Synthese (2009) 169:371–384
Knowledge, Rationality & Action 775–788
DOI 10.1007/s11229-009-9549-1

Bridging learning theory and dynamic epistemic logic

Nina Gierasimczuk

Received: 30 November 2008 / Accepted: 6 April 2009 / Published online: 23 April 2009

Abstract This paper discusses the possibility of modelling inductive inference (Gold 1967) in dynamic epistemic logic (see e.g. van Ditmarsch et al. 2007). The general purpose is to propose a semantic basis for designing a modal logic for learning in the limit. First, we analyze a variety of epistemological notions involved in identification in the limit and match it with traditional epistemic and doxastic logic approaches. Then, we provide a comparison of learning by erasing (Lange et al. 1996) and iterated epistemic update (Baltag and Moss 2004) as analyzed in dynamic epistemic logic. We show that finite identification can be modelled in dynamic epistemic logic, and that the elimination process of learning by erasing can be seen as iterated belief-revision modelled in dynamic doxastic logic. Finally, we propose viewing hypothesis spaces as temporal frames and discuss possible advantages of that perspective.

Keywords Identification in the limit · Learning by erasing · Induction · Learning by elimination · Co-learning · Finite identifiability · Dynamic epistemic logic · Dynamic doxastic logic · Epistemic update · Belief revision

1 Introduction

In this paper we discuss two perspectives on the phenomenon of epistemic change: formal learning theory (see e.g. Jain et al. 1999) and belief-revision theory in its interrelation with dynamic epistemic logic. Formal learning theory (LT) has grown out of the classical framework of identification in the limit (Gold 1967). The latter was introduced as a model of language acquisition, more specifically—grammar inference. Initially

N. Gierasimczuk (✉)
Institute for Logic, Language and Computation,
Universiteit van Amsterdam, Amsterdam, The Netherlands
e-mail: nina.gierasimczuk@gmail.com

restricted to syntax and underappreciated in semantic considerations, eventually the idea of identification has been applied also to the acquisition of natural language semantics (Tiede 1999; Costa Florêntio 2002; Gierasimczuk 2007) and modelling the process of scientific inquiry (Kelly 1996). Moreover, the design of model-theoretic learning (Osherson et al. 1997) and its application to belief-revision theory (Martin and Osherson 1998) can be viewed as an important step towards involving even more semantics.

In parallel, a perspective that focuses more on the epistemological than operational aspects of belief-change has been developed. After a precise language to discuss epistemic states of agents has been established (Hintikka 1962), the need of formalizing dynamics of knowledge emerged. The belief-revision AGM framework (Alchourrón et al. 1985) constitutes an attempt to talk about the dynamics of epistemic states. Belief-revision policies thus explained have been successfully modelled in dynamic epistemic logic (see van Ditmarsch et al. 2007) and in the above-mentioned model-theoretic learning (Martin and Osherson 1998).

In this paper we present a way to translate some notions of LT into the dynamic epistemic logic (DEL) and dynamic doxastic logic (DDL) paradigm. First, we recall the basics of formal learning theory in its set-theoretical version, concentrating on a special case—learning by erasing. We extensively discuss epistemological notions involved in learning in the limit, stressing the points that we view as important from a modal logic perspective. Then, we lay out a way to model finite identification in dynamic epistemic logic and learning by erasing in dynamic doxastic logic. In the end, we briefly mention a way to treat hypotheses as temporal frames.

This paper is meant as an introduction of LT notions to DEL community. Therefore, we assume that the reader is familiar with dynamic epistemic and dynamic doxastic logic (see van Ditmarsch et al. 2007). We do not discuss them extensively, instead we focus on formal learning theory.

2 Learning theory

2.1 Identification in the limit

The primary concern of learning theory is the process of identification in the limit (Gold 1967). The idea was originally motivated by the need of formalizing the process of inferring general conclusions from partial, inductively given information, like in the case of language learning (inferring grammars from sentences) and scientific inquiry (drawing general conclusions from partial experiments). We can think of these processes as games between Scientist and Nature. At the start we have a class of possible worlds together with a class of hypotheses (possible descriptions of worlds). Different hypotheses may describe the same world. We assume that both Scientist and Nature know what all the possibilities are, i.e., they both have access to the initial class. Nature chooses one of those possible worlds to be the actual one. Scientist has to guess which one it is. Scientist receives information about the world in an inductive manner. The stream of data is infinite and contains only and all the elements from the chosen reality. Each time Scientist receives a piece of information he answers with one of the hypotheses from the initial class. We say that Scientist identifies Nature's choice in the limit if after some finite number of guesses his answers stabilize on a correct hypothesis. Moreover, to discuss more general identifiability, we require that the same is true for all the possible worlds from the initial class, i.e., regardless of which element from the class is chosen by Nature to be true, Scientist can identify it in the limit on the basis of data about the actual world.

