a$ cares equally much about nature and $b$'s beliefs thereof, every element of $D$ may be given equal weight; If she cares less about $b$'s beliefs, $D$ may be split in two, with formulas that contains $B_b$-operators given a strictly lower weight than those that do not.

**4.5.2.** *Representative Descriptors.*

EXAMPLE 8 (Degrees of Bisimilarity). Contrary to our syntactic approach to metric construction, a natural semantic approach rests on *bisimulations*. In particular, the notion of *n-bisimularity* may be used to define a semantically based metric on quotient spaces of pointed Kripke models where degrees of bisimilarity translate to closeness in space—the more bisimilar, the closer:

Let $X$ be a set of pointed Kripke models for which modal equivalence and bisimilarity coincide[4] and let $\leftrightarrow_n$ relate $x, y \in X$ iff $x$ and $y$ are $n$-bisimilar. Then

$$(1) \qquad d_B(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } x \leftrightarrow_n y \text{ for all } n \\ \frac{1}{n+1} & \text{if } n \text{ is the least intenger such that } x \not\leftrightarrow_n y \end{cases}$$

is a metric on $\boldsymbol{X}_{\mathcal{L}}$.[5] We refer to $d_B$ as the $n$-**bisimulation metric**.

For $X$ and $\mathcal{L}$ based on a finite signature, the $n$-bisimulation metric is a special case of the approach we suggest. I.e., for each set of pointed Kripke model $X$,

---

[4]That all models in $X$ are *image-finite* is a sufficient condition, cf. the Hennessy-Milner Theorem. See e.g. [20] or [38].

[5]The metric is inspired by [37], defining a distance between theories of first-order logic using quantifier depth. We return to this in Section 5.4. A further bisimulation based metric is the "$n$-Bisimulation-based Distance" of [23], which yields a pseudo-metric on sets of finite, pointed Kripke models, see also Example 10 below.

there exists a descriptor $D$ such that $d_B \in \mathcal{D}_{(X,D)}$. To see this, first recall that each model in $X$ has a **characteristic formula** up to $n$-bisimulation: For each $x \in X$, there exists a $\varphi_{x,n} \in \mathcal{L}$ such that for all $y \in X$, $y \vDash \varphi_{x,n}$ iff $x \leftrightarroweq_n y$, cf. [38,48]. Given that both $\Phi$ and $\mathcal{I}$ are finite, so is, for each $n$, the set $D_n = \{\varphi_{x,n} : x \in X\} \subseteq \mathcal{L}$. Pick the descriptor to be $D = \bigcup_{n \in \mathbb{N}} D_n$. Then $D$ is $\mathcal{L}$-representative for $X$, so $\boldsymbol{X}_D$ is identical to $\boldsymbol{X}_{\mathcal{L}}$, cf. Lemma 3.

Let the weight function $b$ be given by

$$b(\varphi) = \tfrac{1}{2}\left(\tfrac{1}{n+1} - \tfrac{1}{n+2}\right) \text{ for } \varphi \in D_n.$$

Then $d_b$, defined by $d_b(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=0}^{\infty} b(\varphi_k) d(\boldsymbol{x}, \boldsymbol{y})$, is a metric on $\boldsymbol{X}_{\mathcal{L}}$, cf. Proposition 4.

As models $x$ and $y$ will, for all $n$, either agree on all members of $D_n$ or disagree on exactly 2 (namely $\varphi_{n,x}$ and $\varphi_{n,y}$) and as, for all $k \le n$, $y \vDash \varphi_{n,x}$ implies $y \vDash \varphi_{k,x}$, and for all $k \ge n$, $y \nvDash \varphi_{n,x}$ implies $y \nvDash \varphi_{k,x}$, we obtain that

$$d_b(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } x \leftrightarroweq_n y \text{ for all } n \\ \sum_{k=n_0}^{\infty} 2 \cdot \tfrac{1}{2}\left(\tfrac{1}{k+1} - \tfrac{1}{k+2}\right) = \tfrac{1}{n_0+1} & n_0 = \min_{n \in \mathbb{N}}\{x \not\leftrightarroweq_n y\} \end{cases}$$

which is exactly $d_B$.

REMARK. We may pick the descriptor $D$ to be $K$-representative, where $K$ is the minimal normal modal logic: Let $X$ be some space of pointed $K$-models models in which every $K$-consistent set is satisfied in some $x \in X$ and follow the construction of $D$ above. Then $D$ is $K$-representative. Moreover, for every set of models $Y$, the map $f : \boldsymbol{Y}_D \to \boldsymbol{X}_D$ sending $\boldsymbol{y} \in \boldsymbol{Y}_D$ to the unique $\boldsymbol{x} \in \boldsymbol{X}_D$ with $\boldsymbol{y} \vDash \varphi$ iff $\boldsymbol{x} \vDash \varphi$ for all $\varphi \in D$ is injective and distance preserving and thus allows identifying $Y_D$ with a subset of $X_D$. Hence, the $d_b$ constructed on $X$ is exactly $d_B$ on any set $Y$ of $\mathcal{L}$-models.

REMARK. The construction given for encoding the $n$-bisimulation metric only works when the set of atoms and number of modalities are both finite: In case either is infinite, there is no metric in $\mathcal{D}_{(X,D)}$ for a descriptor $D \subseteq \mathcal{L}$ that is equivalent to the $n$-bisimulation metric, cf. Section 5.4.

EXAMPLE 9 (Close to Home, Close to Heart). The distances $d_B$ and $d_b$ from the previous example do not reflect all differences between models. For example, if two models are not $n$-bisimilar due only to atomic disagreement $n$ steps from the designated state, then it does not matter on how many atoms or how many worlds at distance $n$ they disagree: Their distance will be $\frac{1}{n+1}$ in all cases. Likewise, no differences they exhibit beyond the $n$th step will influence their distance: Only the first difference matters. I.e., the metric is extremal, in the terminology of Sokolsky *et al.*.

We may construct a metric which retains the feature of $d_b$ that differences further from the designated state weighs less than differences closer, but which, in an aggregate manner, assigns a positive weight to every modally expressible difference. In a slogan:

> *All and only modally expressible difference matters, but the further you have to go to find it, the less it matters.*

We might further this requirement by demanding that the new metric be a refinement of the bisimulation metric. That is, disagreement at distance $n$ from the designated state should weigh heavier than all disagreements at distance $n+1$ or beyond combined. On a set of models $X$ for language $\mathcal{L}$ with finite signature, a metric that lives up to the slogan may be defined as follows:

For descriptor, choose first for each equivalence class $\varphi_X := \{\psi \in \mathcal{L} \colon X \vDash \psi \leftrightarrow \varphi\}$ a **shallowest representative**, i.e., a formula in $\varphi_X$ which is of lowest *modal depth* among the members of $\varphi_X$. Denote the chosen shallowest representative of $\varphi_X$ by $\underline{\varphi}$ and let $D = \{\underline{\varphi} \colon \varphi \in \mathcal{L}\}$.

Let $\{D_n\}_{n \in \mathbb{N}}$ be a partition of $D$ by modal depth: For $n \in \mathbb{N}$, let $D_n$ contain the $\underline{\varphi}$ in $D$ that are of modal depth $n$. Since there are only finitely many equivalence classes of formulas of modal depth $n$, each $D_n$ is finite. Define a weight function $c$ by

$$c(\underline{\varphi}) = \frac{1}{|D_n|} \frac{1}{\prod_{k<n} |D_k|} \frac{1}{2^n} \text{ for } \underline{\varphi} \in D_n.$$

Then $d_c$ is a metric on $\boldsymbol{X}_{\mathcal{L}}$.

The first term ensures that disagreement on any formula in $D_n$ contributes $\frac{1}{\prod_{k<n} |D_k|} \frac{1}{2^n}$ to the distance between models. The second term ensures that the summed weight of all formulas in $D_j$ for $j > n$ is less than or equal to the weight of any $D_n$ formula, even when $|D_j| > |D_n|$. The third term ensures that the weights form a convergent series and that the inequality between disagreement levels is strict: One disagreement on a single formula of modal depth $n$ adds more to the distance between two models than do disagreement on all formulas of modal depth $n+1$ and above combined. Formally, for all $n$,

$$(2) \qquad \frac{1}{2^n} \frac{1}{|D_n|} \frac{1}{\prod_{k<n} |D_k|} > \sum_{m=n+1}^{\infty} \frac{1}{2^m} \frac{|D_m|}{|D_m|} \frac{1}{\prod_{k<n} |D_k|}.$$

Given this features, the metric $d_c$ captures both aspects the slogan:

1. Given that every $\mathcal{L}$-expressible property satisfiable on $X$ is given positive weight, and that only disagreement on expressible properties contribute to the distance between models, all and only modally expressible differences matter.

2. That further distance from the designated world should imply less importance of difference is captured as Equation (2) implies that for any $x, y, z \in X$, if $x$ and $y$ are not $n$-modally equivalent but $x$ and $z$ are, then $d_c(\boldsymbol{x}, \boldsymbol{y}) > d_c(\boldsymbol{x}, \boldsymbol{z})$.

**4.5.3.** *Metrics on Finite Sets: Relations to Caridroit* et al.

EXAMPLE 10 (Metrics on Finite Sets). As a last example, consider the case where $X$ and $\mathcal{L}$ are such that $\boldsymbol{X}_{\mathcal{L}}$ is of finite cardinality. This may happen when $X$ itself is finite, as is explicitly assumed in [23] where Caridroit *et al.* define their six distances between finite pointed KD45 Kripke models. It also happens in a language with a single operator and a finite set of atoms where the accessibility relation in models is an equivalence relation. In these settings, for *any* metric $d$ on $\boldsymbol{X}_{\mathcal{L}}$ there is a metric $d_w \in \mathcal{D}_{(X,D)}$ equivalent with $d$ up to linear shift. Somewhat weaker than strict numerical equivalence, this entails that the spaces $(\boldsymbol{X}_{\mathcal{L}}, d)$ and $(\boldsymbol{X}_{\mathcal{L}}, d_w)$ are *quasi-isometric* to each other.

PROPOSITION 11. *Let $(\boldsymbol{X}_{\mathcal{L}}, d)$ be a finite metric modal space. Then there exists a descriptor $D \subseteq \mathcal{L}$ finitely representative over $X$ , a metric $d_w \in \mathcal{D}_{(X,D)}$ and some $c \in \mathbb{R}$ such that $d_w(\boldsymbol{x}_D, \boldsymbol{y}_D) = d(\boldsymbol{x}_{\mathcal{L}}, \boldsymbol{y}_{\mathcal{L}}) + c$ for all $\boldsymbol{x} \neq \boldsymbol{y} \in \boldsymbol{X}_{\mathcal{L}}$. As a consequence, $(\boldsymbol{X}_D, d_w)$ are $(\boldsymbol{X}_{\mathcal{L}}, d)$ quasi-isometric to each other.*

PROOF. Since $\boldsymbol{X}_{\mathcal{L}}$ is finite, there is some $\varphi_{\boldsymbol{x}} \in \mathcal{L}$ for each $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$ such that for all $y \in X$, if $y \vDash \varphi_{\boldsymbol{x}}$, then $y \in \boldsymbol{x}$. Moreover, let $\varphi_{\{\boldsymbol{x},\boldsymbol{y}\}}$ denote the formula $\varphi_{\boldsymbol{x}} \vee \varphi_{\boldsymbol{y}}$ which holds true in $\boldsymbol{z} \in \boldsymbol{X}_{\mathcal{L}}$ iff $\boldsymbol{z} = \boldsymbol{x}$ or $\boldsymbol{z} = \boldsymbol{y}$. Let $D = \{\varphi_{\boldsymbol{x}} \colon \boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}\} \cup \{\varphi_{\{\boldsymbol{x},\boldsymbol{y}\}} \colon \boldsymbol{x} \neq \boldsymbol{y} \in \boldsymbol{X}_{\mathcal{L}}\}$. It follows that $\boldsymbol{X}_D = \boldsymbol{X}_{\mathcal{L}}$, hence $D$ is finitely representative over $X$.

Next, partition the finite set $\boldsymbol{X}_{\mathcal{L}} \times \boldsymbol{X}_{\mathcal{L}}$ according to the metric $d$: Let $S_1, ..., S_k$ be the unique partition of $\boldsymbol{X}_{\mathcal{L}} \times \boldsymbol{X}_{\mathcal{L}}$ that satisfies, for all $i, j \leq k$

1. If $(\boldsymbol{x}, \boldsymbol{x}') \in S_i$ and $(\boldsymbol{y}, \boldsymbol{y}') \in S_i$, then $d(\boldsymbol{x}, \boldsymbol{x}') = d(\boldsymbol{y}, \boldsymbol{y}')$, and
2. If $(\boldsymbol{x}, \boldsymbol{x}') \in S_i$ and $(\boldsymbol{y}, \boldsymbol{y}') \in S_j$ for $i < j$, then $d(\boldsymbol{x}, \boldsymbol{x}') < d(\boldsymbol{y}, \boldsymbol{y}')$.

For $i \leq k$, let $b_i$ denote $d(\boldsymbol{x}, \boldsymbol{y})$ for any $(\boldsymbol{x}, \boldsymbol{y}) \in S_i$. Define a weight function $w : D \to \mathbb{R}_{>0}$ by

$$w(\varphi_{\boldsymbol{x}}) = \sum_{i=1}^{k} \sum_{\substack{(\boldsymbol{y},\boldsymbol{z}) \in S_i \\ \boldsymbol{x} \neq \boldsymbol{y}, \boldsymbol{z}}} \frac{1 + b_k - b_i}{4}$$

$$w(\varphi_{\{\boldsymbol{x},\boldsymbol{y}\}}) = 2 \cdot \frac{1 + b_k - b_i}{4} \text{ for the } i \text{ with } (\boldsymbol{x}, \boldsymbol{y}) \in S_i$$

Note that by symmetry, $(\boldsymbol{x}, \boldsymbol{y}) \in S_i$ implies $(\boldsymbol{y}, \boldsymbol{x}) \in S_i$, thus $w(\varphi_{\{\boldsymbol{x},\boldsymbol{y}\}})$ is well-defined. We get for each $\boldsymbol{x}$ that

$$w(\varphi_{\boldsymbol{x}}) + \sum_{\boldsymbol{y} \neq \boldsymbol{x}} w(\varphi_{\{\boldsymbol{x},\boldsymbol{y}\}}) = \sum_{i=1}^{k} \sum_{\substack{(\boldsymbol{y},\boldsymbol{z}) \in S_i \\ \boldsymbol{x} \notin \{\boldsymbol{y},\boldsymbol{z}\}}} \frac{1 + b_k - b_i}{4} + \sum_{i=1}^{k} \sum_{\substack{(\boldsymbol{y},\boldsymbol{z}) \in S_i \\ \boldsymbol{x} \in \{\boldsymbol{y},\boldsymbol{z}\}}} \frac{1 + b_k - b_i}{4}$$

$$= \sum_{i=1}^{k} \sum_{(\boldsymbol{y},\boldsymbol{z}) \in S_i} \frac{1 + b_k - b_i}{4}$$

