Formal Semantics of Temporal Epistemic Reflection+

Wiebe van der Hoek[a,b]
John-Jules Meyer[a,b]
Jan Treur[b]

[a] Utrecht University, Department of Computer Science
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands.
Email: {wiebe,jj}@cs.ruu.nl

[b] Free University Amsterdam, Department of Mathematics and Computer Science
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands.
Email: treur@cs.vu.nl.

Abstract
In this paper we show how a formal semantics can be given to reasoning processes in meta-level architectures that reason about (object level) knowledge states and changes of them. Especially the attention is focused on the upward and downward reflections in these architectures. Temporalized epistemic logic is used to specify meta-level reasoning processes and the outcomes of these.

1 Introduction

Meta-level architectures often are used either to model dynamic control of the object level inferences, or to extend the inference relation of the object level. In [Tre92] we introduced formal semantics for meta-level architectures of the first kind based on temporal models. It may be considered quite natural that for such a dynamic type of reasoning system the temporal element of the reasoning should be made explicit in the formal semantics. For the use of meta-level architectures to extend the object level inference relation the situation looks different. In principle one may work out formal semantics in terms of (the logic behind) this extended, non-classical inference relation; e.g., as in the literature for nonmonotonic logics. However, much discussion is possible about this case. Some papers argue that also in the case of a non-monotonic logic the semantics have to make the inherent temporal element explicit; approaches are described in, e.g., [Gab82], [ET93]. In the current paper we adopt this line.

In principle a downward reflection that extends the inference relation of the object level theory disturbs the (classical) object level semantics: facts (assumptions) are added that are not logically entailed by the available object level knowledge. Adding a temporal dimension (in the spirit of [FG92]) enables one to obtain formal semantics

+ This work is partially supported by ESPRIT III Basic Research Action 6156 (DRUMS II).

of *downward reflection* in a dynamic sense: as a *transition* from the current object level theory to a next one (where the reflected assumption has been added).

In [MH93a] a metaphor of a meta-level architecture was exploited to define a nonmonotonic logic, called Epistemic Default Logic. It was shown how *upward reflection* can be formalized by a *nonmonotonic entailment relation based on epistemic states* (according to [HM84]), and the meta-level reasoning process by a (monotonic) epistemic logic. Compared to a meta-level architecture, what was still missing was a formalization of the step where the conclusions of the meta-level actually were used to change the object level, i.e., where formulas φ are added to the object level knowledge, in order to be able to reason further with them at the object level. This should be achieved by the downward reflection step. In the current paper we introduce a formalization of this downward reflection step in the reasoning pattern as well. We formalize the semantics of the architecture by means of entailment on the basis of temporalized Kripke models. Thus, a formalization is obtained of the reasoning pattern as a whole, consisting of a process of generating possible default assumptions by meta-level reasoning and actually assuming them by downward reflection (a similar pattern as generated by the so-called BMS-architecture introduced in [TT91]).

The temporal epistemic meta-level architecture described here is a powerful tool to reason with dynamic assumptions: it enables one to introduce and retract additional assumptions during the reasoning, on the basis of explicit epistemic information on the current knowledge state and meta-knowledge to determine adequate additional assumptions. This covers a whole class of reasoning patterns of meta-level architectures; we call them *temporal epistemic meta-level architectures* (TEMA). In such an architecture upward reflection is restricted to the transfer to the meta-level of epistemic information: information about what currently is known and what is not known at the object level, and downward reflection is restricted to introducing additional information (assumptions) to the object level, based on the conclusions derived at the meta-level. In this architecture a number of interesting reasoning patterns can be formalized in temporal semantics in a quite intuitive and transparent manner. In [Tre90], [Tre91] various applications of the architecture are shown, including: hypothetical reasoning, the method of indirect proof, reasoning about knowledge, reasoning about actions integrated with executing them.

The formalization of downward reflection was inspired by [Tre92], where it is pointed out how temporal models can provide an adequate semantics of meta-level architectures for dynamic control, and [ET93] where similar ideas have been worked out to obtain a linear time temporal semantics for default logic. The general idea is that conclusions derived at the meta-level essentially are statements about the state of the object level reasoning at the next moment of time. Thus, downward reflection is a shift of time in a (reasoning) process that is described by temporal logic.