To formalize this setting we need to make the notion of *stream of data* clear. In learning theory such streams are often called 'environments'.[1]

Let us consider E—the set of all computably enumerable sets. Let $C \subseteq E$ be some class of c.e. sets. For each S in C we consider Turing machines h_n which generate S and in such a case we say that n is an index of S. The Turing machines will function as the conjectures that Scientist makes. It is well-known that each S has infinitely many indices. Let us take I_S to be the set of all indices of the set S, i.e, $I_S = \{n | h_n \, generates \, S\}$.

Definition 1 (Environment)
By environment of S, ε, we mean an infinite sequence of elements from S such that it enumerates all and only the elements from S, allowing repetitions.

Definition 2 (Notation)
We will use the following notation:

- ε_n is the n-th element of ε;
- $\varepsilon | n$ is a sequence $(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1})$;
- SEQ denotes a set of all finite initial segments of all environments;
- $set(\varepsilon)$ is a set of elements that occur in ε;
- h_n will refer to a hypothesis, i.e., a finite description of a set, a Turing machine generating S;
- L is a learning function—a map from finite data sequences to indexes of hypotheses, $L : SEQ \to I_{H_C}$.

The structure of the identifiability in the limit can be formulated by the following chain of definitions:

Definition 3 (Identification in the limit, LIM)
We say that a learning function L:

1. identifies $S \in C$ in the limit on ε iff there is a number k, such that for co-finitely many m, $L(\varepsilon | m) = k$ and $k \in I_S$;
2. identifies $S \in C$ in the limit iff it identifies S in the limit on every ε for S;
3. identifies C in the limit iff it identifies in the limit every $S \in C$.

The notion of identifiability can be strengthened in various ways. One radical case is to introduce a finiteness condition for identification.

[1] We are concerned here only with sequences of positive information (texts).

Definition 4 (Finite identification, FIN)
We say that a learning function L:

1. finitely identifies $S \in C$ on ε iff, when inductively given ε, at some point L outputs a single k, such that $k \in I_S$, and stops;
2. finitely identifies $S \in C$ iff it finitely identifies S on every ε for S;
3. finitely identifies C iff it finitely identifies every $S \in C$.

To clarify the notion of identification let us give some examples. First let us consider finite identifiability of some infinite class of finite sets.

Example 1 Let $C_0 = \{\{0, 1\}, \{0, 2\}, \{0, 3\}, \ldots\}$. The corresponding class of hypotheses[2] is $H_{C_0} = \{h_1, h_2, h_3, \ldots\}$. C_0 is finitely identifiable by the following function $L : SEQ \rightarrow H_{C_0}$:

$$L(\varepsilon|n) = \begin{cases} \text{is undefined if } set(\varepsilon|n) = \{0\}, \\ \max(set(\varepsilon|n)) \text{ otherwise.} \end{cases}$$

In other words, L outputs the correct hypothesis as soon as it receives a number different than 0, and the procedure ends.

There are simple classes of sets, even finite classes of finite sets, which are not finitely identifiable.

Example 2 Let $C_1 = \{\{1\}, \{1, 2\}, \{1, 2, 3\}\}$, with the corresponding class of hypotheses being $H_{C_1} = \{h_1, h_2, h_3\}$. C_1 is not finitely identifiable. To see that, assume that $\{1, 2\}$ is chosen to be the actual world. Then Scientist can never conclusively decide that h_2 is true. For all he knows, 3 might appear in the future, so he has to leave the h_3-possibility open.

C_1 is, however, identifiable in the limit by the following function $L : SEQ \rightarrow H_{C_1}$:

$$L(\varepsilon|n) = h_m, \text{ such that } m = \max(set(\varepsilon|n)).$$

Let us now extend the class C_1 so that it becomes infinite. We will consider such infinite cases in the following two examples.