For simplicity, let $a$ denote $\sum_{i=1}^{k} \sum_{(\boldsymbol{y},\boldsymbol{z}) \in S_i} \frac{1 + b_k - b_i}{4}$, the rightmost term of the previous equation. Next, note that two models $\boldsymbol{x}$ and $\boldsymbol{y}$ differ on exactly the formulas $\varphi_{\boldsymbol{x}}, \varphi_{\boldsymbol{y}}$ and all $\varphi_{\{\boldsymbol{x},\boldsymbol{z}\}}$ and $\varphi_{\{\boldsymbol{y},\boldsymbol{z}\}}$ for $\boldsymbol{z} \neq \boldsymbol{x}, \boldsymbol{y}$. In particular,

$$d_w(\boldsymbol{x}, \boldsymbol{y}) = w(\varphi_{\boldsymbol{x}}) + w(\varphi_{\boldsymbol{y}}) + \sum_{\boldsymbol{z} \neq \boldsymbol{x}, \boldsymbol{y}} w(\varphi_{\{\boldsymbol{x},\boldsymbol{z}\}}) + \sum_{\boldsymbol{z} \neq \boldsymbol{x}, \boldsymbol{y}} w(\varphi_{\{\boldsymbol{y},\boldsymbol{z}\}})$$

$$= w(\varphi_{\boldsymbol{x}}) + w(\varphi_{\boldsymbol{y}}) + \sum_{\boldsymbol{z} \neq \boldsymbol{x}} w(\varphi_{\{\boldsymbol{x},\boldsymbol{z}\}}) + \sum_{\boldsymbol{z} \neq \boldsymbol{y}} w(\varphi_{\{\boldsymbol{y},\boldsymbol{z}\}}) - 2w(\varphi_{\{\boldsymbol{x},\boldsymbol{y}\}})$$

$$= 2a - 4 \cdot \frac{1 + b_k - b_i}{4}$$

$$= 2a + b_i - 1 - b_k$$

where $i$ is such that $\{\boldsymbol{x}, \boldsymbol{y}\} \in S_i$. Denoting $2a - 1 - b_k$ by $c$, we get that $d_w(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{x}, \boldsymbol{y}) + c$ whenever $\boldsymbol{x} \neq \boldsymbol{y}$.                                    ⊣

**§5. Topological Properties.** Given a set of pointed Kripke models $X$ and a descriptor $D \subseteq \mathcal{L}$, Proposition 4 states that for any weight function $w$, $d_w$ is a metric on the modal space $\boldsymbol{X}_D$, the quotient of $X$ under $D$-equivalence. Hence $(\boldsymbol{X}_D, d_w)$ is a *metric space*. Any such metric space induces a *topological space* $(\boldsymbol{X}_D, \mathcal{T}_w)$ with a basis consisting of the open $\epsilon$-balls of $(\boldsymbol{X}_D, d_w)$. That is, the basis of the $d_w$ **metric topology** $\mathcal{T}_w$ on $\boldsymbol{X}_D$ is $\{B_{d_w}(\boldsymbol{x}, \varepsilon) : \boldsymbol{x} \in \boldsymbol{X}_D, \epsilon > 0\}$ with $B_{d_w}(\boldsymbol{x}, \varepsilon) = \{\boldsymbol{y} \in \boldsymbol{X}_D : d_w(\boldsymbol{x}, \boldsymbol{y}) < \varepsilon\}$. In this section, we investigate the topological properties of such spaces.

REMARK. Our main interest lies with modal logics, hence the focus. However, many of the following results do not require specifics from modal logic or the nature of pointed Kripke models, but may, *mutatis mutandis*, be applied more generally. We do not elaborate on this point further.

**5.1. Stone-like Topologies.** In fixing a descriptor $D$ for $X$, one also fixes the family of metrics $\mathcal{D}_{(X,D)}$. The members of $\mathcal{D}_{(X,D)}$ vary in their numerical metrical properties as is illustrated by the different metrics given in Example 9. Topologically, however, all members of $\mathcal{D}_{(X,D)}$ are equivalent. To show this, we will work with the following generalization of the *Stone topology*:

DEFINITION. Let $D$ be a descriptor for $X$. Define the **Stone-like topology** on $\boldsymbol{X}_D$ to be the topology $\mathcal{T}_D$ given by the subbasis of all sets $\{\boldsymbol{x} \in \boldsymbol{X}_D : x \vDash \varphi\}$ and $\{\boldsymbol{x} \in \boldsymbol{X}_D : x \vDash \neg\varphi\}$ for $\varphi \in D$.

Note that, as $D$ need not be closed under conjunction, this subbasis is, in general, not a basis of the topology. When $D \subseteq \mathcal{L}$ is $\mathcal{L}$-representative over $X$, $\boldsymbol{X}_D$ is identical to $\boldsymbol{X}_\mathcal{L}$, and the Stone-like topology $\mathcal{T}_D$ on $\boldsymbol{X}_D$ is a coarsening of the **Stone topology** on $\boldsymbol{X}_\mathcal{L}$ given by the basis of sets $\{\boldsymbol{x} \in \boldsymbol{X}_\mathcal{L} : x \vDash \varphi\}$, $\varphi \in \mathcal{L}$. If $D$ is finitely $\mathcal{L}$-representative over $X$ then $\mathcal{T}_D$ is identical to the Stone topology on $\boldsymbol{X}_\mathcal{L}$.

We can now state the promised proposition:

PROPOSITION 12. *The metric topology $\mathcal{T}_w$ of any metric $d_w \in \mathcal{D}_{(X,D)}$ on $\boldsymbol{X}_D$ is the Stone-like topology $\mathcal{T}_D$.*

PROOF. We recall that for topologies $\mathcal{T}$ and $\mathcal{T}'$ on some set $X$, if $\mathcal{T}' \supseteq \mathcal{T}$, then $\mathcal{T}'$ is said to be **finer than** $\mathcal{T}$ and $\mathcal{T}$ to be **coarser than** $\mathcal{T}'$. This is the case iff for each $x \in X$ and each basis element $B \in \mathcal{T}$ with $x \in B$, there exists a $B' \in \mathcal{T}'$ with $x \in B' \subseteq B$, cf. [49, Lemma 13.3]. It hence suffices to show that for any $d_w \in \mathcal{D}_{(X,D)}$ the topology $\mathcal{T}_w$ is both coarser and finer than $\mathcal{T}_D$.

We start by showing that $\mathcal{T}_w$ is finer than $\mathcal{T}_D$: It suffices to show the claim for all elements of a subbasis of $\mathcal{T}_D$. Let $\boldsymbol{x} \in \boldsymbol{X}_D$ and let $B_D$ be a subbasis element of $\mathcal{T}_D$ which contains $\boldsymbol{x}$. Then $B_D$ is of the form $\{\boldsymbol{y} \in \boldsymbol{X}_D : y \vDash \varphi\}$ or $\{\boldsymbol{y} \in \boldsymbol{X}_D : y \vDash \neg\varphi\}$ for some $\varphi \in D$. Without loss of generality, we assume the former. As $\boldsymbol{x} \in B_D$, $\boldsymbol{x} \vDash \varphi$. This formula $\varphi$ is assigned a strictly positive weight $w(\varphi)$ in the metric $d_w$. The open ball $B(\boldsymbol{x}, w(\varphi))$ of radius $w(\varphi)$ around $\boldsymbol{x}$ is a basis element of $\mathcal{T}_w$ and contains $\boldsymbol{x}$. Moreover, $B(\boldsymbol{x}, w(\varphi)) \subseteq B_D$. To see this, assume $\boldsymbol{y} \in B(\boldsymbol{x}, w(\varphi))$, but $y \nvDash \varphi$. Then $d_w(\boldsymbol{x}, \boldsymbol{y}) \geq w(\varphi)$, which implies $\boldsymbol{y} \notin B(x, w(\varphi))$, contrary to our assumption. Hence $\mathcal{T}_w$ is finer than $\mathcal{T}_D$.

For the reverse direction, we show that $\mathcal{T}_D$ is finer than $\mathcal{T}_w$: Let $B$ be a basis element of $\mathcal{T}_w$ which contains $\boldsymbol{x}$. As $B$ is a basis element, it is of the form $B(\boldsymbol{y}, \delta)$ for some $\delta > 0$. Let $\epsilon = \delta - d_w(\boldsymbol{x}, \boldsymbol{y})$. Note that $\epsilon > 0$. Let $\varphi_1, \varphi_2, \dots$ be an enumeration of $D$. Since $\sum_{i=1}^{|D|} w(\varphi_i) < \infty$, there is some $n$ such that $\sum_{i=n}^{|D|} w(\varphi_i) < \epsilon$. For $j < n$, let $\chi_j = \varphi_i$ if $x \vDash \varphi_j$ and $\neg\chi_j$ otherwise and let $\chi = \bigwedge_{j<n} \chi_j$. By construction, all $z$ with $z \vDash \chi$ agree with $x$ on the truth values of $\varphi_1, \dots, \varphi_{n-1}$ and thus $d_w(\boldsymbol{x}, \boldsymbol{z}) < \epsilon$. By the triangular inequality, this implies $d_w(\boldsymbol{y}, \boldsymbol{z}) < \delta$ and hence $\{\boldsymbol{z} : z \vDash \varphi\} \subseteq B$. Furthermore, since $\mathcal{T}_D$ is generated by

a subbasis containing $\{\boldsymbol{x} \in \boldsymbol{X}_D : x \vDash \varphi\}$ and $\{\boldsymbol{x} \in \boldsymbol{X}_D : x \vDash \neg\varphi\}$ for $\varphi \in D$, we have $\{\boldsymbol{z} : z \vDash \varphi\} \in \mathcal{T}_D$ as desired.                                                                  $\dashv$

As for any set of models $X$ and any descriptor $D$ the set $\mathcal{D}_{(X,D)}$ is non-empty, we get:

COROLLARY 13. *Any Stone-like topology $\mathcal{T}_D$ on a space $\boldsymbol{X}_D$ is metrizable.*

**5.2. Stone Spaces.** The Stone topology is well-known, but typically defined on the set of ultrafilters of a Boolean algebra, which it turns into a **Stone space**: A *totally disconnected, compact, Hausdorff* topological space.[6] The first property goes hand-in-hand with an abundance of clopen sets, to which Section 5.3 is dedicated. Compactness has several implications used throughout: It plays a key role in determining which sets are clopen, and in Section 6, it is related to convergence. Finally, being Hausdorff, stating that any two points are topologically distinguishable, entails that the limits of convergent sequences are unique.

When equipping modal spaces with Stone-like topologies, Stone spaces often result. That the resulting topological spaces are Hausdorff follows as each Stone-like topology is metrizable, cf. the previous section. We show that the Stone-like topology is also totally disconnected and identify sufficient conditions for its compactness.

PROPOSITION 14. *For any descriptor $D$, the space $(\boldsymbol{X}_D, \mathcal{T}_D)$ is totally disconnected.*

PROOF. Let $\boldsymbol{x} \neq \boldsymbol{y} \in \boldsymbol{X}_D$. We must find open sets $U, V$ with $\boldsymbol{x} \in U$ and $\boldsymbol{y} \in Y$ such that $U \cap V = \emptyset$ and $U \cup V = \boldsymbol{X}_D$. Since $\boldsymbol{x} \neq \boldsymbol{y}$, there exists some $\varphi \in D$ such that $\boldsymbol{x} \vDash \varphi$ while $\boldsymbol{y} \nvDash \varphi$ or vice versa. The sets $A = \{\boldsymbol{z} \in \boldsymbol{X}_D : z \vDash \varphi\}$ and $\overline{A} = \{\boldsymbol{z}' \in \boldsymbol{X}_D : z \vDash \neg\varphi\}$ are both open in the Stone-like topology, $A \cap \overline{A} = \emptyset$, and $A \cup \overline{A} = \boldsymbol{X}_D$. As $\boldsymbol{x} \in A$ and $\boldsymbol{y} \in \overline{A}$ or vice versa, this shows that the space $(\boldsymbol{X}_D, \mathcal{T}_D)$ is totally disconnected.                                                  $\dashv$

The space $(\boldsymbol{X}_D, \mathcal{T}_D)$, $D \subseteq \mathcal{L}$, is moreover compact when two requirements are satisfied: First, there exists a logic $\Lambda$ sound with respect to $X$ which is (logically) **compact:** An arbitrary set $A \subseteq \mathcal{L}$ of formulas is $\Lambda$-consistent iff every finite subset of $A$ is. Many modal logics are compact, including every *basic modal logic*, cf. e.g. [15], but not all are: Examples of non-compact modal logics include logics with a *common knowledge operator* [30, 7.3] or with *Kleene star* as a *PDL constructor* [20, 4.8]. As second requirement, we must assume the set $X$ sufficiently rich in model diversity. In short, we require that every $\Lambda$-consistent subset of $D$ has a model in $X$.

DEFINITION. Let $D \subseteq \mathcal{L}$ be a descriptor and $\Lambda$ be sound with respect to $X$. Then $X$ is $\Lambda$-**saturated** with respect to $D$ if for all subsets $A, A' \subseteq D$ such that the set $B = A \cup \{\neg\varphi : \varphi \in A'\}$ is $\Lambda$-consistent, there exists a model $x$ in $X$ such that $x \vDash \psi$ for all $\psi \in B$. If $D$ is also $\mathcal{L}$-representative over $X$, then $X$ is $\Lambda$-**complete**.

---

[6]A topological space is **compact** if every open cover has a finite open subcover, **Hausdorff** if any two distinct points have disjoint open neighborhoods, and **totally disconnected** if any two distinct points have disjoint open neighborhoods that jointly cover the entire space.

For logical compactness, $\Lambda$-saturation is a sufficient richness conditions, cf. the proposition below. To ease notation, we will write $\overline{D}$ for $D \cup \{\neg\varphi : \varphi \in D\}$ in the remainder of this paper.

PROPOSITION 15. *If $\Lambda$ is compact and $X$ is $\Lambda$-saturated with respect to $D \subseteq \mathcal{L}$, then the space $(\boldsymbol{X}_D, \mathcal{T}_D)$ is compact.*

PROOF. Note that a basis of the topology $\mathcal{T}_D$ is given by the family of all sets $\{\boldsymbol{x} \in \boldsymbol{X}_D : \boldsymbol{x} \vDash \chi\}$ with $\chi$ of the form $\chi = \psi_1 \wedge \ldots \wedge \psi_n$ where $\psi_i \in \overline{D}$ for all $i \leq n$. To show that $(\boldsymbol{X}_D, \mathcal{T}_D)$ is compact, it suffices to show that every open cover consisting of basic open sets has a finite subcover. Suppose for a contradiction that $\{\{\boldsymbol{x} \in \boldsymbol{X}_D : \boldsymbol{x} \vDash \chi_i\} : i \in I\}$ is a cover of $\boldsymbol{X}_D$ that contains no finite subcover. This implies that every finite subset of $\{\neg\chi_i : i \in I\}$ is satisfied in some $\boldsymbol{x} \in \boldsymbol{X}_D$ and hence consistent, i.e., the set $\{\neg\chi_i : i \in I\}$ is finitely $\Lambda$-consistent. By compactness of $\Lambda$, the set $\{\neg\chi_i : i \in I\}$ is thus also $\Lambda$-consistent. Hence, by saturation, there is an $\boldsymbol{x} \in \boldsymbol{X}_D$ such that $\boldsymbol{x} \vDash \neg\chi_i$ for all $i \in I$. But then $\boldsymbol{x}$ cannot be in $\{\boldsymbol{x} \in \boldsymbol{X}_D : x \vDash \chi_i\}$ for any $i \in I$, contradicting the assumption that $\{\{\boldsymbol{x} \in \boldsymbol{X}_D : \boldsymbol{x} \vDash \chi_i\} : i \in I\}$ is a cover of $\boldsymbol{X}$.               $\dashv$

Propositions 14 and 15 jointly yield the following:

COROLLARY 16. *Let $\Lambda \subseteq \mathcal{L}$ be a compact modal logic sound and complete with respect to the class of pointed Kripke models $\mathcal{C}$. Then $(\mathcal{C}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ is a Stone space.*

PROOF. The statement follows immediately the propositions of this section when $\mathcal{C}_{\mathcal{L}}$ is ensured to be a set using *Scott's trick* [57].               $\dashv$

REMARK. When $D$ is $\mathcal{L}$-representative for $X$ and $\boldsymbol{X}_D$ is $\Lambda$-saturated, one obtains a very natural space. Such a space contains a (unique) point satisfying each maximal $\Lambda$-consistent set of formulas. It is thus homeomorphic to the space of all complete $\Lambda$-theories under the Stone topology of $\mathcal{L}$. Such spaces have been widely studied, see e.g. [37,59]. When we call such spaces $\Lambda$-complete, it reflects that the joint requirement ensures that the logic $\Lambda$ is complete with respect to the set $X$, but that the obligation of sufficiency lies on the set $X$ to be inclusive enough for $\Lambda$, not on $\Lambda$ to be restrictive enough for $X$.