To get the idea we will use the following informal description of an example reasoning pattern, involving diagnostics (of a car that cannot start) by elimination of hypotheses (for a more extensive description, see [Tre90]). Later on (in Section 7.3) we will come back to this example to show how it can be formalized by our approach. Suppose at the object level we have the following (causal) knowledge about a car:

- *if the battery is empty*
  then the lights cannot burn
- *if the battery is empty, or the sparking-plugs are tuned up badly*
  then the car does not start
- *the car does not start*

At the meta-level we control a diagnostic reasoning process to find out whether it can be excluded that an empty battery is the cause of the problems (the only hypothesis that we fully consider in this example). To this end we have the following (simplified) meta-knowledge that enables to propose hypotheses for elimination, and to reject them if indeed they turn out to fail:

- *if it is known that the car does not start*
  and it is not known whether the hypothesis "the battery is empty" holds
  then "the battery is empty" is an adequate hypothesis to focus on
- *if it is known that the car does not start*
  and it is known that the hypothesis "the battery is empty" does not hold
  and it is not known whether the hypothesis
  "the sparking-plugs are tuned up badly" holds
  then "the sparking-plugs are tuned up badly" is an adequate hypothesis to focus on
- *if we have focused on a hypothesis X*
  and assuming X we have derived that the observable Y should be the case
  and it has been observed that Y is not the case,
  then our hypothesis X should be rejected
- *it has been observed that the lights can burn*

Using these two knowledge bases in a temporal epistemic meta-level architecture we can perform the following dynamic reasoning pattern:

1. Draw all conclusions that are possible at the object-level
2. Reflect upwards that at the object level it is not known whether the battery is empty
3. Draw the conclusion at the meta-level that 'battery is empty' is an adequate hypothesis to focus on
4. Reflect this hypothesis downward: introduce it at the object level as additional information (assumption)
5. Draw the conclusion at the object level that the lights cannot burn
6. Reflect upwards the information that at the object level it has been found that the lights cannot burn
7. Use the observation at the meta-level that the lights can burn, and notice the contradictory situation
8. At the meta-level draw the conclusion that the focus hypothesis 'battery is empty' should be rejected
9. Reflect downwards that 'battery is empty' is considered to be not true.

This example shows some of the possibilities of temporal epistemic reflections. Notice especially the manner in which we treat epistemic states and additional information in a dynamic manner. This implies that during the reasoning the (states of) object level and meta-level both change in (reasoning) time; temporal logic is exploited to describe these changes in a logical manner.

Our approach provides a formalization of upward reflection by a nonmonotonic entailment relation based on epistemic states and downward reflection by an entailment relation based on temporal models. As in the literature reflections usually either are kept rather restricted (and static), or are described in a procedural or syntactic manner (e.g., by means of generalized inference rules; see [GTG93]), the novelty of our approach is that it introduces a semantic formalization for reflections involving epistemic and temporal aspects.

In this paper in Section 2 the basic notions of epistemic logic are presented. In Section 3 we do the same for temporal logic and the notion of temporalization of a logic system. In Section 4 we show how upward reflection can be formalized based on epistemic states. In Section 5 we introduce the logic used to formalize the meta-level reasoning: S5P*. In Section 6 we define a labelled branching time temporal formalization of downward reflection. In Section 7 we show how the overall formalization can be obtained from its parts and how it works for the running example reasoning pattern. In Section 8 we draw some conclusions. In this version, proofs are omitted; they can be found in the full paper [HMT94b].

2 Basic notions and properties of Epistemic Logic

2.1 Epistemic Logic

2.1. DEFINITION (epistemic formulas). Let **P** be a set of propositional constants (atoms); $\mathbf{P} = \{p_k \mid k \in I\}$, where I is either a finite or countably infinite set. The set FORM of *epistemic formulas* φ, ψ,... is the smallest set containing **P**, closed under the classical propositional connectives and the epistemic operator *K*, where $K\varphi$ means that φ is known. An *objective formula* is a formula without any occurrences of the modal operator *K*. For $\Gamma \subseteq$ FORM, we denote by Prop(Γ) the set of objective formulas in Γ.

2.2. DEFINITION (S5-Kripke models). An S5-model is a structure $\mathbf{M} = \langle M, \pi \rangle$ where M is a non-empty set, π is a truth assignment function of type $M \to (\mathbf{P} \to \{\mathbf{t}, \mathbf{f}\})$ such that for all $m_1, m_2 \in M$: $\pi(m_1) = \pi(m_2) \Rightarrow m_1 = m_2$. The class of S5-models is denoted by Mod(S5).