Example 3 Let $C_2 = \{\{1\}, \{1, 2\}, \{1, 2, 3\}, ...\}$. The corresponding class of hypotheses is $H_{C_2} = \{h_1, h_2, h_3, ...\}$. This class is identifiable in the limit by the same function that identifies the finite class from Example 2, namely $L : SEQ \rightarrow H_{C_2}$:

$$L(\varepsilon|n) = h_m, \text{ such that } m = \max(set(\varepsilon|n)).$$

Now, let us see what happens when we add the set of all natural numbers to the class C_2.

[2] Finite sets of numbers can be taken to be their own finite descriptions, i.e., the hypothesis which corresponds to the finite set can simply enumerate all elements from the set.

Example 4 Let $C_3 = \{\mathbb{N}\} \cup \{\{1\}, \{1, 2\}, \{1, 2, 3\}, ...\}$. The corresponding class of hypotheses is $H_{C_3} = \{h_0, h_1, h_2, h_3, \ldots, \}$, such that h_0 = 'there are all natural numbers'. C_3 is not identifiable in the limit.

A short argument in favor of the above claim is as follows. Let us assume that there is a function L that identifies C_3. Then, there is a k and n, such that for all $m \geq n$, $L(\varepsilon|m) = k$. Now, if $k \in \{1, 2, 3, \ldots\}$, then L cannot identify the set $\mathbb{N}$. On the other hand, if $k = 0$ then L cannot identify $h_{max(set(\varepsilon|n))}$. So, we get a contradiction, L cannot identify C_3.

We already know that there are infinite classes of finite sets such that they are identifiable in the limit. In the next example we will show an infinite class of infinite sets that is identifiable in the limit.

Example 5 Let $C_4 = \{S_n | S_n = \mathbb{N} - \{n\}, n \in \mathbb{N}\}$. The corresponding class of hypotheses is $H_{C_4} = \{h_n =$ 'all positive integers except n'$|n \in \mathbb{N}\}$. C_4 is identifiable in the limit by the learning function $L : SEQ \rightarrow H_{C_4}$:

$$L(\varepsilon|n) = h_m \text{ such that } m = \min(\mathbb{N} - set(\varepsilon|n)).$$

To complete the picture, we add the set of all natural numbers to C_4. As a result we get a class that is not identifiable in the limit.

Example 6 This time let $C_5 = C_4 \cup \{\mathbb{N}\}$. C_5 is not identifiable in the limit. Here the argument is very similar to the one in Example 4.

2.2 Learning by erasing

Learning by erasing (Lange et al. 1996) is an epistemologically intuitive modification of the identification in the limit model. Very often the cognitive process of converging to a correct conclusion consists of eliminating those possibilities that are falsified during the inductive inquiry. Accordingly, in the formal model the outputs of the learning function are negative, i.e., the function each time eliminates a hypothesis, instead of explicitly guessing one that is supposed to be correct. A special case of learning by erasing is co-learning (Freivalds and Zeugmann 1995). The set $S \in C$ is co-learnable iff there is a function which stabilizes by eliminating all indexes from I_{H_C} except just one, from I_S. The difference between this approach and the usual identification is in the interpretation of the positive guess of the learning function. In learning by erasing one adds an ordering of the initial hypothesis space. This allows one to interpret the actual positive guess of the learning-by-erasing function to be the least hypothesis (in the given ordering) not yet eliminated.

Let us give now the two definitions that explain the notion of learning by erasing.

Definition 5 (Function stabilization)
In learning by erasing we say that a function stabilizes to number k on environment ε iff for co-finitely many $n \in \mathbb{N}$:

$$k = \min\{\mathbb{N} - \{L(\varepsilon|0), \ldots, L(\varepsilon|n)\}\}.$$

Definition 6 (Learning by erasing, e-learning)
We say that a learning function L:

1. learns $S \in C$ by erasing on ε iff L stabilizes to k on ε and $k \in I_S$;
2. learns $S \in C$ by erasing iff it learns by erasing S from every ε for S;
3. learns C by erasing iff it learns by erasing every $S \in C$.

3 Properties of identification

In this section we discuss learning-theoretic properties of identification in the limit that we view as relevant for modelling in dynamic epistemic logic.