**5.2.1.** *Compact Subspaces.* As the intersection of an arbitrary family of closed sets is itself a closed set and as every closed subspace of a compact space is compact [49, Theorems 17.1 and 26.2], we obtain the following, making use of the fact that $\{\boldsymbol{y} \in \boldsymbol{X}_D : \boldsymbol{y} \vDash \varphi\} = \boldsymbol{X}_D \backslash \{\boldsymbol{y} \in \boldsymbol{X}_D : \boldsymbol{y} \vDash \neg\varphi\}$ is closed for any $\varphi \in D$.

COROLLARY 17. *Let $A \subseteq D$ and let $Y = \{\boldsymbol{y} \in \boldsymbol{X_D} : y \vDash \varphi \text{ for all } \varphi \in A\}$. If $(\boldsymbol{X}_D, \mathcal{T}_D)$ is compact, then $\boldsymbol{Y}_D$ is compact under the subspace topology.*

Moreover, the subspace topology on $\boldsymbol{Y}_D$ induced by the Stone-like topology on $\boldsymbol{X}_D$ is again the Stone-like topology $\mathcal{T}'_D$ on $\boldsymbol{Y}_D$.

**5.3. Open, Closed and Clopen Sets in Stone-like Topologies.** In this section, we characterize the open, closed and clopen sets of Stone-like topologies relative to the set of formulas $\mathcal{L}$. With this, we hope to paint a picture of the structure of Stone-like topologies, helpful in understanding closed subspaces and limit points.

By definition, the Stone-like topology $\mathcal{T}_D$ is generated by the subbasis $\mathcal{S}_D = \{[\varphi]_D, [\neg\varphi]_D : \varphi \in D\}$. All subbasis elements are clearly clopen: If $U$ is of the form $[\varphi]_D$ for some $\varphi \in D$, then the complement of $U$ is the set $[\neg\varphi]_D$, which again is a subbasis element. Hence both $[\varphi]_D$ and $[\neg\varphi]_D$ are clopen. For the converse, we introduce the following:

DEFINITION. The Stone-like topology $\mathcal{T}_D$, $D \subseteq \mathcal{L}$, on the modal space $\boldsymbol{X}_D$ **reflects** the language $\mathcal{L}$ if for every set $Y \subseteq \boldsymbol{X}_D$, $Y$ is clopen in $\mathcal{T}_D$ iff $Y = [\varphi]_D$ for some $\varphi \in \mathcal{L}$.

We obtain the following:

PROPOSITION 18.   *Let $\Lambda$ be a logic sound with respect to the set of pointed Kripke models $X$. If $\Lambda$ is compact and $D$ is $\Lambda$-representative, then $[\varphi]_D$ is clopen in $\mathcal{T}_D$ for every $\varphi \in \mathcal{L}$. If $(\boldsymbol{X}_D, \mathcal{T}_D)$ is also topologically compact, then $\mathcal{T}_D$ reflects $\mathcal{L}$.*

PROOF. To show that under the assumptions, $[\varphi]_D$ is clopen in $\mathcal{T}_D$, for every $\varphi \in \mathcal{L}$, we first show the claim for the special case where $X$ is such that every $\Lambda$-consistent set $\Sigma$ is satisfied in some $x \in X$. By Proposition 12, it suffices to show that $\{\boldsymbol{x} \in \boldsymbol{X}_D : x \vDash \varphi\}$ is open for $\varphi \in \mathcal{L}\backslash D$. Fix such $\varphi$. As $D$ is $\Lambda$-representative, $\boldsymbol{X}_D$ is identical to $\boldsymbol{X}_{\mathcal{L}}$, cf. Lemma 3. Hence $[\varphi] := \{\boldsymbol{x} \in X_D : x \vDash \varphi\}$ is well-defined. To see that $[\varphi]$ is open, pick $\boldsymbol{x} \in [\varphi]$ arbitrarily. We find an open set $U$ with $\boldsymbol{x} \in U \subseteq [\varphi]$: Let $D_x = \{\psi \in \overline{D} : x \vDash \psi\}$. As witnessed by $x$, the set $D_x \cup \{\varphi\}$ is $\Lambda$-consistent. As $D$ is $\Lambda$-representative, $D_x$ thus semantically entails $\varphi$ over $X$. Hence, no model $y \in X$ satisfies $D_x \cup \{\neg\varphi\}$. By the choice of $X$, $\boldsymbol{X}_D$ is $\Lambda$-saturated with respect to $D$. This implies that the set $D_x \cup \{\neg\varphi\}$ is $\Lambda$-inconsistent. By the compactness of $\Lambda$, a finite subset $F$ of $D_x \cup \{\neg\varphi\}$ is inconsistent. W.l.o.g. we can assume that $\neg\varphi \in F$. Inconsistency of $F$ implies that $\psi_1 \wedge \ldots \wedge \psi_n \to \varphi$ is a theorem of $\Lambda$. On the semantic level, this translates to $\bigcap_{i \leq n}[\psi_i] \subseteq [\varphi]$. As each $[\psi_i]$ is open, $\bigcap_{i \leq n}[\psi_i]$ is an open neighborhood of $\boldsymbol{x}$ contained in $[\varphi]$.
Next, we prove the general case. Let $X$ be any set of $\Lambda$-models and let $Y$ be such that every $\Lambda$-consistent set $\Sigma$ is satisfied in some $y \in Y$. Then the function $f : \boldsymbol{X}_D \to \boldsymbol{Y}_D$ that sends $\boldsymbol{x} \in \boldsymbol{X}_D$ to the unique $\boldsymbol{x} \in \boldsymbol{Y}_D$ with $x \vDash \varphi \Leftrightarrow y \vDash \varphi$ for all $\varphi \in \mathcal{L}$ is a continuous map from $(\boldsymbol{X}_D, \mathcal{T}_D)$ to $(\boldsymbol{Y}_D, \mathcal{T}_D)$ with $f^{-1}(\{\boldsymbol{y} \in \boldsymbol{Y}_D : y \vDash \varphi\}) = \{\boldsymbol{x} \in \boldsymbol{X}_D : \boldsymbol{x} \vDash \varphi\}$. By the first part, $\{\boldsymbol{y} \in \boldsymbol{Y}_D : y \vDash \varphi\}$ is clopen in $\boldsymbol{Y}_D$. As the continuous pre-image of clopen sets is clopen, this shows that $\{\boldsymbol{x} \in \boldsymbol{X}_D : \boldsymbol{x} \vDash \varphi\}$ is clopen.
We turn to the second claim, that if $(\boldsymbol{X}_D, \mathcal{T}_D)$ is also topologically compact, then $\mathcal{T}_D$ reflects $\mathcal{L}$. It suffices to show that if $O \subseteq \boldsymbol{X}_D$ is clopen, then $O$ is of the form $[\varphi]_D$ for some $\varphi \in \mathcal{L}$. So assume $O$ is clopen. As $O$ and its complement $\overline{O}$ are open, there are formulas $\psi_i, \chi_i \in D$ for $i \in \mathbb{N}$ such that $O = \bigcup_{i \in \mathbb{N}}[\psi_i]_D$ and $\overline{O} = \bigcup_{i \in \mathbb{N}}[\chi_i]_D$. Hence $\{[\varphi_i]_D : i \in \mathbb{N}\} \cup \{[\psi_i]_D : i \in \mathbb{N}\}$ is an open cover of $\boldsymbol{X}_D$. By topological compactness, it contains a finite subcover. I.e., there are $I_1, I_2 \subset \mathbb{N}$ finite such that $\boldsymbol{X}_D = \bigcup_{i \in I_1}[\psi_i]_D \cup \bigcup_{i \in I_2}[\psi_i]_D$. In particular, $O = \bigcup_{i \in I_1}[\psi_i]_D = [\bigvee_{i \in I_1}\psi_i]_D$ which is what we had to show.    ⊣

By Proposition 15, two immediate consequences are:

COROLLARY 19. *Let $\Lambda$ be sound with respect to the set of pointed Kripke models $X$. If $\Lambda$ is compact, $D$ is $\Lambda$-representative and $X$ is $\Lambda$-saturated with respect to $D$, then $\mathcal{T}_D$ reflects $\mathcal{L}$.*

COROLLARY 20. *Let $\Lambda \subseteq \mathcal{L}$ be a compact modal logic sound and complete with respect to some class of pointed Kripke models $\mathcal{C}$. Then $\mathcal{T}_\mathcal{L}$ reflects $\mathcal{L}$.*

Compactness is essential to Proposition 18's characterization of clopen sets. Without the assumption of compactness, the clopen sets of Stone topologies do not reflect the underlying logic:

PROPOSITION 21. *Let $\boldsymbol{X}_D$ be $\Lambda$-saturated with respect to $D$ and $D$ be $\Lambda$-representative, but $\Lambda$ not compact. Then there exists a set $U$ clopen in $\mathcal{T}_D$ that is not of the form $[\varphi]_D$ for any $\varphi \in \mathcal{L}$.*

PROOF. As $\Lambda$ is not compact, there exists a $\Lambda$-inconsistent set of formulas $S = \{\chi_i : i \in \mathbb{N}\}$ for which every finite subset is $\Lambda$-consistent. For simplicity of notation, define $\varphi_i := \neg\chi_i$. As $\boldsymbol{X}_D$ is $\Lambda$-saturated with respect to $D$, $\{[\varphi_i]\}_{i\in\mathbb{N}}$ is an open cover of $\boldsymbol{X}_D$ that does not contain a finite subcover. For $i \in \mathbb{N}$ let $\rho_i$ be the formula $\varphi_i \wedge \bigwedge_{k<i} \neg\varphi_k$. In particular, we have that *i*) $[\rho_i] \cap [\rho_j] = \emptyset$ for all $i \neq j$ and *ii*) $\bigcup_{i\in\mathbb{N}}[\rho_i] = \bigcup_{i\in\mathbb{N}}[\varphi_i] = \boldsymbol{X}_D$, i.e., $\{[\rho_i]\}_{i\in\mathbb{N}}$ is a disjoint cover of $\boldsymbol{X}_D$. We further have that $[\rho_i] \subseteq [\varphi_i]$; hence $\{[\rho_i]\}_{i\in\mathbb{N}}$ cannot contain a finite subcover $\{[\rho_i]\}_{i\in I}$ of $\boldsymbol{X}_D$, as the corresponding $\{[\varphi_i]\}_{i\in I}$ would form a finite cover. In particular, infinitely many $[\rho_i]$ are non-empty. Without loss of generality, assume that all $[\rho_i]$ are non-empty. For all $S \subseteq \mathbb{N}$, the set $U_S = \bigcup_{i\in S}[\rho_i]$ is open. As all $[\rho_i]$ are mutually disjoint, the complement of $U_S$ is $\bigcup_{i\notin S}[\rho_i]$ which is also open; hence $U_S$ is clopen. Using again that all $[\rho_i]$ are mutually disjoint and non-empty, we have that $U_S \neq U_{S'}$ whenever $S \neq S'$. Hence, $\{U_S : S \subseteq \mathbb{N}\}$ is an uncountable family of clopen sets. As $\mathcal{L}$ is countable, there must be some element of $\{U_S : S \subseteq \mathbb{N}\}$ which is not of the form $[\varphi]$ for any $\varphi \in \mathcal{L}$.            $\dashv$

Working with a logical language thus imposes a natural structure through Stone-like topologies. This is especially so for suitable descriptors in the compact setting. In that case, the clopen sets are the ones characterizable as truth sets of a logical formula; the open, but not closed, sets are those characterizable by only by infinitary disjunction; and the closed, but not open, sets are those characterizable only by infinitary conjunction.

**5.4. Relations to the $n$-Bisimulation Topology.** In Example 8, we showed that $\mathcal{D}_{(X,\mathcal{L})}$ includes the semantically based $n$-bisimulation metric $d_B$ for modal languages with finite signature. The metric topology induced by the $n$-bisimulation metric is referred to as the $n$-**bisimulation topology**, $\mathcal{T}_B$. A basis for this topology is given by all subsets of $\boldsymbol{X}_\mathcal{L}$ of the form

$$B_{\boldsymbol{x}n} = \{\boldsymbol{y} \in \boldsymbol{X}_\mathcal{L} : y \leftrightarroweq_n x\}$$

for $\boldsymbol{x} \in \boldsymbol{X}_\mathcal{L}$ and $n \in \mathbb{N}$.

By Proposition 12, Example 8 and the fact that the set $D$ constructed in the latter is finitely $\mathcal{L}$-representative over $X$, we obtain the following:

COROLLARY 22. *If $\mathcal{L}$ has finite signature, then the $n$-bisimulation topology $\mathcal{T}_B$ is the Stone(-like) topology $\mathcal{T}_\mathcal{L}$.*

This is not the case in general:

PROPOSITION 23. *If $\mathcal{L}$ is based on an infinite set of atoms, then the $n$-bisimulation topology $\mathcal{T}_B$ is strictly finer than the Stone(-like) topology $\mathcal{T}_{\mathcal{L}}$ on $\boldsymbol{X}_{\mathcal{L}}$.*

PROOF. To see that the Stone(-like) topology is not as fine as the $n$-bisimulation topology, consider the basis element $B_{\boldsymbol{x}0}$ of the latter, containing exactly those elements $\boldsymbol{y}$ such that $y$ and $x$ are 0-bisimilar, i.e., share the same atomic valuation. Clearly, $\boldsymbol{x} \in B_{\boldsymbol{x}0}$. There is no formula $\varphi$ for which the Stone basis element $B = \{\boldsymbol{z} \in \boldsymbol{X} : z \vDash \varphi\}$ contains $\boldsymbol{x}$ and is contained in $B_{\boldsymbol{x}0}$: This would require that $\varphi$ implied every atom or its negation, requiring the strength of an infinitary conjunction.

For the inclusion of the Stone(-like) topology in the $n$-bisimulation topology, consider any $\varphi \in \mathcal{L}$ and the corresponding Stone basis element $B = \{\boldsymbol{y} \in \boldsymbol{X} : y \vDash \varphi\}$. Assume $\boldsymbol{x} \in B$. Let the modal depth of $\varphi$ be $n$. Then for every $\boldsymbol{z} \in B_{\boldsymbol{x}n}$, $z \vDash \varphi$. Hence $\boldsymbol{x} \in B_{\boldsymbol{x}n} \subseteq B$.                    ⊣

The discrepancy in induced topologies results as the $n$-bisimulation metric, in the infinite case, introduces distinctions that are not finitely expressible in the language: If there are infinitely many atoms, there does not exist a characteristic formula $\varphi_{x,n}$ satisfied only by models $n$-bisimilar with $x$.