The set of states in an S5-model represents a collection of alternative states that are considered (equally) possible on the basis of (lack of) knowledge. We shall use S5-models as representations of the knowledge of agents.

2.3 REMARK For any $m \in M$, the function $\pi(m)\colon p \to \pi(m)(p)$ is a valuation. Since we have that in an S5-model it holds that $\pi(m_1) = \pi(m_2) \Leftrightarrow m_1 = m_2$, we may identify states m with their valuations π(m), and write $m \equiv \pi(m)$ for $m \in M$. So, without loss of generality, we may consider S5-models of the form $\mathbf{M} = \langle M, \pi \rangle$ with $M \subseteq \mathbf{W}$.

2.4. DEFINITION (submodels and union of S5-models). We define a subset relation on S5-models by: $\mathbf{M}_1 \subseteq \mathbf{M}_2$ iff $M_1 \subseteq M_2$. Moreover, if $\mathbf{M}_1 = \langle M_1, \pi_1 \rangle$ and $\mathbf{M}_2 = \langle M_2, \pi_2 \rangle$ are two S5-models, their union is defined as: $\mathbf{M}_1 \cup \mathbf{M}_2 = \langle M, \pi \rangle$, where $M = M_1 \cup M_2$ and $\pi(m) = \pi_i(m)$ if $m \in M_i$, $i = 1, 2$.

2.5. DEFINITION (interpretation of epistemic formulas). Given $\mathbf{M} = \langle M, \pi \rangle$, we define the relation $(\mathbf{M}, m) \models \varphi$ by induction on the structure of the epistemic formula φ:

$$(\mathbf{M}, m) \models p \quad \Leftrightarrow \quad \pi(m)(p) = \mathbf{t} \text{ for } p \in \mathbf{P}$$
$$(\mathbf{M}, m) \models \varphi \wedge \psi \quad \Leftrightarrow \quad (\mathbf{M}, m) \models \varphi \text{ and } (\mathbf{M}, m) \models \psi$$
$$(\mathbf{M}, m) \models \neg\varphi \quad \Leftrightarrow \quad (\mathbf{M}, m) \not\models \varphi$$
$$(\mathbf{M}, m) \models K\varphi \quad \Leftrightarrow \quad (\mathbf{M}, m') \models \varphi \text{ for all } m' \in M$$

2.6. DEFINITION (validity).

i φ is *valid in* $\mathbf{M} = \langle M, \pi \rangle$, denoted $\mathbf{M} \models \varphi$, if for all $m \in M$: $(\mathbf{M}, m) \models \varphi$.

ii φ is *valid*, notation $\text{Mod}(\mathbf{S5}) \models \varphi$, if $\mathbf{M} \models \varphi$ for all **S5**-models $\mathbf{M}$.

Validity w.r.t. S5-models can be axiomatized by the system **S5**:

2.7. DEFINITION (system **S5**). The logic **S5** consists of the following:

Axioms:

(A1) All propositional tautologies

(A2) $(K\varphi \wedge K(\varphi \rightarrow \psi)) \rightarrow K\psi$ — *Knowledge is closed under logical consequence.*

(A3) $K\varphi \rightarrow \varphi$ — *Known facts are true.*

(A4) $K\varphi \rightarrow KK\varphi$ — *An agent knows that he knows something.*

(A5) $\neg K\varphi \rightarrow K\neg K\varphi$ — *An agent knows that he does not know something.*

Derivation rules:

(R1) $\vdash \varphi, \vdash \varphi \rightarrow \psi \Rightarrow \vdash \psi$ — *Modus Ponens*

(R2) $\vdash \varphi \Rightarrow \vdash K\varphi$ — *Necessitation*

That φ is a theorem derived by the system **S5** is denoted by $\mathbf{S5} \vdash \varphi$.

2.8. THEOREM (Soundness and completeness of **S5**). $\mathbf{S5} \vdash \varphi \Leftrightarrow \text{Mod}(\mathbf{S5}) \models \varphi$

2.2 Epistemic States and Stable Sets

In this paper we simply define an epistemic state as an S5-model:

2.9. DEFINITION. An *epistemic state* is an S5-model $\mathbf{M} = \langle M, \pi \rangle$. The set M is the set of epistemic alternatives allowed by the epistemic state $\mathbf{M}$.