3.1 Different nature of data and conclusions

The word "learning" is used to cover a variety of epistemic processes. One of them is the epistemic update of the form of one-step *learning that* φ, followed by a modification of the set of beliefs—with an attempt to add φ as it is. This scenario is usually modelled and analyzed in AGM framework and in dynamic epistemic logic. In the present setting the incoming information is spread over more than one step. Moreover, the incoming pieces of data are of a different nature than the actual thing being learned. Typically, at each finite step environment gives only partial information about a potentially infinite set. The relationship between data and hypothesis is like between sentences and grammars, natural numbers as such and Turing machines. Namely, if we know the hypothesis we can infer what kind of possible data are going to appear, but in principle we will not be able to make a conclusive inference from data to hypotheses.

3.2 Inductive, step-by-step process

The process of restricting the hypothesis space to only those hypotheses that are consistent with the incoming data resembles the event of public announcement. As we have seen above, this type of identification, learning by erasing, is in fact one of the possible learning scenarios. The important difference is that in DEL a scheme of arbitrary sequences of public announcements can be shrunk to one. This feature is absent in learning theory, since the point of convergence to a correct hypothesis is unknown and in general uncomputable. Therefore, finite sequences of data cannot be seen as a single announcement of a given hypothesis (the one that is guessed). Furthermore, which hypothesis is announced by the data heavily depends on the shape of the initial set of hypotheses. For instance, let us consider two simple classes of sets: $C_1 = \{\{1\}, \{2\}\}$ and $C_2 = \{\{1\}, \{1, 2\}\}$, and let h_1 be a hypothesis corresponding to the set $\{1\}$. In this case the single event of publicly announcing !1 is equivalent to announcing h_1 in case C_1 has been the initial set of hypotheses, but it does not announce h_1 when Scientist has to pick from C_2, since the other hypothesis is still possible.

3.3 Class of hypotheses

The procedure of learning starts with a class of hypotheses which represents the background knowledge of the scientist. This fact has at least two important consequences for our considerations. One is that the Scientist expects that one of them is true, and in the framework it is guaranteed that he is right—Nature indeed chooses one from the class fixed in the beginning. Second is that we cannot think of a learning scenario as of the process of simply verifying or falsifying a single hypothesis, although those two processes can be viewed as important components of identification (Gierasimczuk 2009). The fact of picking *one element from a class* is an important factor in learnability analysis. It allows considering learnability on two levels: single hypotheses and classes determined by some external properties.

3.4 Infinite procedures

The learning theory framework is defined for potentially infinite universes, but even for finite worlds the sequences of data are infinite. The reason for this is that we want to account for situations when Scientist does not know the finiteness or size of the entity he investigates. If the initial class of hypotheses is not drastically restrictive, Scientist can never know whether all the elements have already been enumerated. This leads to infinite procedures and conditions defined in the limit. Our temporal and epistemic models should reflect these properties. They should allow talking about epistemic states as invariant from some point onwards, without specifying when this happens.

3.5 Non-introspective knowledge

Learning in this framework leads to a form of knowledge—eventually Scientist declares a hypothesis that is true, he believes that it is true, and moreover he has a justification (though often very limited) to choose it. This "knowledge" is strictly operational (work-in-progress) and does not entail introspection. It is preceded by a sequence of belief changes. In general, Scientist's perspective does not allow him to point out the successful guess, he does not know whether he will not be forced to change his guess again in the light of future data.

3.6 Single agent

Although science as well as learning seem to be at least a two-player game, we are concerned here only with the role of Scientist/Learner. We assume Nature/Teacher to be an objective, uninvolved machine that makes a choice and gives out data.

3.7 Positive, true and readable data

We assume the following features of the incoming data.

1. Truthfulness, i.e., Scientist receives only true data, no false information is included. This assumption leads to the priority of incoming data over the background preference of Scientist.

2. Positiveness, i.e., Scientist receives only elements of a positive presentation (text) of the object being learned. Alternatively, together with the positive also negative information could be included (informant), e.g., for set learning the graph of the characteristic function of the set could be enumerated.
3. Readability, i.e., there is complete clarity about what information Scientist receives. A further step is to analyze the situation of uncertainty about the incoming information.

3.8 Summary

Below we sum up the above-explained features of learning in the limit.