**Non-compactness.** Even if $\boldsymbol{X}_{\mathcal{L}}$ is compact in the Stone(-like) topology, it need not be compact in the $n$-bisimulation topology: Let $\mathcal{L}$ be based on an infinite set of atoms $\Phi$ and $X$ a set of pointed models $\Lambda$-saturated with respect to $\mathcal{L}$ for some compact logic $\Lambda$. Then $\boldsymbol{X}_{\mathcal{L}}$ is compact in the Stone(-like) topology. It is not compact in the $n$-bisimulation topology: $\{B_{\boldsymbol{x}0} : x \in X\}$ is an open cover of $\boldsymbol{X}_{\mathcal{L}}$ which contains no finite subcover.

**Relations to Goranko's metric.** Corollary 22 and Proposition 23 jointly relate our metrics to the metric introduced by Goranko in [37] on first-order theories. The straight-forward alteration of that metric to suit a modal space $\boldsymbol{X}_{\mathcal{L}}$ is

$$d_g(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{y} \\ \frac{1}{n+1} & \text{if } n \text{ is the least intenger such that } n(\boldsymbol{x}) \neq n(\boldsymbol{y}) \end{cases}$$

where $n(\boldsymbol{x})$ is the set of formulas of modal depth $n$ satisfied by $\boldsymbol{x}$.

The induced topology of this metric is exactly the $n$-bisimulation topology. Hence, for languages with finite signature, every metric in the family $\mathcal{D}_{(X,\mathcal{L})}$ induces the same topology as $d_g$, but the induced topologies differ on languages with infinitely many atoms.

Goranko notes that his topological approach to prove relative completeness may, with a bit of work, be applied in a modal logical setting.[7] Replacing, in our approach, the modal space $\boldsymbol{X}_{\mathcal{L}}$ with the quotient space of $X$ under bisimulation would, we venture to claim, supply the stepping stone. We omit a detour into the details in favor of working with Stone-like topologies.

**§6. Convergence and Limit Points.** We next turn to dynamic aspects of Stone(-like) topologies. In particular, we focus on the nature of *convergent*

---

[7]See §6, especially the final paragraph.

*sequences* in Stone(-like) topologies and such topologies' *isolated points*, i.e., points no sequence can approach gradually.

**6.1. Convergence.** Recall that, with $(X, \mathcal{T})$ a topological space, a sequence of points $x_1, x_2, ...$ from $X$ is said to **converge** to the point $x \in X$ in the topology $\mathcal{T}$ iff for every open set $U \in \mathcal{T}$ containing $x$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \in U$. We denote a sequence $x_1, x_2, ...$ by $(x_n)_{n \in \mathbb{N}}$, but often omit the outer subscript.

For a Stone-like topology $\mathcal{T}_D$ on a modal space $\boldsymbol{X}_D$, this standard notion of convergence captures the geometrical intuition of a sequence gradually approaching a unique final destination:

PROPOSITION 24. *Let $(\boldsymbol{X}_D, \mathcal{T}_D)$ be a modal space with its Stone-like topology. Then any sequence of points $(\boldsymbol{x}_n)$ of $\boldsymbol{X}_D$ converges to at most one point.*

PROOF. As every Stone-like topology is metrizable, cf. Proposition 12, it is also Hausdorff. This implies the desired, cf. e.g. [49, Thm. 17.10]. ⊣

If the sequence $(\boldsymbol{x}_n)$ converges to $\boldsymbol{x}$, we are thus justified in writing $(\boldsymbol{x}_n) \to \boldsymbol{x}$ and referring to $\boldsymbol{x}$ as the **limit** of $(\boldsymbol{x}_n)$. In general, it is not ensured that a limit exists, a theme we return to in discussing compactness and the common knowledge operator below.

Convergence in Stone-like topologies also satisfies a natural logical intuition, namely that a sequence and its limit should eventually agree on every formula of the language used to describe them. This intuition is captured by the notion of *logical convergence*, introduced in [42]. For a modal space $\boldsymbol{X}_D$, it is defined as follows: A sequence of points $\boldsymbol{x}_1, \boldsymbol{x}_2, ...$ **logically converges** to $\boldsymbol{x}$ in $\boldsymbol{X}_D$ iff for every $\psi \in \{\varphi, \neg\varphi \colon \varphi \in D\}$ for which $x \vDash \psi$, there exists some $N \in \mathbb{N}$ such that $x_n \vDash \psi$, for all $n \geq N$.

The following proposition identifies a tight relationship between (topological) convergence, Stone-like topologies and logical convergence:

PROPOSITION 25. *Let $\boldsymbol{X}_D$ be a modal space and $\mathcal{T}$ a topology on $\boldsymbol{X}_D$. Then the following are equivalent:*

1. *A sequence $\boldsymbol{x}_1, \boldsymbol{x}_2, ...$ of points from $\boldsymbol{X}_D$ converges to $\boldsymbol{x}$ in $(\boldsymbol{X}_D, \mathcal{T})$ if, and only if, $\boldsymbol{x}_1, \boldsymbol{x}_2, ...$ logically converges to $\boldsymbol{x}$ in $\boldsymbol{X}_D$.*
2. *$\mathcal{T}$ is the Stone-like topology $\mathcal{T}_D$ on $\boldsymbol{X}_D$.*

PROOF. $2 \Rightarrow 1$: This is shown, *mutatis mutandis*, in [42, Prop. 2].
$1 \Rightarrow 2$: We first show that under the assumption of 1., the topology $\mathcal{T}$ contains $\mathcal{T}_D$, by showing that $\mathcal{T}$ contains the subbasis of $\mathcal{T}_D$: I.e., for all $\varphi \in D$, $[\varphi], [\neg\varphi] \in \mathcal{T}$. We only show the claim for $[\varphi]$, the proof for $[\neg\varphi]$ being equivalent. We show $[\varphi]$ open in $\mathcal{T}$ by showing it compliment, $[\neg\varphi]$, closed in $\mathcal{T}$, which is done by showing that $[\neg\varphi]$ contains all its limit points: Assume the sequence $(\boldsymbol{x}_i) \subseteq [\neg\varphi]$ converges to $\boldsymbol{x}$ in $(\boldsymbol{X}_D, \mathcal{T})$. For each $i \in \mathbb{N}$, we have $\boldsymbol{x}_i \vDash \neg\varphi$. As convergence is assumed to imply logical convergence, then also $\boldsymbol{x} \vDash \neg\varphi$. Hence $\boldsymbol{x} \in [\neg\varphi]$, so $[\neg\varphi]$ is closed in $\mathcal{T}$. Hence $\mathcal{T}_D \subseteq \mathcal{T}$.

Secondly, we show the reverse inclusion: That $\mathcal{T}_D$ contains $\mathcal{T}$. This is done by showing that for every element $\boldsymbol{x}$ of any open set $U$ of $\mathcal{T}$ there is a basis element $B$ of $\mathcal{T}_D$ such that $\boldsymbol{x} \in B \subseteq U$. Let $U \in \mathcal{T}$ and let $\boldsymbol{x} \in U$. Enumerate the set $\{\psi \in \overline{D} \colon x \vDash \psi\}$ as $\psi_1, \psi_2, \ldots$, and consider all conjunctions of finite prefixes $\psi_1$,

$\psi_1 \wedge \psi_2$, $\psi_1 \wedge \psi_2 \wedge \psi_3, \ldots$ of this enumeration. If for some $k$, $[\psi_1 \wedge \cdots \wedge \psi_k] \subseteq U$, then $B = [\psi_1 \wedge \cdots \wedge \psi_k]$ is the desired $\mathcal{T}_D$ basis element as $\boldsymbol{x} \in [\psi_1 \wedge \cdots \wedge \psi_k] \subseteq U$. If there exists no $k \in \mathbb{N}$ such that $[\psi_1 \wedge \cdots \wedge \psi_k] \subseteq U$, then for each $m \in \mathbb{N}$, we can pick an $\boldsymbol{x}_m$ such that $\boldsymbol{x}_m \in [\psi_1 \wedge \cdots \wedge \psi_m] \setminus U$. The sequence $(\boldsymbol{x}_m)_{m \in \mathbb{N}}$ then logically converges to $\boldsymbol{x}$. Hence, by assumption, it also converges topologically to $\boldsymbol{x}$ in $\mathcal{T}$. Now, for each $m \in \mathbb{N}$, $\boldsymbol{x}_m$ is in $U^c$, the compliment of $U$. However, $\boldsymbol{x} \notin U^c$. Hence, $U^c$ is not closed in $\mathcal{T}$, so $U$ in not open in $\mathcal{T}$. This contradicts our assumption, rendering the case that there is no $k \in \mathbb{N}$ such that $[\psi_1 \wedge \cdots \wedge \psi_k] \subseteq U$ impossible. Hence $\mathcal{T}_D \subseteq \mathcal{T}$.                                                        $\dashv$

In [42], the satisfaction of point 1 was used as motivation for working with Stone-like topologies. Proposition 25 shows that this choice of topology was necessary, if one wants the logical intuition satisfied. Moreover, it provides a third way of inducing Stone-like topologies, different from inducing them from a metric or a basis, namely through *sequential convergence*.[8]

From the properties of Stone-like topologies, we immediately obtain general results concerning convergence. Proposition 26 provides a selection of these. As its proof shows, these properties are neither novel nor specific to Stone-like topologies.

Recall that a for a topological space $(X, \mathcal{T})$, a point $x$ in $X$ is a **limit point** of the set $A \subseteq X$ iff for every open neighborhood $U \in \mathcal{T}$ containing $x$, there is some point $y \in U \setminus \{x\}$ contained in $A$. In contrast, a point $x$ is an **isolated** point of $A$ if there exists an open neighborhood $U$ of $x$ that contains no (other) points from $A$. Hence, $x$ is isolated in $X$ iff $\{x\}$ is open in $\mathcal{T}$. Lastly, we call $x$ in $X$ a **limit point** of the sequence $(\boldsymbol{x}_n)_{n \in \mathbb{N}}$ if for each open neighborhood $U$ the set $\{n \in \mathbb{N} : x_n \in U\}$ is infinite.

PROPOSITION 26. *Let $(\boldsymbol{X}_D, \mathcal{T}_D)$ be a modal space with it's Stone-like topology. Then:*

1. *A point in $(\boldsymbol{X}_D, \mathcal{T}_D)$ is a limit point of $A \subseteq \boldsymbol{X}_D$ iff it is not isolated in $A$.*
2. *A sequence $(\boldsymbol{x}_n)$ in $A \subseteq \boldsymbol{X}_D$ converges to an isolated point $\boldsymbol{x}$ in $A \subseteq \boldsymbol{X}_D$ iff for some $N$, for all $k > N$, $\boldsymbol{x}_k = \boldsymbol{x}$.*
3. *A point $\boldsymbol{x}$ is in the closure of $A \subseteq \boldsymbol{X}_D$ iff there exists a sequence $(\boldsymbol{x}_n)$ in $A$ which converges to $\boldsymbol{x}$.*
4. *A point $\boldsymbol{x} \in \boldsymbol{X}_D$ is a limit point of the sequence $(\boldsymbol{x}_n)$ iff there exists a subsequence of $(\boldsymbol{x}_n)$ converging to $\boldsymbol{x}$.*
5. *Let $f : \boldsymbol{X}_D \to \boldsymbol{X}_D$. Then $f$ is continuous iff if $\boldsymbol{x}_n \to \boldsymbol{x}$, then $f(\boldsymbol{x}_n) \to f(\boldsymbol{x})$.*
6. *$\mathcal{T}_D$ is compact iff every sequence in $\boldsymbol{X}_D$ has a convergent subsequence iff every infinite subset $A \subseteq \boldsymbol{X}_D$ has a limit point.*

PROOF. We omit the proof for the simple points 1, 2 and 4. For 3, see [49, Lemma 21.2]; for 5, see [49, Theorem 21.3]; and for 6, see [49, Theorem 28.2].    $\dashv$

It is beyond the scope of this paper to seek further conditions for convergence. However, motivated by Proposition 26.2, we present results on the existence of isolated points. In particular, Proposition 26.2 may be of interest in information dynamics: If, e.g., the goal of a given protocol holds only at isolated points, then

---

[8]For more on this approach to topologies, see the historical overview in [35].

the protocol will either be successful in finite time or not at all. This is related to common knowledge in Section 6.2.2 below.

**6.2. Isolated Points.** For Stone-like topologies, the existence of isolated points is tightly connected with the expressive power of the underlying descriptor. Say that a point $\boldsymbol{x} \in \boldsymbol{X}_D$ is **characterizable** by $D$ in $\boldsymbol{X}_D$ if there exists a finite set of formulas $A \subseteq \overline{D}$ such that for all $\boldsymbol{y} \in \boldsymbol{X}_D$, if $\boldsymbol{y} \vDash \varphi$ for all $\varphi \in A$, then $\mathbf{y} = \boldsymbol{x}$. We obtain the following:

PROPOSITION 27. *Let* $(\boldsymbol{X}_D, \mathcal{T}_D)$ *be a modal space with it's Stone-like topology. Then* $\boldsymbol{x} \in \boldsymbol{X}_D$ *is an isolated point of* $\boldsymbol{X}_D$ *iff* $\boldsymbol{x}$ *is characterizable by* $D$ *in* $\boldsymbol{X}_D$.

PROOF. *Left-to-right*: Assume $\{\boldsymbol{x}\}$ is open in $\mathcal{T}_D$. Since $\{\boldsymbol{x}\}$ is a singleton, it necessarily is contained in the basis of $\mathcal{T}_D$. Hence, it is a finite intersection of subbasic elements, i.e., $\{\boldsymbol{x}\} = \bigcap_{\varphi \in A}[\varphi]$ for some finite $A \subseteq \overline{D}$. Then $A$ characterizes $\boldsymbol{x}$. *Right-to-left*: Let $A$ be the set characterizing $\boldsymbol{x}$ in $\boldsymbol{X}_D$. For each $\varphi \in A$, $[\varphi]$ is open in $\mathcal{T}_D$ by definition. As $A$ is finite, also $\bigcap_{\varphi \in A}[\varphi]$ is open. Hence $\{\boldsymbol{x}\} \in \mathcal{T}_\mathcal{D}$. ⊣

**6.2.1.** *Perfect Spaces.* Recall that a topological space $(X, \mathcal{T})$ is called **perfect** if it contains no isolated points. In perfect spaces, every point is the limit of some sequence, and may hence be approximated arbitrarily well. The property is enjoyed by several natural classes of modal spaces under their Stone(-like) topologies, cf. Corollary 29. Proposition 27 implies that a space of models is perfect iff no points $\boldsymbol{x} \in \boldsymbol{X}_D$ is $D$-characterizable. If $X$ is sound and $\Lambda$-saturated with respect to some logic $\Lambda$ and descriptor $D$, we get the following:

PROPOSITION 28. *Let* $D \subseteq \mathcal{L}$, *let* $\Lambda$ *be a logic and let* $X$ *a set of* $\Lambda$-models $\Lambda$-saturated with respect to $D$. *Then* $(\boldsymbol{X}_D, \mathcal{T}_D)$ *is perfect iff for every finite* $\Lambda$-consistent set $A \subseteq \overline{D}$ *there is some* $\psi \in D$ *such that both* $\psi \wedge \bigwedge_{\chi \in A} \chi$ *and* $\neg\psi \wedge \bigwedge_{\chi \in A} \chi$ *are* $\Lambda$-consistent.