2.10. DEFINITION. Let $\mathbf{M}$ be an **S5** model. Then $K(\mathbf{M})$ is the set of facts known in $\mathbf{M}$: $K(\mathbf{M}) = \{\varphi \mid \mathbf{M} \models K\varphi\}$. We call $K(\mathbf{M})$ the *theory of* $\mathbf{M}$ or *knowledge in* $\mathbf{M}$.

We mention here that the knowledge in $\mathbf{M}$ are exactly the validities in $\mathbf{M}$ (Cf. [MH94]), i.e. we have $K(\mathbf{M}) = \{\varphi \mid \mathbf{M} \models K\varphi\} = \{\varphi \mid \mathbf{M} \models \varphi\}$.

2.11. LEMMA. For any S5 models $\mathbf{M}_1$ and $\mathbf{M}_2$:

i If $\mathbf{M}_1 \subseteq \mathbf{M}_2$ then $\text{Prop}(K(\mathbf{M}_2)) \subseteq \text{Prop}(K(\mathbf{M}_1))$.

ii If the set of atoms **P** is finite, then also
$\text{Prop}(K(\mathbf{M}_2)) \subseteq \text{Prop}(K(\mathbf{M}_1)) \Rightarrow \mathbf{M}_1 \subseteq \mathbf{M}_2$.

2.12. PROPOSITION (Moore [Moo85]).

i The theory $\Sigma = K(\mathbf{M})$ of an epistemic state $\mathbf{M}$ is a so-called *stable set*, i.e., satisfies the following properties:

(St 1) all instances of propositional tautologies are elements of Σ;

(St 2) if $\varphi \in \Sigma$ and $\varphi \rightarrow \psi \in \Sigma$ then $\psi \in \Sigma$;

(St 3) $\varphi \in \Sigma \Leftrightarrow K\varphi \in \Sigma$

(St 4) $\varphi \notin \Sigma \Leftrightarrow \neg K\varphi \in \Sigma$

(St 5) Σ is propositionally consistent.

ii Every stable set Σ of epistemic formulas determines an **S5**-Kripke model $\mathbf{M}_\Sigma$ for which it holds that $\Sigma = K(\mathbf{M}_\Sigma)$. Moreover, if **P** is a finite set, then $\mathbf{M}_\Sigma$ is the unique **S5**-Kripke model with this property.

2.13. PROPOSITION. A stable set is uniquely determined by its objective formulas.

3 Basic notions and properties of Temporal Logic

We start (following [FG92]) by defining the temporalized models associated to any class of models and apply it to the classes of models as previously discussed. In contrast to the reference as mentioned we use labelled flows of time. We use one fixed set L of labels, viz. $L = 2^I$, the powerset of some index set I. However, in most definitions we do not use this fact, but only refer to (elements τ of) L.

3.1 Flows of time

3.1. DEFINITION (discrete labelled flow of time).

Suppose L is a set of labels. A *(discrete) labelled flow of time* (or *lft*) , labelled by L is a pair $\mathbf{T} = (T, (<_\tau)_{\tau \in L})$ consisting of a non-empty set T of time points and a collection of binary relations $<_\tau$ on T. Here for s, t in T and τ in L the expression $s <_\tau t$ denotes that t is a (immediate) *successor* of s with respect to an arc labelled by τ. Sometimes it is convenient to leave the indices out of consideration and use just the binary relation $s < t$ denoting that $s <_\tau t$ for some τ (for some label τ they are connected). Thus we have that $< = \cup_\tau <_\tau$. We also use the (non-reflexive) transitive closure « of this binary relation: $« = <^+$.

We will make additional assumptions on the flow of time; for instance that it describes a discrete tree structure, with one root and where time branches in the direction of the future.

3.2. DEFINITION (labelled time tree)

An lft $\mathbf{T} = (T, (<_\tau)_{\tau \in L})$ is called a labelled time tree (ltt) if the following conditions are satisfied (recall that $< = \cup_\tau <_\tau$):

i   the graph ⟨T,<⟩ is a directed rooted tree.
ii   Successor existence: Time points have at least one <-successor.
iii   Label-deterministic: For every label τ there is at most one τ-successor.