1. Data and conclusions are of a different nature.
2. Learning is an inductive, step-by-step process.
3. Learning starts with a class of hypotheses.
4. Learning is a potentially infinite procedure, defined in the limit.
5. Learning results in operational, non-introspective knowledge.
6. We analyze single-agent learning.
7. Environments include only true, positive and readable information.

4 Learning as epistemic update

The properties of learning in the limit discussed in Sect. 3 have many important consequences for modelling learning in DEL. One of them is that we are forced to provide models which can handle two sorts of objects: pieces of incoming information and hypotheses. To establish a bridge between those two different ontologies we treat a hypothesis as the set of histories of events that it predicts, e.g., if we take a hypothesis h to be 'There are all natural numbers except 3' it predicts that the environment will enumerate all the natural numbers except 3.

The possible worlds in our epistemic model are identified with hypotheses. Unlike in the classical DEL approach, the event models are announcements of data corresponding to elements of the sets being learned, and not hypotheses themselves.

A further difference is in the number of agents. In DEL and DDL the update is defined for multi-agent epistemic cases. As we have already explained, in the present paper we are concerned only with the role of Scientist (Learner). We recognize the possibility and potential of analyzing two or more agents in the contexts of inductive inference (an example is given by Dégremont and Gierasimczuk 2009). However, for the sake of simplicity our DEL and DDL models are going to account for only one agent.

Let us again fix C to be a class of sets, and for each $S_n \in C$ we consider h_n to be a hypothesis that describes S_n. In learning by erasing we can take the initial epistemic model to represent the background knowledge of Scientist together with his uncertainty about which world is the actual one. Let us take the initial epistemic frame to be

$$M = \langle H_C, \sim \rangle,$$

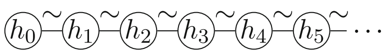

Fig. 1 Initial epistemic model

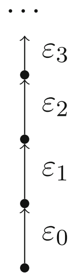

Fig. 2 Environment ε consistent with h_3

Fig. 3 Event model E_0 of the announcement of ε_0

where H_C is a possibly infinite[3] set of worlds (hypotheses that are considered possible) and $\sim \subseteq H_C \times H_C$ is an uncertainty relation for Scientist. Since we assume that the initial hypothesis space is arbitrary, we also do not require any particular preference of the scientist over H_C. Hence, we can for now assume that $\sim$ is a universal, binary equivalence relation over H_C. The initial epistemic state of the Scientist is depicted in Fig. 1. This model corresponds to the starting point of the scientific discovery process. Each world represents a hypothesis from the initial set determined by the background knowledge. In the beginning Scientist considers all of them possible. The model also reflects the fact that Scientist is given the class of hypotheses H_C. In other words he knows what the alternatives are.

Next, Nature decides on some state of the world by choosing one possibility from C. Let us assume that as a result h_3 correctly describes the chosen world. Then, she decides on some particular environment ε, consistent with Nature's choice. We picture this enumeration in Fig. 2 below.

The sequence ε is successively given to Scientist. Let us focus now on the first step of the procedure. We have the uncertainty range of the scientist, it runs through the whole set of hypotheses H_C. A piece of data ε_0 is given to the scientist. This fact can be represented by the event model $E_0 = \langle \{e\}, \rightarrow, pre \rangle$, where $e \rightarrow e$ and $pre(e) = \varepsilon_0$ (see Fig. 3).

Scientist, when confronted with the announcement of ε_0 updates his epistemic state accordingly. We will represent the process formally by the product update $M \otimes E_0$. The result of the product update is again an epistemic model $M' = \langle H_C{}', \sim' \rangle$, where:

[3] We can effectively deal with the epistemic update and identification in infinite domains by using special enumeration strategies (for explanation and examples see Gierasimczuk 2009).

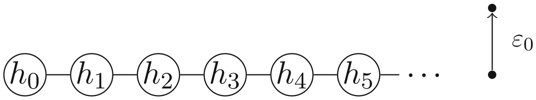

Fig. 4 Confrontation with data

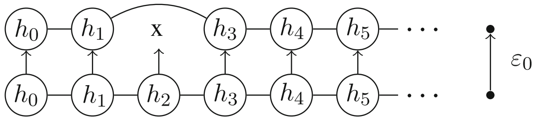

Fig. 5 Epistemic update

1. $H_C{}' = \{(h_n, e) | h_n \in H_C \& e \in |E_0| \& pre(e) \in S_n)\}$;
2. $\sim' = \sim |H_C{}'$.