PROOF. For the left to right direction, assume that $(\boldsymbol{X}_D, \mathcal{T}_D)$ is perfect. Let $A \subseteq \{\varphi, \neg\varphi \colon \varphi \in D\}$ be finite and $\Lambda$-consistent. We have to show that there is some $\psi \in D$ such that $\psi \wedge \bigwedge_{\chi \in A} \chi$ and $\neg\psi \wedge \bigwedge_{\chi \in A} \chi$ are both $\Lambda$-consistent. As $X_D$ is $\Lambda$-saturated with respect to $D$, there is some $\boldsymbol{x} \in X_D$ with $x \vDash \bigwedge_{\chi \in A} \chi$. As $(\boldsymbol{X}_D, \mathcal{T}_D)$ is perfect, we have that $\bigcap_{\varphi \in A}[\varphi]_D \supsetneq \{\boldsymbol{x}\}$, i.e. there is some $\boldsymbol{y} \neq \boldsymbol{x} \in \bigcap_{\varphi \in A}[\varphi]_D$. By construction, this implies that there is some $\psi \in D$ such that $\boldsymbol{x} \vDash \psi$ and $\boldsymbol{y} \nvDash \psi$ or vice versa. Either way, $\boldsymbol{x}$ and $\boldsymbol{y}$ witness that $\psi \wedge \bigwedge_{\chi \in A} \chi$ and $\neg\psi \wedge \bigwedge_{\chi \in A} \chi$ are both $\Lambda$-consistent. For the right to left direction pick $\boldsymbol{x} \in X_D$. We show that $\boldsymbol{x}$ is not an isolated point. By Proposition 27, it suffices to show that $\boldsymbol{x}$ is not characterizable by $D$ in $\boldsymbol{X}_D$. For a contradiction, assume that there is some finite set $A \subseteq \overline{D}$ characterizing $\boldsymbol{x}$. Hence, there is some $\psi \in D$ such that both $\psi \wedge \bigwedge_{\chi \in A} \chi$ and $\neg\psi \wedge \bigwedge_{\chi \in A} \chi$ are $\Lambda$-consistent. As $X$ is $\Lambda$-saturated, there are some $y, z \in X$ with $y \vDash \psi \wedge \bigwedge_{\chi \in A} \chi$ and $z \vDash \neg\psi \wedge \bigwedge_{\chi \in A} \chi$. As $\psi \in D$ we have that $\boldsymbol{y} \neq \boldsymbol{z}$. In particular $\boldsymbol{x} \neq \boldsymbol{y}$ or $\boldsymbol{x} \neq \boldsymbol{z}$, contradicting the assumption that $A$ characterizes $\boldsymbol{x}$. ⊣

Note that if $D$ is closed under negations and disjunctions, we could relax the assumption to stating that for any $\Lambda$-consistent $\varphi \in D$ there is some $\psi \in D$ such

that $\varphi \wedge \psi$ and $\varphi \wedge \neg\psi$ are both $\Lambda$-consistent. This property is enjoyed by many classic modal logics. In particular, we have:

COROLLARY 29. *For the following modal logics, $(X_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ is perfect if $X$ is saturated with respect to $\mathcal{L}$: i) the normal modal logic $K$ with an infinite set of atoms, as well as ii) $KD$, iii) $KD45^n$ for $n \geq 2$ and iv) $S5^n$ for $n \geq 2$.*

Apart from the relation to convergence, it may be noted that any perfect space that is additionally compact is homeomorphic to the Cantor set (as every totally disconnected compact metric space is; see e.g. [47, Ch. 12]).

**6.2.2.** *Imperfect Spaces.* It is not difficult to find $\Lambda$-complete spaces $(\boldsymbol{X}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ that contain isolated points. We provide two examples. The first shows that, when working in a language with finite signature, then e.g. for the minimal normal modal logic $K$, the $K$-complete space will have an abundance of isolated points.

PROPOSITION 30. *Let $\mathcal{L}$ have finite signature $(\Phi, \mathcal{I})$ and let $\Lambda$ be such that $\bigvee_{i \in \mathcal{I}} \Diamond_i \top$ is not a theorem. If $(\boldsymbol{X}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ is $\Lambda$-complete, then it contains an isolated point. If the logic is exactly $K$, then it contains infinitely many isolated points.*

PROOF. Since $\bigvee_{i \in \mathcal{I}} \Diamond_i \top$ is not a theorem, there is some an atomic valuation $\varphi$ such that the formula $\varphi \wedge \bigwedge_{i \in \mathcal{I}} \Box_i \bot$ is consistent. This formula completely characterizes the point $\boldsymbol{x}$ in $\boldsymbol{X}_{\mathcal{L}}$ for which every pointed Kripke model $x \in \boldsymbol{x}$ has the valuation encoded by $\varphi$ at the designated state and all relations from that state are empty. This point is clearly isolated. If $\Lambda$ is exactly $K$, there are for each $n \in \mathbb{N}$ only finitely many modally different models satisfying $\psi_n = \bigwedge_{i \in \mathcal{I}} \left( \bigwedge_{m < n} \Diamond_i^m \top \wedge \neg\Diamond_i^n \top \right)$, hence $[\psi_n]$ is finite in $\boldsymbol{X}_{\mathcal{L}}$. This, together with the fact that $(\boldsymbol{X}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ is Hausdorff implies that any $\boldsymbol{x} \in [\psi_n]$ is characterizable by $\mathcal{L}$ making $\boldsymbol{x}$ isolated, cf. Proposition 27.                    ⊣

For the second example, we turn to *epistemic logic with common knowledge*. Let $(\Phi, \mathcal{I})$ be a finite signature with $\mathcal{I} = \{1, ..., n, G\}$. Let $EQ$ be the class of pointed Kripke models for $(\Phi, \mathcal{I})$ where for each $i \in \mathcal{I} \backslash \{G\}$, $R(i)$ is an equivalence relation and $R(G)$ the transitive closure of $\bigcup_{i \leq n} R(i)$. Let $\mathcal{L}$ be the language based on $(\Phi, \mathcal{I})$ and S5C the appropriate multi-agent epistemic logic with common knowledge axioms.[9] Then:

PROPOSITION 31. *The set of isolated points in $(\boldsymbol{EQ}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ is infinite.*

PROOF. This follows from Proposition 27 as every finite model of $EQ$ is characterizable by a single formula cf. [12, Lemma 3.4].[10]                    ⊣

Proposition 31 may be of interest in relation to discussions concerning the *attainability of common knowledge*, cf. [33, 34, 39]: If a set of agents is seeking to attain common knowledge of the atomic valuation $\varphi$ of the designated state, then their target point $\boldsymbol{x}$ will be characterized by $\Box_G \varphi$, and hence be isolated by Proposition 27. Hence, by Proposition 26.2, common knowledge must be

---

[9]See e.g. [12, 33]. The resulting logic is non-compact; see e.g. [30].
[10]See also [6, Prop. 2.4] for a similar result and [62] for a concise presentation of related material.

attained in finite time if it is to be attained in the limit. In a nutshell, common knowledge cannot be attained gradually.[11]   Taking $L$ to be the set of all $\mathcal{L}$ formulas not containing the common knowledge operator $\Box_G$, then $\boldsymbol{EQ}_L$ will be identical to $\boldsymbol{EQ}_{\mathcal{L}}$, but $\boldsymbol{x}$ will not be isolated in $(\boldsymbol{EQ}_L, \mathcal{T}_L)$. Hence, if the common knowledge operator is not reflected in the topology, the agents may converge to their target in the limit without converging in finite time. There is an interplay here between expressibility, non-compactness, topology and formal epistemology which we hope to address in later work.

**6.2.3.** *Discrete Spaces.* Finally, we note that the opposite extreme of perfect spaces are also realizable: A topological space is called **discrete** when every point is isolated.

PROPOSITION 32. *Let $\mathcal{L}$ be the mono-modal language over a finite atom set $\Phi$ and let $X$ be S5-complete. Then $(\boldsymbol{X}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}})$ is discrete.*

PROOF. The set $\boldsymbol{X}_{\mathcal{L}}$ is finite. As it is also metrizable, it is discrete.                    ⊣

**§7. Maps and Model Transformations.** In dynamic epistemic logic, dynamics are introduced by transitioning between pointed Kripke models from some set $X$ using a possibly partial map $f : X \to X$ often referred to as a *model transformer*. Many model transformers have been suggested in the literature, the most well-known being *truthful public announcement* [50], $!\varphi$, which maps $x$ to $x_{|\varphi}$, restricting $[\![x]\!]$ to the truth set of $\varphi$. Truthful public announcements are a special case of a rich class of model transformers definable through a particular graph operation, *product update*, of pointed Kripke models with *action models*. Due to their generality, popularity and wide applicability, we focus on the general class of maps on modal spaces induced by action models applied using product update.

An especially general version of action models is *multi-pointed* action models with *postconditions*. Postconditions allow action states in an action model to change the valuation of atoms [16,29], thereby also allowing the representation of information about ontic change. Permitting multiple points allows the actual action state to depend on the pointed Kripke model to be transformed, thus generalizing single-pointed action models. Multi-pointed action models are also referred to as *epistemic programs* in [6], and allow encodings akin to *knowledge-based programs* [33] of interpreted systems, cf. [55]. Allowing for multiple points renders the class of action models *Turing complete* [21], even when restricting to postcondition free product models. [43].

**7.1. Action Models and Product Update.** A **multi-pointed action model** is a tuple $\Sigma\Gamma = ([\![\Sigma]\!], \mathsf{R}, pre, post, \Gamma)$ where $[\![\Sigma]\!]$ is a non-empty set of **actions**. The map $\mathsf{R} : \mathcal{I} \to \mathcal{P}([\![\Sigma]\!]^2)$ assigns an **accessibility relation** $\mathsf{R}_i$ on $[\![\Sigma]\!]$ to each agent $i \in \mathcal{I}$. The map $pre : [\![\Sigma]\!] \to \mathcal{L}$ assigns to each action a **precondition**,

---

[11]In fact, this hold true for any common knowledge the agents may hope to attain, not only about the full valuation: Proposition 25 implies that common knowledge about some $\psi$, if reached at the limit, will already be attained in finite time whenever $\Box_G\psi \in D$ or $D$ is $S5C$-representative.

and the map $post : \llbracket \Sigma \rrbracket \to \mathcal{L}$ assigns to each action a **postcondition**,[12] which must be $\top$ or a conjunctive clause[13] over $\Phi$. Finally, $\Gamma \subseteq \llbracket \Sigma \rrbracket$ is a non-empty set of **designated actions**.

To obtain well-behaved total maps on modal spaces, we must invoke a set of mild, but non-standard, requirements: Let $X$ be a set of pointed Kripke models. Call $\Sigma\Gamma$ **precondition finite** if the set $\{pre(\sigma) \in \mathcal{L} : \sigma \in \llbracket \Sigma \rrbracket\}$ is finite (up to logical equivalence). This is needed for our proof of continuity. Call $\Sigma\Gamma$ **exhaustive over** $X$ if for all $x \in X$, there is a $\sigma \in \Gamma$ such that $x \vDash pre(\sigma)$. This conditions ensures that the action model $\Sigma\Gamma$ is universally applicable on $X$. Finally, call $\Sigma\Gamma$ **deterministic over** $X$ if $X \vDash pre(\sigma) \wedge pre(\sigma') \to \bot$ for any two preconditions $\sigma, \sigma' \in \Gamma, \sigma \neq \sigma'$. Together with exhaustivity, this condition ensures that the product of $\Sigma\Gamma$ and any $Ms \in X$ is a (single-)pointed Kripke model, i.e., that the actual state after the updates is well-defined and unique.

Let $\Sigma\Gamma$ be exhaustive and deterministic over $X$ and let $Ms \in X$. Then the **product update** of $Ms$ with $\Sigma\Gamma$, denoted $Ms \otimes \Sigma\Gamma$, is the pointed Kripke model $(\llbracket M\Sigma \rrbracket, R', \llbracket \cdot \rrbracket', s')$ with

$$\llbracket M\Sigma \rrbracket = \{(s, \sigma) \in \llbracket M \rrbracket \times \llbracket \Sigma \rrbracket : Ms \vDash pre(\sigma)\}$$
$$R' = \{((s, \sigma), (t, \tau)) : (s, t) \in R_i \text{ and } (\sigma, \tau) \in \mathsf{R}_i\}, \text{ for all } i \in \mathcal{I}$$
$$\llbracket p \rrbracket' = \{(s, \sigma) : s \in \llbracket p \rrbracket, post(\sigma) \nvDash \neg p\} \cup \{(s, \sigma) : post(\sigma) \vDash p\}, \text{ for all } p \in \Phi$$
$$s' = (s, \sigma) : \sigma \in \Gamma \text{ and } Ms \vDash pre(\sigma)$$

Call $\Sigma\Gamma$ **closing over** $X$ if for all $x \in X$, $x \otimes \Sigma\Gamma \in X$. With exhaustivity and deterministicality, this ensures that $\cdot \otimes \Sigma\Gamma$ induces a well-defined total map on $X$.

**7.2. Clean Maps on Modal Spaces.** If two pointed Kripke models $x$ and $y$ are $\mathcal{L}$-modally equivalent, then so are $x \otimes \Sigma\Gamma$ and $y \otimes \Sigma\Gamma$ for any product model $\Sigma\Gamma$, cf. [6].[14] Hence, action models applied using product update yield natural maps $\cdot \otimes \Sigma\Gamma$ on modal spaces $\boldsymbol{X}_\mathcal{L}$. The class of maps of interest in the present is the following:

DEFINITION. Let $\boldsymbol{X}_\mathcal{L}$ be a modal space. A map $\boldsymbol{f} : \boldsymbol{X}_\mathcal{L} \to \boldsymbol{X}_\mathcal{L}$ is called **clean** if there exists a precondition finite, multi-pointed action model $\Sigma\Gamma$ closing, deterministic and exhaustive over $X$ such that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}$ iff $\boldsymbol{x} \otimes \Sigma\Gamma = \boldsymbol{y}$ for all $\boldsymbol{x} \in \boldsymbol{X}_\mathcal{L}$.

REMARK. Replacing $\boldsymbol{X}_\mathcal{L}$ with $\boldsymbol{X}_D$ for an arbitrary descriptor $D \subseteq \mathcal{L}$ in the definition of clean maps will not in general result in well-defined maps $\cdot \otimes \Sigma\Gamma$ on $\boldsymbol{X}_D$. E.g.: Let $p$ and $q$ be atoms of $\mathcal{L}$ and let $D = \{p, \neg p\}$. Let $\Sigma\Gamma$ have $\llbracket \Sigma \rrbracket = \Gamma = \{\sigma, \tau\}$ with $pre(\sigma) = q, pre(\tau) = \neg q$ and $post(\sigma) = \top, post(\tau) = p$. Then for $x \vDash p \wedge q$ and $y \vDash p \wedge \neg q$ we have $y \in \boldsymbol{x} \in \boldsymbol{X}_D$, but $y \otimes \Sigma\Gamma \notin \boldsymbol{x} \otimes \boldsymbol{\Sigma\Gamma}$. For $\mathcal{L}$-representative descriptors $D$ over $X$, clean maps are, of course, well-defined.