3.3. DEFINITION (branch and path)

a   A *branch* in an lft $\mathbf{T}$ is an lft $\mathbf{B} = (T', (<'_{\tau})_{\tau \in L})$ with (i) $T' \subseteq T$, (ii) $s <'_{\tau} t \Rightarrow s <_{\tau} t$, (iii) every $s \in T'$ has at most one <'-successor $t \in T'$, (iv) for all $s, t \in T'$: $s <_{\tau} t \Rightarrow s <'_{\tau} t$, and (v) every element of T that is in between elements of T' is itself in T': for all $s' \in T', t \in T, u' \in T'$: s' « t « u' $\Rightarrow t \in T'$.
   b   A branch in an ltt $\mathbf{T} = (T,(<_{\tau})_{\tau \in L})$ is *maximal* if contains the root r of T.

c   A *path* is a finite sequence of successors: $s_0, \ldots, s_n$ such that: $s_i < s_{i+1}$ for all $0 \leq i \leq n-1$. We call s_0 the starting point and s_n the end point of the path.

3.4. PROPOSITION Any branch of an lft $\mathbf{T}$ is an ltt.

3.5. DEFINITION (time stamps). Given an ltt $(T, (<_{\tau})_{\tau \in L})$. A mapping $|\cdot| : T \rightarrow \mathbf{N}$ is called a *time stamp mapping* if for the root r it holds that $|r| = 0$, and for all time points s, t it holds $s < t \Rightarrow |t| = |s| + 1$.

3.2 Temporal models

We first define our temporal formulas:

3.6. DEFINITION (temporal formulas).
Given a logic **L**, *temporal formulas over* (the language of) **L** are defined as follows:
i   if φ is a formula of **L** then $C\varphi$ is a temporal formula (also called a *temporal atom*)
ii   if φ and ψ are temporal formulas, then so are:

$$\neg\varphi,\ \varphi \wedge \psi,\ \varphi \rightarrow \psi,\ X_{\exists,\tau}\varphi,\ X_{\exists}\varphi,\ X_{\forall,\tau}\varphi,\ X_{\forall}\varphi,\ F_{\exists}\varphi,\ F_{\forall}\varphi,\ G_{\exists}\varphi,\ G_{\forall}\varphi.$$

Note how the *C*-operator acts as a kind of ‘separator’ between the basic language for **L** and the actual temporal formulas: from the temporal language point of view, formulas of the form $C\varphi$ may be conceived as a kind of ‘atoms’; in the truth-definition, occurrences of the *C* enforce a ‘shift’ in the evaluation of formulas, taking us from a temporal model __ to some of it snapshots $\_\_{}_t$, which are in their turn models for **L**:

3.7. DEFINITION (temporal models)

a   Let MOD be a class of models, and $\mathbf{T} = (T, (<_{\tau})_{\tau \in L})$ a labelled flow of time. A *temporal* MOD-*model over* $\mathbf{T}$ is a mapping __: $T \rightarrow$ MOD. For $t \in T$ we sometimes denote __(t) (the snapshot at time point t) by $\_\_{}_t$.

b   If we apply a) to the classes of models ModSet(**PC**) or Mod(**S5**) we call these temporalized models *temporal valuation-set-models* (abbreviated *temporal V-models*) and *temporal* S5-*models over* $\mathbf{T}$, respectively. Similarly for the class of S5P*-models that will be introduced in Section 5.

c   Given an lft $\mathbf{T}$, the temporal formulas are interpreted on MOD-models as follows:
   i   Conjunction and implication are defined as usual; moreover
   __, $s \models \neg\varphi$ iff not __, $s \models \varphi$;

ii The temporal operators are interpreted as follows:

1) $C\varphi$ means that in the current state φ is true, i.e.
__, s ⊨ $C\varphi$ iff $\_\__s$ ⊨ φ

2) $X_{\exists,\tau}\varphi$ means that φ is true in some τ-successor state i.e.,
__, s ⊨ $X_{\exists,\tau}\varphi$ iff there exists a time point t with s $<_\tau$ t such that __, t ⊨ φ

3) $X_\exists\varphi$ means that there is a τ with some τ-successor in which φ is true.
__, s ⊨ $X_\exists\varphi$ iff there exists a time point t with s < t such that __, t ⊨ φ

4) $X_{\forall,\tau}\varphi$, meaning that φ is true in all τ-successor states, i.e.,
__, s ⊨ $X_{\forall,\tau}\varphi$ iff for all time points t with s $<_\tau$ t it holds __, t ⊨ φ

5) $X_\forall\varphi$ means that φ is true in all immediate successors:
__, s ⊨ $X_\forall\varphi$ iff for all time points t with s < t it holds __, t ⊨ φ

6) $F_\exists\varphi$ means that φ is true in some future state, i.e.,
__, s ⊨ $F_\exists\varphi$ iff there exists a time point t with s « t such that __, t ⊨ φ

7) $F_\forall\varphi$, means that for all future paths there is a time point where φ is true:
__, s ⊨ $F_\forall\varphi$ iff for all branches _ starting in s there is a t in __with __, t ⊨ φ

8) $G_\exists\varphi$ means that φ is true along some future path, i.e.,
__, s ⊨ $G_\exists\varphi$ iff there exists a branch _ starting in s,
with __, t ⊨ φ for all t in _.