We use here event models similar in spirit to those of public announcements (Batlag et al. 1998). They consist of only one state with a pre-condition determined by the piece of data that is given. In Fig. 4 Scientist's confrontation with ε_0 is depicted.

Scientist tests each hypothesis with ε_0. If a hypothesis is consistent with it, it remains as a possibility, if it is not consistent, it is eliminated (see Fig. 5). Let us assume that ε_0 is not consistent with h_2.

This epistemic update can be iterated infinitely many times along ε resulting in an infinite sequence of models which according to the lines of DEL can be called ε-Generated Epistemic Model (see e.g. van Benthem et al. 2007).

Definition 7 (Generated epistemic model)
The generated epistemic model $(M)^{\varepsilon}$, with $\varepsilon = \varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$, is the result of epistemic update $M \otimes E_0 \otimes E_1 \otimes E_2 \otimes \ldots$, where for each n, the event E_n corresponds to the announcement of ε_n.

Let us now see a simple example of finite identification of a single hypothesis in the special case of set learning.

Example 7 Let us take $H_C = \{h_1, h_2, h_3\}$, such that h_n corresponds to $\{1, ..., n\}$. Nature makes her choice regarding what the world is like. We assume that as a result h_3 holds. Then, Nature chooses an enumeration $\varepsilon = 1, 2, 1, 3, 2, \ldots$. After the first piece of data, 1, the uncertainty range of the scientist includes the whole H_C. After the second, 2, scientist eliminates h_1 since it does not contain the event 2 and now he hesitates between h_2 and h_3. The third piece, 1, does not change anything, however the next one, 3, eliminates h_2. Uncertainty is eliminated as well. He knows that the only hypothesis that can be true is h_3. Therefore, we can say that he learned it conclusively, with certainty.

The above discussion allows drawing the following conclusion.

Proposition 1 *Finite identifiability can be modelled in DEL.*

The argument for the above proposition is as follows. For the translation of the FIN framework into DEL semantics we take:

– epistemic states for hypotheses;
– infinite sequences of announcements for environments;
– epistemic update for the progress in eliminating uncertainty over the hypothesis space.

Scientist succeeds in finite identification of S from ε iff there is a finite initial segment of ε, $\varepsilon|n$, such that the domain of the $\varepsilon|n$-generated model contains only one hypothesis, h_k, and $k \in I_S$. In other words, there is a finite step of the iterated epistemic update along ε, that eliminates Scientist's uncertainty.

4.1 Learning by erasing in DDL

From Scientist's point of view the process of learning has a few components that are very important in logical modelling. The first is of course the current conjecture—a hypothesis that is considered appropriate in a given step of the procedure. The second is the set of those hypotheses that were used in the past and have already been discarded. The third part is the set of hypotheses that are still considered possible, but are for some reason less probable than the chosen one.

Let us consider the following example of a learning scenario, in which the uncertainty is never eliminated.

Example 8 As you probably observed, in Example 7 scientist was very lucky. Let us assume for a moment that nature had chosen h_2, and had fixed the enumeration $\varepsilon = 1, 2, 1, 2, 2, 2, 2, \ldots$ In this case Scientist's uncertainty can never be eliminated.[4]

This example indicates that the central element of the identification in the limit model is the unavoidable presence of uncertainty. The limiting framework allows however introducing some kind of *operational* knowledge, which is uncertainty-proof.

To model the algorithmic nature of the learning process that includes the actual guess and other not-yet-eliminated possibilities, we enrich the epistemic model with some preference relation $\leq: H_C \times H_C$. The relation $\leq$ represents some preference over the set of hypotheses, e.g., if Scientist is an occamist, the preference would be defined according to the simplicity of hypotheses. In the initial epistemic state the uncertainty of the scientist again ranges over all of H_C. This time however the class is ordered and Scientists current belief is the most preferred hypothesis. Therefore, we consider the initial epistemic state of Scientist to be:

$$M = \langle H_C, \sim, \leq \rangle.$$

The procedure of erasing hypotheses that are inconsistent with successively incoming data is the same as in the previous section. This time however let us introduce the current-guess state which is interpreted as the actual conjecture of the Scientist. It is always the one that is most preferred—the smallest one according to $\leq$. In doxastic