---

[12]The precondition of $\sigma$ specify the conditions under which $\sigma$ is executable, while its post-condition dictates the posterior values of a finite, possibly empty, set of atoms.

[13]I.e. a conjunction of literals, where a literal is an atom or a negated atom.

[14]Baltag and Moss [6] show that multi-pointed action models applied using product update preserve bisimulation, which in turn implies that they preserve modal equivalence.

Below, we show that clean maps are continuous with respect to the Stone(-like) topology on $\boldsymbol{X}_{\mathcal{L}}$.

REMARK. By Proposition 12 and Lemma 3, the following analysis equally applies to the Stone(-like) topology on $\boldsymbol{X}_D$ for any descriptor $D$ that is $\mathcal{L}$-representative over $X$.

In general, the same clean map may be induced by several different action models. In showing clean maps continuous, we will make use of the following:

LEMMA 33. *Let $\boldsymbol{f} : \boldsymbol{X}_{\mathcal{L}} \to \boldsymbol{X}_{\mathcal{L}}$ be a clean map based on $\Sigma\Gamma$. Then there exists a precondition finite, multi-pointed action model $\Sigma'\Gamma'$ deterministic over $X$ that also induces $\boldsymbol{f}$ such that for all $\sigma, \sigma' \in [\![\Sigma']\!]$, either $\vDash pre(\sigma) \wedge pre(\sigma') \to \bot$ or $pre(\sigma) = pre(\sigma')$.*

PROOF. Assume we are given any precondition finite, multi-pointed action model $\Sigma\Gamma$ deterministic over $X$ generating $\boldsymbol{f}$. We construct an equivalent action model, $\Sigma'\Gamma'$, with the desired property.

For the preconditions, note that for every finite set of formulas $S = \{\varphi_1 \ldots \varphi_n\}$ there is some finite set formulas $S' = \{\psi_1, \ldots, \psi_m\}$ with the properties that any $\psi_i \neq \psi_j \in S'$ are mutually inconsistent and that for each $\varphi \in S$ there is some $J(\varphi) \subseteq S'$ such that $\vdash \bigvee_{\psi \in J(\varphi)} \psi \leftrightarrow \varphi$. One suitable candidate for such a set $S'$ is $\{\bigwedge_{k \leq n} \chi_k \colon \chi_k \in \{\varphi_k, \neg\varphi_k\}\}$: The disjunction of all conjunctions with $\chi_k = \varphi_k$ is equivalent with $\varphi_k$.

By assumption, $S = \{pre(\sigma) \colon \sigma \in [\![\Sigma]\!]\}$ is finite. Let $S'$ and $J(\varphi)$ be as above. Construct $\Sigma'\Gamma'$ as follows: For every $\sigma \in [\![\Sigma]\!]$ and every $\psi \in J(pre(\sigma))$, the set $[\![\Sigma']\!]$ contains a state $e^{\sigma,\psi}$ with $pre(e^{\{\sigma,\psi\}}) = \psi$ and $post(e^{\{\sigma,\psi\}}) = post(\sigma)$. Let $R'$ be given by $(e^{\sigma,\psi}, e^{\sigma',\psi'}) \in R'$ iff $(\sigma, \sigma') \in R$. Finally, let $\Gamma' = \{e^{\{\sigma,\psi\}} \colon \sigma \in \Gamma\}$.

The resulting multi-pointed action model $\Sigma'\Gamma'$ is again precondition finite and deterministic over $X$ while satisfying that for all $\sigma, \sigma' \in [\![\Sigma']\!]$, either $\vDash pre(\sigma) \wedge pre(\sigma') \to \bot$ or $\vDash pre(\sigma) \leftrightarrow pre(\sigma')$. Moreover, for any $x \in X$, the models $x \otimes \Sigma\Gamma$ and $x \otimes \Sigma'\Gamma'$ are bisimilar witnessed by the relation connecting $(s, \sigma) \in [\![x \otimes \Sigma\Gamma]\!]$ and $(s', e^{\sigma',\psi}) \in [\![x \otimes \Sigma'\Gamma']\!]$ iff $s = s'$ and $\sigma = \sigma'$. Hence, the maps $\boldsymbol{f}$, $\boldsymbol{f}' : \boldsymbol{X}_{\mathcal{L}} \to \boldsymbol{X}_{\mathcal{L}}$ defined by $\boldsymbol{x} \mapsto \boldsymbol{x} \otimes \boldsymbol{\Sigma\Gamma}$ and $\boldsymbol{x} \mapsto \boldsymbol{x} \otimes \boldsymbol{\Sigma'\Gamma'}$ are identical. $\dashv$

**7.3. Continuity of Clean Maps.** We show that the metrics introduced earlier square well with respect to the analysis of dynamics induced by clean maps by showing the latter continuous in the induced topology:

PROPOSITION 34. *Any clean map $\boldsymbol{f} : \boldsymbol{X}_{\mathcal{L}} \to \boldsymbol{X}_{\mathcal{L}}$ is uniformly continuous with respect to the metric topology generated by any $d_w \in \mathcal{D}_{(X,D)}$ where $D$ is finitely $\mathcal{L}$-representative with respect to $X$.*

In the proof, we make use of the following lemma:

LEMMA 35. *Let $(\boldsymbol{X}_{\mathcal{L}}, d_w)$ be a metric space, $d_w \in \mathcal{D}_{(X,D)}$ for $D$ finitely $\mathcal{L}$-representative with respect to $X$. Then*

1. *For every $\epsilon > 0$, there are formulas $\chi_1, \ldots, \chi_l \in \mathcal{L}$ such that every $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$ satisfies some $\chi_i$, and whenever $\boldsymbol{x} \vDash \chi_i$ and $\boldsymbol{z} \vDash \chi_i$ for some $i \leq l$, then $d_w(\boldsymbol{y}, \boldsymbol{z}) < \epsilon$.*

2. *For every $\varphi \in \mathcal{L}$, there is a $\delta$ such that for all $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$, if $\boldsymbol{x} \vDash \varphi$ and $d_w(\boldsymbol{x}, \boldsymbol{y}) < \delta$, then $\boldsymbol{y} \vDash \varphi$.*

PROOF OF LEMMA 35. For 1., note that there is some $n > 0$ for which $\sum_{k=n}^{\infty} w(\varphi_k) < \epsilon$. Let $J_1, ..., J_{2^{n-1}}$ be an enumeration of the subsets of $\{1, ..., n - 1\}$. For each $i \in \{1, ..., 2^{n-1}\}$ let the formula $\chi_i$ be $\bigwedge_{j \in J_i} \varphi_j \wedge \bigwedge_{j \notin J_i} \neg \varphi_j$. Then each $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$ must satisfy some $\chi_i$ with $i \leq 2^{n-1}$. Moreover, whenever $\boldsymbol{y} \vDash \chi_i$ and $\boldsymbol{z} \vDash \chi_i$, $d_w(\boldsymbol{y}, \boldsymbol{z}) = \sum_{k=1}^{\infty} w(\varphi_k) d_k(\boldsymbol{y}, \boldsymbol{z}) = \sum_{k=n}^{\infty} w(\varphi_k) d_k(\boldsymbol{y}, \boldsymbol{z}) < \epsilon$. For 2., let $\varphi \in \mathcal{L}$ be given. Since $D$ is finitely $\mathcal{L}$-representative with respect to $X$, there is $\{\psi_i\}_{i \in I} \subseteq D$ finite such that for all sets $S = \{\psi_i\}_{i \in J} \cup \{\neg \psi_i\}_{i \in I \setminus J}$ with $J \subseteq I$ either $\varphi$ or $\neg \varphi$ is entailed by $S$ in $X$. Then $\delta := \min_{i \in I} w(\psi_i)$ yields the desired. $\dashv$

PROOF OF PROPOSITION 34. We show that $\boldsymbol{f}$ is uniformly continuous, using the $\varepsilon$-$\delta$ formulation of continuity.

Assume that $\epsilon > 0$ is given. We have to find some $\delta > 0$ such that for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_{\mathcal{L}}$ $d_w(\boldsymbol{x}, \boldsymbol{y}) < \delta$ implies $d_w(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{f}(\boldsymbol{y})) < \epsilon$. By Lemma 35.1, there exist $\chi_1, \ldots, \chi_l$ such that $\boldsymbol{f}(\boldsymbol{x}) \vDash \chi_i$ and $\boldsymbol{f}(\boldsymbol{y}) \vDash \chi_i$ implies $d_w(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{f}(\boldsymbol{y})) < \epsilon$ and for every $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$ there is some $i \leq l$ with $\boldsymbol{f}(\boldsymbol{x}) \vDash \chi_i$. We use $\chi_1, \ldots, \chi_l$ to find a suitable $\delta$:

**Claim:** There is a function $\boldsymbol{\delta} : \mathcal{L} \to (0, \infty)$ such that for any $\varphi \in \mathcal{L}$, if $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi$ and $d_w(\boldsymbol{x}, \boldsymbol{y}) < \boldsymbol{\delta}(\varphi)$, then $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi$.

Clearly, setting $\delta = \min\{\boldsymbol{\delta}(\chi_i) : i \leq l\}$ yields a $\delta$ with the desired property. Hence the proof is completed by a proof of the claim. The claim is shown by constructing the function $\boldsymbol{\delta} : \mathcal{L} \to (0, \infty)$. This construction will proceed by induction over the complexity of $\varphi$. To begin with, fix a precondition finite, multi-pointed action model $\Sigma\Gamma$ closing, deterministic and exhaustive over $X$ such that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}$ iff $\boldsymbol{x} \otimes \Sigma\Gamma = \boldsymbol{y}$. To be explicit, the function $\boldsymbol{\delta} : \mathcal{L} \to (0, \infty)$ we construct depends on this action model. More precisely, $\boldsymbol{\delta}$ depends on the set $\{pre(\sigma) : \sigma \in [\![\Sigma]\!]\}$. The below construction of $\boldsymbol{\delta}$ is a simultaneous induction over all multi-pointed action models that are closing, deterministic and exhaustive over $X$ with set of preconditions exactly $\{pre(\sigma) : \sigma \in [\![\Sigma]\!]\}$. By Lemma 33, we can assume that for all $\sigma, \sigma' \in [\![\Sigma]\!]$ it holds that $pre(\sigma) = pre(\sigma')$ or $\vdash pre(\sigma) \wedge pre(\sigma') \to \bot$. By working with an extended language that contains $\neg, \wedge, \vee, \Box_i$ and $\Diamond_i$ as primitives, we can assume, without loss of generality that all negations in $\varphi$ immediately precede atoms.

If $\varphi$ is an atom or negated atom: By Lemma 35.2, there exists for any $\sigma \in [\![\Sigma]\!]$ some $\delta_\sigma$ such that whenever $\boldsymbol{x} \vDash pre(\sigma)$ and $d_w(\boldsymbol{x}, \boldsymbol{y}) < \delta_\sigma$ we also have that $\boldsymbol{y} \vDash pre(\sigma)$. Likewise, there is some $\delta_0$ such that whenever $\boldsymbol{x} \vDash \varphi$ and $d_w(\boldsymbol{x}, \boldsymbol{y}) < \delta_0$ we also have that $\boldsymbol{y} \vDash \varphi$. By assumption, the set $\{pre(\sigma) : \sigma \in [\![\Sigma]\!]\}$ is finite. Let $S = \{\delta_0\} \cup \{\delta_\sigma : \sigma \in [\![\Sigma]\!]\}$. We can thus set $\boldsymbol{\delta}(\varphi) = \min(S)$. To see that this $\boldsymbol{\delta}$ is as desired, assume $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi$. Let $x = Ms \in \boldsymbol{x}$. There is a unique $\sigma \in \Gamma$ in the deterministic, multi-pointed action model $(\Sigma, \Gamma)$ such that $(s, \sigma)$ is the designated state of $\boldsymbol{x} \otimes \Sigma\Gamma$. In particular, we have that $x \vDash pre(\sigma)$. By our choice of $\boldsymbol{\delta}(\varphi)$, we get that $d_w(\boldsymbol{x}, \boldsymbol{y}) < \boldsymbol{\delta}(\varphi)$ implies $\boldsymbol{y} \vDash pre(\sigma)$. For $y = Nt \in \boldsymbol{y}$, we thus have that $(t, \sigma)$ is the designated state of $Nt \otimes \Sigma\Gamma$. Moreover, we have $\boldsymbol{x} \vDash \varphi \Leftrightarrow \boldsymbol{y} \vDash \varphi$. Together, these imply that $\boldsymbol{f}(\boldsymbol{Nt}) \vDash \varphi$, i.e. $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi$.

If $\varphi$ is $\varphi_1 \wedge \varphi_2$, set $\boldsymbol{\delta}(\varphi) = \min(\boldsymbol{\delta}(\varphi_1), \boldsymbol{\delta}(\varphi_2))$ To show that this is as desired, assume $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi_1 \wedge \varphi_2$. We thus have $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi_1$ and $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi_2$. By induction, this implies that whenever $d_w(\boldsymbol{x}, \boldsymbol{y}) < \boldsymbol{\delta}(\varphi)$, we have $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi_1$ and $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi_2$ and hence $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi_1 \wedge \varphi_2$.

If $\varphi$ is $\varphi_1 \vee \varphi_2$, set $\delta(\varphi) = \min(\boldsymbol{\delta}(\varphi_1), \boldsymbol{\delta}(\varphi_2))$ To show that this is as desired, assume $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi_1 \vee \varphi_2$. We thus have $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi_1$ or $\boldsymbol{f}(\boldsymbol{x}) \vDash \varphi_2$. By induction, this implies that whenever $d_w(\boldsymbol{x}, \boldsymbol{y}) < \boldsymbol{\delta}(\varphi)$ we have $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi_1$ or $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi_2$ and hence $\boldsymbol{f}(\boldsymbol{y}) \vDash \varphi_1 \vee \varphi_2$.

If $\varphi$ is $\Diamond_r \varphi_1$: By Lemma 35.1, there are $\chi_1, \dots, \chi_l$ such that every $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$ satisfies some $\chi_i$ and whenever $\mathbf{z} \vDash \chi_i$ and $\mathbf{z}' \vDash \chi_i$ for some $i \leq l$ we have $d_w(\mathbf{z}, \mathbf{z}') < \boldsymbol{\delta}(\varphi_1)$.

Now, let $F = \{\Diamond_r(pre(\sigma) \wedge \chi_i) \colon \sigma \in \llbracket \Sigma \rrbracket, i \leq l\} \cup \{pre(\sigma) \colon \sigma \in \llbracket \Sigma \rrbracket\}$. By assumption, $F$ is finite. By Lemma 35.2, for each $\psi \in F$ there is some $\delta_\psi$ such that $\boldsymbol{x} \vDash \psi$ and $d_w(\boldsymbol{x}, \boldsymbol{y}) < \delta_\psi$ implies $\boldsymbol{y} \vDash \psi$. Set $\boldsymbol{\delta}(\varphi) = \min\{\delta_\psi \colon \psi \in F\}$.