9) $G_\forall\varphi$, means that φ is true all future states i.e.
__, s ⊨ $G_\forall\varphi$ iff for all time points t with s « t it holds __, t ⊨ φ.

4 Formalizing Upward Reflection Using Epistemic States

In order to let the meta-level manipulate the information that is (explicitly or implicitly) encoded at the object-level, somehow it has to be reflected upward what is known at the object-level, and what is not. The former, i.e. to reflect upward what is known, is straightforward: if an objective formula φ is true at the object-level, we simply reflect this as Kφ being true. More interestingly, we also want to reflect upward those facts that are (currently) *not* known at the object level. Moreover, we somehow want to implement the idea that the facts that are true at the object level is *all that is known* at the current time point.

The converse relation of ⊆ on Kripke models (Cf. Definition 2.4), will play an important role in the sequel. $\mathbf{M}_1 \supseteq \mathbf{M}_2$ means that the model $\mathbf{M}_2$, viewed as a representation of the knowledge of an agent, involves a *refinement* of the knowledge associated with model $\mathbf{M}_1$. This has to be understood as follows: in the model $\mathbf{M}_2$ less (or the same) states are considered possible by the agent as compared by the model $\mathbf{M}_1$. So, in the former case the agent has less doubts about the true nature of the world. It will turn out that our definitions below will work in such a way that this means that with respect to model $\mathbf{M}_2$ the agent has at least the knowledge associated with model $\mathbf{M}_1$, and possibly more. So in a transition of $\mathbf{M}_1$ to $\mathbf{M}_2$ we may say that knowledge is gained by the agent. Thus the relation '⊇' acts as an knowledge ordering on the set of S5-models.

We already noted in Section 2, that the set of states in an S5-model represents the states that are considered possible on the basis of (lack of) knowledge). We also mentioned, that the validities of such a model exactly determine what was called an epistemic state. On the basis of such epistemic states, Halpern & Moses define an entailment relation $\mid\!\sim$ with which one can infer what is known, and, more importantly, what is unknown in such epistemic states.

4.1. DEFINITION. Given a set $M \subseteq \mathbf{W}$, we define the *associated S5-model* $\Phi(M)$, given by $\Phi(M) = \langle M, \pi \rangle$ with $\pi: M \times \mathbf{P} \to \{\mathbf{t}, \mathbf{f}\}$ such that $\pi: (m, p) \mapsto m(p)$.

4.2. DEFINITION. Given some objective formula φ, we define M_φ as the set of valuations satisfying φ, i.e., $M_\varphi = \{m \in \mathbf{W} \mid m \models \varphi\}$. We denote the epistemic state $\Phi(M_\varphi)$ associated with M_φ by $\mathbf{M}_\varphi$.

4.3. PROPOSITION. $\mathbf{M}_\varphi = \bigcup\{\mathbf{M} \mid \mathbf{M} \models \varphi\} = \bigcup\{\mathbf{M} \mid \mathbf{M} \models K\varphi\}$.

Thus, in order to get $\mathbf{M}_\varphi$, we can consider all S5-models of φ and take their union to obtain one 'big' S5-model. We denote the mapping $\varphi \mapsto \mathbf{M}_\varphi$ by μ: $\mu(\varphi) = \mathbf{M}_\varphi$.

4.4. DEFINITION (Nonmonotonic epistemic entailment).
For $\varphi \in$ Prop(FORM), and $\psi \in$ FORM: $\varphi \mid\!\sim \psi \Leftrightarrow \psi \in K(\mathbf{M}_\varphi)$.