[4] As we are interested here in learning by erasing, we assume a suitable underlying ordering of the hypothesis space. In this case it is: h_1, h_2, h_3. However, note that this type of identification is not order independent. If the initial ordering was: h_1, h_3, h_2, then Scientist would not stabilize on the correct hypothesis.

logic a set of most preferred hypotheses is almost invariably interpreted as the one that the agent *believes* in. Let us go back to Example 8, where Nature chose a world consistent with h_2. After seeing 1 and eliminating h_1, Scientist's attention focuses on h_2, then h_2 is his current belief. It is the most preferred hypothesis, and as such can be reiterated as long as it is consistent with ε. In this particular case, since Nature chose a world consistent with h_2, it will never be contradicted, so Scientist will always be uncertain between h_2 and h_3. However, his preference directs him to believe in the correct hypothesis, without him being aware of the correctness. Therefore, we claim the following.

Proposition 2 *Learning by erasing can be modelled in DDL.*

The argument for the above proposition is as follows. For the translation we take:

- epistemic states for hypotheses;
- infinite sequences of announcements for environments;
- epistemic update for the progress in eliminating uncertainty over the hypothesis space;
- preference relation for the underlying hypothesis space;
- in each step of the procedure, the most preferred hypothesis for the actual positive guess of the learning function.

Scientist learns S by erasing from ε iff there is n such that for every $m > n$, the most preferred state of the domain of the $\varepsilon|m$-generated epistemic model is h_k, and $k \in I_S$.

5 Hypotheses as protocols. A note on a temporal view

Recently, the correspondence results between dynamic epistemic logic and epistemic temporal logic have been obtained (van Benthem et al. 2007). In this section we indicate that such a double perspective (DEL + ETL) on the dynamics of learning is very relevant for our purposes.

In formal learning theory each hypothesis from the given class is associated with the corresponding set of environments. The latter can be seen as possible "streams of events" or "histories" that may occur if the relevant hypothesis is true. A history can in its turn be represented as a branch in the tree of all possible courses of events. Accordingly, hypotheses can be viewed as trees (sets of histories).

Let us take Σ to be the set of events[5] (all possible elementary information). Accordingly, Σ^* is the set of all finite sequences over Σ (all finite strings of events). Let now us take a class of hypotheses H_C and $h_n \in H_C$. The hypothesis h_n is an epistemic temporal frame

$$\langle \Sigma, H_n, \sim \rangle,$$

where Σ is the set of events; $H_n \subseteq \Sigma^*$ is a protocol (says which sequences of events are allowed), that is closed under non-empty prefixes; and $\sim$ is a binary universal

[5] In general, Σ can be infinite. This will allow us to consider many interesting, and still effective learning cases.

relation on H_n. Such ETL frame indicates which sequences of data can be expected when the corresponding hypothesis is true. This way of thinking allows viewing the class of hypotheses H_C as a forest of such protocols.

The temporal treatment of hypotheses can give an important contribution to DEL. Typically, the epistemic updates are memoryless in the sense that each time an event occurs the epistemic model is modified and in the next step the procedure starts all over again. Temporal treatment of epistemic states themselves allows putting additional conditions on the *sequences* of events that can lead to a particular change in an epistemic model.

This view has other possible implications. Let us only mention that it allows considering properties of classes of hypotheses in the limit, in (possibly) infinite universes. An example of such hypotheses are generalized quantifiers. Some of their formal features, like verifiability and falsifiability in the limit (Gierasimczuk 2009), allow a temporal analysis. This analysis can be semantically based on the so-called number triangle representation (van Benthem 1986). Some work on learnability of generalized quantifiers exists, and relating its results to temporal validities would be very interesting.

If we turn the above-defined frame into a temporal model, we create the possibility of a temporal logic for hypothesis spaces. This is even more relevant if we recall that learnability is defined for *classes* of hypotheses. We can mention here, for instance, Angluin's tell-tale set construction (Angluin 1980) which constitutes the basic condition of learnability of a recursive class of recursive sets. Constructing a temporal logic for hypothesis spaces allows expressing this condition as a temporal validity and link it to properties of dynamic epistemic models for learning (Dégremont and Gierasimczuk 2009).