To show that this is as desired, assume $\boldsymbol{f}(\boldsymbol{x}) \vDash \Diamond_r \varphi_1$ and let $\boldsymbol{y}$ be such that $d_w(\boldsymbol{x}, \boldsymbol{y}) < \boldsymbol{\delta}(\varphi)$. We have to show that $\boldsymbol{f}(\boldsymbol{y}) \vDash \Diamond_r \varphi_1$. Let $x = Ms \in \boldsymbol{x}$ and let the designated state of $x \otimes \Sigma\Gamma$ be $(s, \sigma)$. Since $x \otimes \Sigma\Gamma \vDash \Diamond_r \varphi_1$, there is some $(s', \sigma')$ in $\llbracket x \otimes \Sigma\Gamma \rrbracket$ with $(s, \sigma) R_r (s', \sigma')$. In particular $x \vDash \Diamond_r(pre(\sigma') \wedge \chi_i)$ for $\sigma' \in \llbracket \Sigma \rrbracket$ and some $i \leq l$. Thus also $\boldsymbol{y} \vDash \Diamond_r(pre(\sigma') \wedge \chi_i)$. Hence, for $y = Nt \in \boldsymbol{y}$, there is some $t' \in \llbracket y \rrbracket$ accessible from $y$'s designated state $t$ that satisfies $pre(\sigma') \wedge \chi_i$. By determinacy and the fact that $\vDash pre(\sigma) \wedge pre(\sigma') \rightarrow \bot$ whenever $pre(\sigma) \neq pre(\sigma')$, there is a unique $\tilde{\sigma} \in \Gamma$ with $pre(\tilde{\sigma}) = pre(\sigma')$. Let $\Gamma' = \Gamma \setminus \{\tilde{\sigma}\} \cup \{\sigma\}$. Then, $\Sigma\Gamma'$ is a precondition finite, multi-pointed action model deterministic over $X$. Let $\boldsymbol{f}'$ be the model transformer induced by $\Sigma\Gamma'$. As $\boldsymbol{f}'$ has the same set $\{pre(\sigma) \colon \sigma \in \llbracket \Sigma \rrbracket\}$ as $\boldsymbol{f}$, our induction hypothesis applies to $\boldsymbol{f}'$. Consider the models $Ms'$ and $Nt'$. We have that $Ms' \vDash \chi_i$ and $Nt' \vDash \chi_i$ jointly imply $d_w(\boldsymbol{Ms'}, \boldsymbol{Nt'}) < \boldsymbol{\delta}(\varphi_1)$ which, in turn, implies that $\boldsymbol{f}'(\boldsymbol{Ms'}) \vDash \varphi_1$ iff $\boldsymbol{f}'(\boldsymbol{Nt'}) \vDash \varphi_1$. In particular, we obtain that $\llbracket y \otimes \Sigma\Gamma \rrbracket, (t', \sigma') \vDash \varphi_1$. Since $(t, \sigma) R_r (t', \sigma')$ this implies that $\boldsymbol{f}(\boldsymbol{y}) \vDash \Diamond_r \varphi_1$.

If $\varphi$ is $\Box_r \varphi_1$: The construction is similar to the previous case. We only elaborate on the relevant differences. Again, there are some $\chi_1, \dots, \chi_l$ such that every $\boldsymbol{x} \in \boldsymbol{X}_{\mathcal{L}}$ satisfies some $\chi_i$ and whenever $\mathbf{z} \vDash \chi_i$ and $\mathbf{z}' \vDash \chi_i$ for some $i \leq l$ we have $d_w(\mathbf{z}, \mathbf{z}') < \boldsymbol{\delta}(\varphi_1)$.

Now, let $R = \{pre(\sigma) \wedge \chi_i \colon \sigma \in \llbracket \Sigma \rrbracket, i \leq l\}$ and let $F = \{\Box_r(\bigvee_{\psi \in J} \psi) \colon J \subseteq R\} \cup \{pre(\sigma) \colon \sigma \in \llbracket \Sigma \rrbracket\}$. Again, $F$ is finite and for each $\psi \in F$ there is some $\delta_\psi$ such that $\boldsymbol{x} \vDash \psi$ and $d_w(\boldsymbol{x}, \boldsymbol{y}) < \delta_\psi$ implies $\boldsymbol{y} \vDash \psi$. Set $\delta(\varphi) = \min\{\delta_\psi \colon \psi \in F\}$.

To show that this is as desired, assume $\boldsymbol{f}(\boldsymbol{x}) \vDash \Box_r \varphi_1$ and let $\boldsymbol{y}$ be such that $d_w(\boldsymbol{x}, \boldsymbol{y}) < \boldsymbol{\delta}(\varphi)$. We have to show that $\boldsymbol{f}(\boldsymbol{y}) \vDash \Box_r \varphi_1$. Let $Ms \in \boldsymbol{x}$ and $Nt \in \boldsymbol{y}$, let $(t, \sigma)$ be the designated state of $Nt \otimes \Sigma\Gamma$ and assume there is some $(t', \sigma')$ in $\llbracket Nt \otimes \Sigma\Gamma \rrbracket$ with $(t, \sigma) R_r (t', \sigma')$. We have to show that $\varphi_1$ holds at $\llbracket Nt \otimes \Sigma\Gamma \rrbracket, (t', \sigma')$. To this end, note that by construction, $t'$ satisfies $pre(\sigma') \wedge \chi_i$, for some $i \leq l$. By the choice of $\boldsymbol{\delta}(\varphi)$, there is some $s' \in \llbracket M \rrbracket$ with $sR_r s'$ that also satisfies $pre(\sigma') \wedge \chi_i$. Hence $(s', \sigma')$ is in $\llbracket Ms \otimes \Sigma\Gamma \rrbracket$ and $(s, \sigma) R_r (s', \sigma')$. By assumption we have $\llbracket Ms \otimes \Sigma\Gamma \rrbracket, (s', \sigma') \vDash \varphi_1$ and by an argument similar to the last case we get $\llbracket Nt \otimes \Sigma\Gamma \rrbracket, (t', \sigma') \vDash \varphi_1$, which is what we had to show. Hence $\boldsymbol{f}(\boldsymbol{y}) \vDash \Box_r \varphi_1$. $\dashv$

COROLLARY 36. *Any clean map $f : X_{\mathcal{L}} \to X_{\mathcal{L}}$ is continuous with respect to the Stone(-like) topology $\mathcal{T}_{\mathcal{L}}$.*

With Proposition 34, we have provided the foundation for understanding widely used dynamic epistemic logic operations on sets of pointed Kripke models as dynamical systems. As mentioned in the introduction, we refer to [11–14, 42, 55, 56] for broader motivations for this perspective and to [42] for initial results applying the approach.

References

[1] T. Ågotnes, H. van Ditmarsch, and Y. Wang. True Lies. *Synthese*, May 2017.

[2] M. Andersen, T. Bolander, H. van Ditmarsch, and M. H. Jensen. Bisimulation for single-agent plausibility models. In *AI 2013: Advances in Artificial Intelligence*, pages 277–288. Springer, 2013.

[3] G. Aucher. *Perspectives on belief and change.* Doctoral thesis, University of Otago and Université Toulouse III - Paul Sabatier, 2008.

[4] G. Aucher. Generalizing AGM to a Multi-Agent Setting. *Logic Journal of IGPL*, 18(4):530–558, 2010.

[5] A. Baltag, Z. Christoff, J. U. Hansen, and S. Smets. Logical Models of Informational Cascades. In J. van Benthem and F. Liu, editors, *Logic across the University: Foundations and Applications*, Studies in Logic, Volume 47, pages 405–432, London, 2013. College Publications.

[6] A. Baltag and L. S. Moss. Logics for Epistemic Programs. *Synthese*, 139(2):165–224, 2004.

[7] A. Baltag, L. S. Moss, and S. Solecki. The Logic of Public Announcements, Common Knowledge, and Private Suspicions (extended abstract). In *TARK 1998*, pages 43–56. Morgan Kaufmann Publishers, 1998.

[8] A. Baltag and B. Renne. Dynamic epistemic logic. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2016 edition, 2016.

[9] A. Baltag and S. Smets. A Qualitative Theory of Dynamic Interactive Belief Revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 9–58. Amsterdam University Press, 2008.

[10] A. Baltag and S. Smets. Group Belief Dynamics Under Iterated Revision: Fixed Points and Cycles of Joint Upgrades. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '09, pages 41–50, New York, NY, USA, 2009. ACM.

[11] J. van Benthem. Games in Dynamic-Epistemic Logic. *Bulletin of Economic Research*, 53(1):219–249, 2001.

[12] J. van Benthem. "One is a Lonely Number": Logic and Communication. In Z. Chatzidakis, P. Koepke, and W. Pohlers, editors, *Logic Colloquium '02*, Lecture Notes in Logic, 27, pages 95–128. Association for Symbolic Logic, 2002.

[13] J. van Benthem. *Logical Dynamics of Information and Interaction.* Cambridge University Press, 2011.

[14] J. van Benthem. Oscillations, Logic, and Dynamical Systems. In S. Ghosh and J. Szymanik, editors, *The Facts Matter*, pages 9–22. College Publications, 2016.

[15] J. van Benthem and P. Blackburn. Modal Logic: A Semantic Perspective. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier, 2008.

[16] J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, 2006.

[17] J. van Benthem, J. Gerbrandy, T. Hoshi, and E. Pacuit. Merging Frameworks for Interaction. *Journal of Philosophical Logic*, 38(5):491–526, 2009.

[18] J. van Benthem and F. Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17(2):157–182, 2007.

[19] J. van Benthem and S. Smets. Dynamic Logics of Belief Change. In H. van Ditmarsch, J. Y. Halpern, W. van der Hoek, and B. Kooi, editors, *Handbook of Logics for Knowledge and Belief*, pages 299–368. College Publications, 2015.

[20] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

[21] T. Bolander and M. Birkegaard. Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1):9–34, 2011.

[22] T. Bolander, M. H. Jensen, and F. Schwarzentruber. In *Proceedings of IJCAI 2015 (24th International Joint Conference on Artificial Intelligence)*, 2015.

[23] T. Caridroit, S. Konieczny, T. de Lima, and P. Marquis. On Distances Between KD45n Kripke Models and Their Use for Belief Revision. In G. Kaminka et al, editor, *ECAI 2016*. IOS Press, 2016.

[24] Z. Christoff and J. U. Hansen. A Logic for Diffusion in Social Networks. *Journal of Applied Logic*, 13(1):48–77, 2015.

[25] Dalal, Mukesh. Investigations Into a Theory of Knowledge: Preliminary Report. In *AAAi-88 Proceedings*, pages 475–479, 1988.

[26] C. Dégremont. *The Temporal Mind: Observations on the logic of belief change in interactive systems*. PhD thesis, University of Amsterdam, 2010.

[27] J. P. Delgrande. Preliminary considerations on the modelling of belief change operators by metric spaces. In *Proceedings of Tenth International Workshop on Non-Monotonic Reasoning NMR'2004*, pages 118–125, 2004.

[28] M. M. Deza and E. Deza. *Encyclopedia of Distances*. Springer, fourth edition, 2016.

[29] H. van Ditmarsch and B. Kooi. Semantic Results for Ontic and Epistemic Change. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, Vol. 3, pages 87–117. Amsterdam University Press, 2008.

[30] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2008.

[31] J. Eijck, J. Ruan, and T. Sadzik. Action emulation. *Synthese*, 185(S1):131–151, mar 2012.

[32] J. van Eijck, F. Sietsma, and Y. Wang. Composing models. *Journal of Applied Non-Classical Logics*, 21(3-4):397–425, 2011.

[33] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge*. The MIT Press, 1995.

[34] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. Common knowledge revisited. *Annals of Pure and Applied Logic*, 96(1):89–105, 1999.

[35] R. Frič. History of Sequential Convergence Spaces. In C. Aull and R. Lowen, editors, *Handbook of the History of General Topology, Vol. 1*, pages 343–355. Springer, 1997.

[36] J. Gerbrandy and W. Groeneveld. Reasoning about Information Change. *Journal of Logic, Language and Information*, 6(2):147–169, 1997.

[37] V. Goranko. Logical Topologies and Semantic Completeness. In J. van Eijck, V. van Oostrom, and A. Visser, editors, *Logic Colloquium'99*, pages 68–79. Lecture Notes in Logic 17, AK Peters, 2004.

[38] V. Goranko and M. Otto. Model Theory of Modal Logic. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier, 2008.

[39] J. Y. Halpern and Y. Moses. Knowledge and Common Knowledge in a Distributed Environment. *Journal of the ACM*, 37(3):549–587, 1990.

[40] R. W. Hamming. Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 29(2):147–160, 1950.

[41] J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. College Publications, 2nd, 2005 edition, 1962.

[42] D. Klein and R. K. Rendsvig. Convergence, Continuity and Recurrence in Dynamic Epistemic Logic. In A. Baltag, J. Seligman, and T. Yamada, editors, *Logic, Rationality, and Interaction (LORI 2017, Sapporo)*, LNCS, pages 108–122. Springer, 2017.

[43] D. Klein and R. K. Rendsvig. Turing Completeness of Finite, Epistemic Programs. *arXiv:1706.06845*, 2017.

[44] B. Kooi and B. Renne. Generalized Arrow Update Logic. In K. R. Apt, editor, *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Thirteenth Conference (TARK 2011)*, pages 205–211, New York, NY, USA, 2011. ACM.

[45] H. van Lee, R. K. Rendsvig, and S. van Wijk. Intensional Protocols for Dynamic Epistemic Logic. *Submitted*, 2017.

[46] D. Lehmann, M. Magidor, and K. Schlechta. Distance Semantics for Belief Revision. *Journal of Symbolic Logic*, 66(1):1–28, 2001.

[47] E. E. Moise. *Geometric Topology in Dimensions 2 and 3*. Graduate Texts in Mathematics. Springer New York, 1977.

[48] L. S. Moss. Finite models constructed from canonical formulas. *Journal of Philosophical Logic*, 36(6):605–640, 2007.

[49] J. R. Munkres. *Topology*. Prentice-Hall, 2nd edition, 2000.

[50] J. A. Plaza. Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.

[51] R. K. Rendsvig. Towards a Theory of Semantic Competence. Master's thesis, Dept. of Philosophy and Science Studies and Dept. of Mathematics, Roskilde University, 2011.

[52] R. K. Rendsvig. Aggregated Beliefs and Informational Cascades. In D. Grossi and O. Roy, editors, *Logic, Rationality, and Interaction (LORI 2013, Hangzhou)*, LNCS, pages 337–341. Springer, 2013.

[53] R. K. Rendsvig. Diffusion, Influence and Best-Response Dynamics in Networks: An Action Model Approach. In R. de Haan, editor, *Proceedings of the ESSLLI 2014 Student Session*, pages 63–75. arXiv:1708.01477, 2014.

[54] R. K. Rendsvig. Pluralistic ignorance in the bystander effect: Informational dynamics of unresponsive witnesses in situations calling for intervention. *Synthese*, 191(11):2471–2498, 2014.

[55] R. K. Rendsvig. Model Transformers for Dynamical Systems of Dynamic Epistemic Logic. In W. van der Hoek, W. H. Holliday, and W.-f. Wang, editors, *Logic, Rationality, and Interaction (LORI 2015, Taipei)*, LNCS, pages 316–327. Springer, 2015.

[56] T. Sadzik. Exploring the Iterated Update Universe. *ILLC Report PP-2006-263*, pages 1–34, 2006.

[57] D. Scott. Definitions by Abstraction in Axiomatic Set Theory. *Bulletin of the American Mathematical Society*, 61(5), 1955.

[58] O. Sokolsky, S. Kannan, and I. Lee. Simulation-based Graph Similarity. In H. Hermanns and J. Palsberg, editors, *Tools and Algorithms for the Construction and Analysis of Systems. TACAS 2006.*, volume LNCS 3920, pages 426–440. Springer, 2006.

[59] M. H. Stone. The Theory of Representation for Boolean Algebras. *Transactions of the American Mathematical Society*, 40(1):37–111, 1936.