Informally, this means that ψ is entailed by φ, if ψ is contained in the theory (knowledge) of the "largest S5-model" $\mathbf{M}_\varphi$ of φ. Halpern & Moses showed in [HM84] that this "largest model" need not always be a model of φ itself if we allow φ to contain epistemic operators. However, in our case where we only use objective formulas φ, $\mathbf{M}_\varphi$ is always the largest model for φ. Moreover, Halpern & Moses have shown that in this case the theory $K(\mathbf{M}_\varphi)$ of this largest model is a stable set that contains φ and such that for all stable sets Σ containing φ it holds that $\text{Prop}(K(\mathbf{M}_\varphi)) \subseteq \text{Prop}(\Sigma)$, thus $K(\mathbf{M}_\varphi)$ is the "propositionally least" stable set that contains φ. So $\mid\!\sim$ can also be viewed as a *preferential entailment* relation in the sense of Shoham [Sho88], where, in our paper, the preferred models of φ are the largest ones, viz. $\mathbf{M}_\varphi$, where the least objective knowledge is available.

We denote the mapping $\varphi \mapsto K(\mathbf{M}_\varphi)$ by κ: $\kappa(\varphi) = K(\mathbf{M}_\varphi)$, the stable set associated with knowing only φ. Alternatively viewed, $\kappa(\varphi)$ is the $\mid\!\sim$-closure of φ. Note that since $\kappa(\varphi) = K(\mathbf{M}_\varphi)$ is a stable set, it is also propositionally closed. We now give a few examples to show how the entailment $\mid\!\sim$ works: Let p and q be two distinct primitive propositions. Then:

$p \mid\!\sim K(p \vee q)$	$p \mid\!\sim \neg Kq$
$p \mid\!\sim Kp \wedge \neg Kq$	$p \mid\!\sim \neg K\neg q$
$p \wedge q \mid\!\sim K(p \wedge q) \wedge Kp \wedge Kq$	$p \mid\!\sim \neg K(p \wedge q)$
$p \vee q \mid\!\sim K(p \vee q)$	$p \vee q \mid\!\sim \neg Kp \wedge \neg Kq$

Obviously, the entailment relation $\mid\!\sim$ is nonmonotonic; for instance, we have $p \mid\!\sim \neg Kq$, while *not* $p \wedge q \mid\!\sim \neg Kq$; it even holds that $p \wedge q \mid\!\sim Kq$.

5 The Meta-level: The Epistemic Preference Logic S5P*

The "upward reflection" entailment relation $\mid\!\sim$ enables us to derive information about what is known and what is not known. In this section we show how we can use this information to perform meta-level reasoning. To this end we extend our language with operators that indicate that something is a *possible assumption* to be introduced at the object level and thus has a different epistemic status than a *certain fact*. In this way the proverbial "make an assumption" is not made directly in the logic, but a somewhat more cautious approach is taken. The "assuming" itself is part of the downward reflection, to be discussed in Section 6.

Let I be a finite set of indexes. The logic **S5P** (introduced in [MH91] and developed further in [MH92, MH93a,b]) is an extension of the epistemic logic **S5** by means of special modal operators P_i denoting *possible assumption* (w.r.t. situation or frame of mind i), for $i \in I$, and also generalisations P_τ, for $\tau \subseteq I$. Informally, $P_i\varphi$ is read as "φ is a possible assumption (within frame of reference i)". As we shall see below, a frame of reference (or mind) refers to a preferred subset of the whole set S of epistemic alternatives. This operator is very close to the PA-operator of [TT91] and the D-operator of [Doh91]. The generalisation $P_\tau\varphi$ is then read as a possible assumption with respect to the (intersection of the) frames of reference occurring in τ. Also, we have an operator K to denote what is *known* and an operator B to describe what is true (*believed*) *under the hypothesis (under focus)*.

The logic **S5P*** is an extension of **S5P** by allowing an arbitrary set A of additional symbols that denote primitive meta-level propositions; for the moment, one may think of them as a way to express that certain propositions have been observed, or that a given hypothesis is 'in focus'. From a logical point of view, such assertions are just atoms, whose truth is governed by some 'meta truth assignment function'. It may well be, that upon closer examination, there are some logical laws steering the truth of such atoms, but in this stage we will not investigate the inner structure of the given meta-level propositions.

Formally, S5P*-formulas are interpreted on Kripke-structures (called S5P*-models) of the form

$$\mathbf{M} = \langle U, M, \pi, \{M_i \mid i \in I\}, MV\rangle,$$

where:

- U is a collection of states (universe), and $M \subseteq U$ is non-empty (the current focus).
- π: $U \times \mathbf{P} \rightarrow \{\mathbf{t}, \mathbf{f}\}$ is a truth assignment to the primitive propositions per world
- $M_i \subseteq M$ ($i \in I$) are sets ('frames') of preferred worlds
- MV: $A \rightarrow \{\mathbf{t}, \mathbf{f}\}$ is a valuation for the additional primitive meta-level propositions.