6 Conclusions and further work

We have shown that some forms of inductive inference can be modelled in dynamic epistemic logic and dynamic doxastic logic. The claim is supported by a translation of the components of learning into a two-sorted semantics for DEL and DDL. In particular, we show that DEL is an appropriate framework to simulate finite identifiability. On the other hand, learning by erasing, that requires an underlying ordering of the hypothesis space, can be formalized in DDL, where the preference relation is a standard element of any model. The restrictiveness of the DEL and DDL modelling suggests that some extension of the traditional DEL framework should be considered. In particular, other epistemic update policies can be introduced and checked for their LT counterparts.

In Sect. 5, we have proposed to view one component of the learning paradigm, hypothesis spaces, as temporal frames.[6] The continuation of that line of research would be to analyze semantic properties of hypotheses (e.g. generalized quantifiers) in a temporal framework.

[6] Learnability-related temporal logic interpreted on this structures is introduced and extensively discussed in (Dégremont and Gierasimczuk 2009).

Modal analysis of the process of learning can be also continued in the following directions:

– analyzing inductive inference process in game-theoretical terms, and discussing strategies for learning and teaching (for an example see (Gierasimczuk et al. 2009));
– studying the notion of non-introspective operational knowledge and uncertainty in inductive inference.

References

Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic, 50*(2), 510–530.
Angluin, D. (1980). Inductive inference of formal languages from positive data. *Information and Control, 45*(2), 117–135.
Baltag, A., & Moss, L. (2004). Logics for epistemic programs. *Synthese, 139*(2), 165–224.
Batlag, A., Moss, L. S., & Solecki, S. (1998). The logic of public announcements and common knowledge and private suspicions. In *Proceedings of the 7th TARK* (pp. 43–56).
Costa Florêntio, C. (2002). Learning generalized quantifiers. In *Proceedings of the 7th ESSLLI Student Session*.
Dégremont, C., & Gierasimczuk, N. (2009). Can doxastic agents learn? On the temporal structure of learning. ILLC Prepublication (PP) Series PP-2009-14, Amsterdam.
Freivalds, R., & Zeugmann, T. (1995). Co-learning of recursive languages from positive data, RIFIS. Technical report, RIFIS-TR-CS-110, RIFIS, Kyushu University 33.
Gierasimczuk, N. (2007). The problem of learning the semantics of quantifiers. In *Proceedings of the 6th TbiLLC, Vol. 4363 of LNAI* (pp. 117–126). Springer.
Gierasimczuk, N. (2009). Identification through inductive verification. In *Proceedings of the 7th TbiLLC, Vol. 5422 of LNAI* (pp. 193–205). Springer.
Gierasimczuk, N., Kurzen, L., & Velázquez-Quesada, F. (2009). Learning as interaction (manuscript).
Gold, E. (1967). Language identification in the limit. *Information and Control, 10*, 447–474.
Hintikka, J. (1962). *Knowledge and belief. An introduction to the logic of the two notions*. Ithaca: Cornell University Press.
Jain, S., Osherson, D., Royer, J. S., & Sharma, A. (1999). *Systems that learn*. Chicago: MIT Press.
Kelly, K. (1996). *The logic of reliable inquiry*. Oxford: Oxford University Press.
Lange, S., Wiehagen, R., & Zeugmann, T. (1996). Learning by erasing. In *Proceedings of the 7th international workshop on algorithmic learning theory* (pp. 228–241). Springer-Verlag.
Martin, E., & Osherson, D. (1998). *Elements of scientific inquiry*. Cambridge: MIT Press.
Osherson, D., de Jongh, D., Martin, E., & Weinstein, S. (1997). Formal learning theory. In J. van Benthem & A. Ter Meulen (Eds.), *Handbook of logic and language*. Amsterdam: North Holland.
Tiede, H.-J. (1999). Identifiability in the limit of context-free generalized quantifiers. *Journal of Language and Computation, 1*, 93–102.
van Benthem, J. (1986). *Essays in logical semantics*. Dordrecht: D. Reidel.
van Benthem, J., Gerbrandy, J., & Pacuit, E. (2007). Merging frameworks for interaction: DEL and ETL. In *Proceedings of the 11th TARK* (pp. 72–81).
van Ditmarsch, H., van der Hoek, W., & Kooi, B. (2007). *Dynamic epistemic logic*. Springer Netherlands.