[60] J. van Benthem, J. Gerbrandy, and B. P. Kooi. Dynamic Update with Probabilities. *Studia Logica*, 93(1):67–96, 2009.

[61] H. van Ditmarsch, W. van der Hoek, and B. Kooi. Dynamic Epistemic Logic and Knowledge Puzzles. In U. Priss, S. Polovina, and R. Hill, editors, *ICCS 2007*, LNAI 4604, pages 45–58. Springer, 2007.

[62] H. van Ditmarsch, D. Fernandez-Duque, and W. van der Hoek. On the definability of simulation and bisimulation in epistemic logic. *Journal of Logic and Computation*, 24(6):1209–1227, 2014.

[63] H. S. van Lee. A Formalization of the Greater Fools Theory with Dynamic Epistemic Logic. In A. Baltag, J. Seligman, and T. Yamada, editors, *Logic, Rationality, and Interaction (LORI 2017, Sapporo)*, LNCS, pages 585–597. Springer, 2017.

DEPARTMENT OF PHILOSOPHY, BAYREUTH UNIVERSITY

and

DEPARTMENT OF POLITICAL SCIENCE, UNIVERSITY OF BAMBERG

THEORETICAL PHILOSOPHY, LUND UNIVERSITY

and

CENTER FOR INFORMATION AND BUBBLE STUDIES, UNIVERSITY OF COPENHAGEN

**Paper VI**

# CHARACTERIZATIONS OF REDUCTION LAW EXISTENCE

RASMUS K. RENDSVIG

**Abstract.** A core methodology in dynamic epistemic logic is to prove completeness of a dynamic logic by reducing it to a static logic already known to be complete. The reduction is obtained by adding *reduction laws* to the static logic as axioms. This paper establishes two characterization theorems for reduction law existence of dynamic operators in arbitrary logics with Boolean base. The cases covered are compact and non-compact logics. The characterizing features of reducible maps are topological, presented relative to the Stone topology. A short discussion of applications and relations to dynamic epistemic logic follows the results.

**§1. Preliminaries.** Let $L$ be a recursively defined, at most countable language closed under negation (if $\varphi \in L$, then $\neg\varphi \in L$) and conjunction (if $\varphi, \psi \in L$, then $\varphi \wedge \psi \in L$). If the grammar of $L$ allows for quantifiers, then identify $L$ with its set of sentences. Let $\Lambda \subseteq L$ be a logic containing the axioms of classical logic, closed under *modus ponens*. Let $\mathcal{C}(\Lambda)$ be the set of complete theories in $L$ extending $\Lambda$, and assume that for every $\Lambda$-consistent $\varphi \in L$, for some $T \in \mathcal{C}(\Lambda)$, $\varphi \in T$. Then the **Stone topology** $\mathcal{S}(L)$ is defined on $\mathcal{C}(\Lambda)$ by the basis of clopen sets $\mathcal{B}(L) = \{[\varphi] \colon \varphi \in L\}$ where $[\varphi] = \{T \in \mathcal{C}(\Lambda) \colon \varphi \in T\}$.

A logic $\Lambda$ is **compact** if every $\Lambda$-inconsistent set of formulas has a finite $\Lambda$-inconsistent subset. The following lemma, central to the results, is shown in [7]:

LEMMA 1. *The logic $\Lambda$ is compact if, and only if, for any $A \subseteq \mathcal{C}(\Lambda)$, $A$ is clopen in $\mathcal{S}(\Lambda)$ iff there exists a $\varphi \in L$ with $A = [\varphi]$.*

Treat $\mathcal{C}(\Lambda)$ as a semantics for $L$ for which the satisfaction relation $\vDash$ fulfills that $T \vDash \varphi$ iff $\varphi \in T$. It is assumed that $\vDash$ is definable by a set of recursive clauses. The relation is extended to accommodate a dynamic operator: Let $f$ be a map on $\mathcal{C}(\Lambda)$. Let $L_f$ be the language obtained when adding closure under the monadic operator $\langle f \rangle$ to the generating grammar of $L$ (i.e., if $\varphi \in L_f$, then $\langle f \rangle\varphi \in L_f$). Let $\vDash_f$ be the satisfaction relation between $\mathcal{C}(\Lambda)$ and $L_f$ obtained by applying the clauses of $\vDash$ *mutatis mutandis* to $L_f$, together with the following: For all $\varphi \in L_f$, $T \vDash_f \langle f \rangle\varphi$ iff $f(T) \vDash_f \varphi$. For all $\varphi \in L_f$, let $[\![\varphi]\!] = \{T \in \mathcal{C}(\Lambda) \colon T \vDash_f \varphi\}$. The dynamic language $L_f$ is said to be **reducible** to $L$ if for every $\varphi_f \in L_f$, there exists a $\psi \in L$ such that for all $T \in \mathcal{C}(\Lambda)$, $T \vDash_f \varphi_f \leftrightarrow \psi$.

**§2. Results.** The following sufficiency condition for continuity is immediate:

PROPOSITION 1. *For any $\Lambda$ and any map $f$ on $\mathcal{C}(\Lambda)$, if $L_f$ is reducible to $L$, then $f$ is continuous in $\mathcal{S}(L)$.*

PROOF. Assume $L_f$ is reducible to $L$ and let $[\varphi] \in \mathcal{B}(L)$. Then $f^{-1}([\varphi])$ is open in $\mathcal{S}(\Lambda)$: By definition, $f^{-1}([\varphi]) = [\![\langle f \rangle \varphi]\!]$. By reducibility, there exists a $\psi \in L$ such that $\mathcal{C}(\Lambda) \vDash_f \langle f \rangle \varphi \leftrightarrow \psi$. Hence $f^{-1}([\varphi]) = \{T \in \mathcal{C}(\Lambda) \colon T \vDash \psi\} = [\psi]$, which is a basis element of $\mathcal{S}(L)$. ⊣

For compact logics, reducibility is also a necessary condition for continuity:

PROPOSITION 2. *If $\Lambda$ is compact, then for any map $f$ on $\mathcal{C}(\Lambda)$, if $f$ is continuous in $\mathcal{S}(L)$, then $L_f$ is reducible to $L$.*

PROOF. Assume $\Lambda$ compact and $f$ continuous. Then $L_f$ may be shown reducible to $L$ by induction on the complexity of $\varphi_f$. Where the main logical symbol of $\varphi_f$ is not $\langle f \rangle$, the argument is straightforward. Where the main symbol of $\varphi_f$ is $\langle f \rangle$, let $\chi \in L_f$ and assume there exists a $\psi \in L$ for which $\mathcal{C}(\Lambda) \vDash_f \chi \leftrightarrow \psi$. Where $\chi \in L$, this assumption is trivially satisfied; where $\chi \in L_f \backslash L$, the assumption serves as induction hypothesis. In either case, the reduction argument is as follows: Let $\varphi_f = \langle f \rangle \chi$. Then $[\![\langle f \rangle \chi]\!] = f^{-1}([\![\chi]\!])$. By assumption, $[\![\chi]\!] = [\psi]$ for some $\psi \in L$. As $[\psi]$ is clopen, so is $f^{-1}([\![\chi]\!])$ by continuity. Hence, by Lemma 1, there exists a $\psi' \in L$ with $[\psi'] = f^{-1}([\![\chi]\!])$, yielding the desired $\mathcal{C}(\Lambda) \vDash_f \langle f \rangle \chi \leftrightarrow \psi'$. ⊣

Compactness, moreover, is a necessary condition for the characterization of reducibility in terms of continuity:

PROPOSITION 3. *If $\Lambda$ is non-compact, then there exists a map $f$ on $\mathcal{C}(\Lambda)$ continuous in $\mathcal{S}(L)$ for which $L_f$ is not reducible to $L$.*

PROOF. Assume $\Lambda$ is non-compact. By Lemma 1, there exists a clopen set $U$ not identical to $[\psi]$ for any $\psi \in L$. The set $U$ is a countably infinite union of basis elements $\{[\varphi_i]\}_{i \in \mathbb{N}}$. Pick any $[\varphi_n] \subseteq U$. Then $f$ surjectively mapping $U$ to $[\varphi_n]$ and with $f(x) = x$ for all $x \in \mathcal{C}(\Lambda) \backslash U$ is well-defined and continuous. Then $U$ is $[\![\langle f \rangle \varphi_n]\!]$, different from $[\psi]$ for all $\psi \in L$. For every $\psi \in L$, there thus exists a $T \in \mathcal{C}(\Lambda)$ such that $T \vDash \neg \psi \wedge \langle f \rangle \varphi_n$. Hence, $L_f$ cannot be reduced to $L$. ⊣

As a corollary to these three propositions, it follows that compactness characterizes when continuity characterizes reducibility:

THEOREM 1. *The following are equivalent:*
1. *The logic $\Lambda$ is compact.*
2. *For any map $f$ on $\mathcal{C}(\Lambda)$, $f$ is continuous in $\mathcal{S}(L)$ iff $L_f$ is reducible to $L$.*

Open is what strengthening of continuity characterizes reducibility for non-compact logics. One option is the following, requiring that the preimage of any basis element is again a basis element: Call a map $f$ on $\mathcal{C}(\Lambda)$ **confined** in $L$ if for every $\varphi \in L$, there exists a $\psi \in L$ such that $f^{-1}([\varphi]) = [\psi]$.

PROPOSITION 4. *For any $\Lambda$ and any map $f$ on $\mathcal{C}(\Lambda)$, if $f$ is confined in $L$, then $L_f$ is reducible to $L$.*

PROOF. The proof is similar to that of Proposition 2. Where the main symbol of $\varphi_f$ is $\langle f \rangle$, let $\chi \in L_f$ and assume there exists a $\psi \in L$ for which $\mathcal{C}(\Lambda) \vDash_f \chi \leftrightarrow \psi$. Where $\chi \in L$, this assumption is trivially satisfied; where $\chi \in L_f \backslash L$, the assumption serves as induction hypothesis. In either case, the reduction

argument is as follows: Let $\varphi_f = \langle f \rangle \chi$. Then $[\![\langle f \rangle \chi]\!] = f^{-1}([\![\chi]\!])$. By assumption, $[\![\chi]\!] = [\varphi]$ for some $\varphi \in L$. From $f^{-1}([\![\chi]\!]) = f^{-1}([\varphi])$, conclude by confinedness the existence of a $\psi \in L$ for which $f^{-1}([\varphi]) = [\psi]$. This yields the desired $\mathcal{C}(\Lambda) \vDash_f \langle f \rangle \chi \leftrightarrow \psi$. $\dashv$

Every confined map is clearly continuous, but the converse may fail in the non-compact case: The map in the proof of Proposition 3 is exemplary. The argument further shows every non-confined map irreducible: For $\varphi \in L$, any continuous preimage of the clopen set $[\varphi]$ is either a clopen set $[\psi], \psi \in L$, or a clopen set $U$ not identical to $[\psi]$ for any $\psi \in L$. If the latter never obtains, the map is confined. If it does, then it is irreducible. Hence the following:

THEOREM 2. *The following are equivalent:*

1. *The logic $\Lambda$ is non-compact.*
2. *For any map $f$ on $\mathcal{C}(\Lambda)$, $f$ is confined in $L$ iff $L_f$ is reducible to $L$.*

**§3. Discussion.** The results have, for a concise presentation, been stated in terms of theories, following [5]. In simple steps, they are applicable to the standard semantic setting of dynamic epistemic logic [2, 3, 4]: With $\mathsf{S}$ a set of structures of interest for which some $\Lambda \subseteq L$ is sound and complete, let $\mathsf{S}_L$ be the quotient space of $\mathsf{S}$ under $L$-equivalence. Each equivalence class in $\mathsf{S}_L$ may then be identified with the unique theory in $\mathcal{C}(\Lambda)$ which its structures satisfy. Hence, if a map on $\mathsf{S}$ preserves $L$-equivalence (as e.g. by product update does, through bisimulation preservation [1]), it may be identified with the map it induces on $\mathsf{S}_L$ and, by extension, with a map on $\mathcal{C}(\Lambda)$. This allows for the results' application in the semantic realm.

A core approach in dynamic epistemic logic is to prove completeness through reduction. In [8], Plaza introduced the method: Given a static logic $\Lambda \subseteq L$ sound and complete with respect to a class of models $\mathsf{S}$, adding to $\Lambda$ as axioms a set of sound *reduction laws* $\{\langle f \rangle \varphi_i \leftrightarrow \psi_i\}_{i \in I}$ allowing the elimination of the dynamic operator $\langle f \rangle$ by iterated application establishes completeness of the extended, dynamic logic $\Lambda_f \subseteq L_f$ with respect to $\mathsf{S}$. That $L_f$ reduces to $L$ through reduction laws is typically shown using a translation and a complexity measure, yielding a concrete axiom system (see e.g [2, 4]). The present results implies the existence of suitable reduction laws from continuity (confinedness). Have reduction laws been elusive, their existence may instill hope; have reduction laws been sought to draw conclusions concerning expressivity or decidability of the dynamic case from the properties of the static, this may now be done on a non-constructive basis. Conversely, discontinuity implies the impossibility of reduction, which e.g. entails that transitive closure on Kripke models is non-reducible: A simple argument using the e.g. the *n-bisimulation metric* of [7] shows that operation discontinuous.

Finally, the existence of reduction laws for a map always implies its continuity in the Stone topology. Hence, the abundance of reduction systems in the literature may be taken to show continuity of the invoked operations. This, in turn, entails that they may be analyzed as *topological dynamical systems*, cf. the approach in [6, 7]. For motivations for such an endeavor, see [3, 6].

References

[1] ALEXANDRU BALTAG and LAWRENCE S. MOSS, *Logics for Epistemic Programs*, **Synthese**, vol. 139 (2004), no. 2, pp. 165–224.

[2] ALEXANDRU BALTAG, LAWRENCE S. MOSS, and SŁAWOMIR SOLECKI, *The Logic of Public Announcements, Common Knowledge, and Private Suspicions (extended abstract)*, **TARK '98: Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge**, pp. 43–56, Morgan Kaufmann Publishers, 1998.

[3] JOHAN VAN BENTHEM, **Logical Dynamics of Information and Interaction**, Cambridge University Press, 2011.

[4] HANS VAN DITMARSCH, WIEBE VAN DER HOEK, and BARTELD KOOI, **Dynamic Epistemic Logic**, Springer, 2007.

[5] VALENTIN GORANKO, *Logical Topologies and Semantic Completeness*, **Logic Colloquium'99** (J. van Eijck, V. van Oostrom, and A. Visser, editors), AK Peters, 2004, pp. 68–79.

[6] DOMINIK KLEIN and RASMUS K. RENDSVIG, *Convergence, Continuity and Recurrence in Dynamic Epistemic Logic*, **Logic, Rationality, and Interaction (LORI 2017)** (A. Baltag, J. Seligman, and T. Yamada, editors), Springer, 2017, pp. 108–122.

[7] DOMINIK KLEIN and RASMUS K. RENDSVIG, *Metrics for Formal Structure, with an Application to Kripke Models and their Dynamics*, **submitted**, (2017).

[8] JAN A. PLAZA, *Logics of Public Communications*, **Proceedings of the Fourth International Symposium on Methodologies for Intelligent Agents** (M.L. Emrich, M.S. Pfeifer, M. Hadzikadic, and Z.W. Ras, editors), Oak Ridge National Library, 1989, Republished in *Synthese*, 158(2), pp. 165–179, 2007, pp. 201–216.