When writing $\mathbf{M}_1 \subseteq \mathbf{M}_2$ we mean that the set of states of $\mathbf{M}_1$ is a subset of those of $\mathbf{M}_2$. Again we identify states s and their truth assignments $\pi(s)$. We let Mod(**S5P***) denote the collection of Kripke-structures of the above form. Given an S5P*-model $\mathbf{M}$ $= \langle U, M, \pi, \{M_i \mid i \in I\}, MV\rangle$, we call the S5-model $\mathbf{M}_M = \langle M, \pi|M\rangle$ the *focused S5-reduct* of $\mathbf{M}$ and $\mathbf{M}_U = \langle U, \pi|U\rangle$ the *universal* or *general S5-reduct* of $\mathbf{M}$.

5.1. DEFINITION (interpretation of S5P*-formulas).
Given a model $\mathbf{M} = \langle U, M, \pi, \{M_i \mid i \in I\}, MV\rangle$, we give the following truth definition. Let φ be a formula, $\alpha \in A$ and $m \in U$. The cases in which $\varphi \in \mathbf{P}$, $\varphi = (\varphi_1 \wedge \varphi_2)$ or $\varphi = \neg\psi$ are dealt with as in Definition 2.5, for the other cases are as follows:

$(\mathbf{M}, m) \models K\varphi$ iff for all $m' \in U$, $(\mathbf{M}, m') \models \varphi$;
$(\mathbf{M}, m) \models B\varphi$ iff for all $m' \in M$, $(\mathbf{M}, m') \models \varphi$;
$(\mathbf{M}, m) \models P_i\varphi$ iff for all $m' \in M_i$, $(\mathbf{M}, m') \models \varphi$;
$(\mathbf{M}, m) \models P_\tau\varphi$ iff for all $m' \in M_\tau$, $(\mathbf{M}, m') \models \varphi$, where $M_\tau = \bigcap_{i\in\tau} M_i$ and $\tau \subseteq I$.
$(\mathbf{M}, m) \models \alpha$ iff $MV(\alpha) = \mathbf{t}$, for α in A.

We see that the clauses state that $P_i\varphi$ is true if φ is a possible assumption w.r.t. subframe M_i, whereas the latter says that $P_\tau\varphi$ is true if φ is a possible assumption w.r.t. the intersection of the subframes M_i, $i \in \tau$. This intersection is denoted by M_τ. We assume that, for $\tau = \emptyset$, $M_\emptyset = \bigcap_{i\in\emptyset} M_i = M$. So in this special case we get that the P_τ modality coincides with the belief operator B. Validity and satisfiability is defined analogously as before.

It is possible to axiomatize (the theory of) **S5P*** by adjusting the axiom system **S5P** of [MH93b]; we will not go into full details here but give some main principles:

5.2. DEFINITION (system **S5P***). In the following, i ranges over I, and τ over subsets of I. Moreover, $\Box$, $\Box_1$ and $\Box_2$ are variables over $\{K, B, P_i, P_\tau \mid i \in I, \tau \subseteq I\}$.

(B1) All propositional tautologies;

(B2) $(\Box\varphi \wedge \Box(\varphi \to \psi)) \to \Box\psi$;
(B3) $\Box\varphi \to \Box\Box\varphi$;
(B4) $\neg\Box\varphi \to \Box\neg\Box\varphi$;

(B5) $K\varphi \to \varphi$;
(B6) $K\varphi \to \Box\varphi$;

(B7) $\neg\Box_1\bot \to (\Box_1\Box_2\varphi \leftrightarrow \Box_2\varphi)$;

(B8) $P_i\varphi \leftrightarrow P_{\{i\}}\varphi$
(B9) $P_\tau\varphi \to P_{\tau'}\varphi$ $\quad \tau \subseteq \tau'$
(B10) $P_\emptyset\varphi \leftrightarrow B\varphi$
(B11) $\neg B\bot$

(R1) Modus Ponens
(R2) Necessitation for K: $\vdash \varphi \Rightarrow \vdash K\varphi$.

We call the resulting system **S5P***. In the sequel we will write $\Gamma \vdash_{\mathbf{S5P*}} \varphi$ or $\varphi \in \mathrm{Th}_{\mathbf{S5P*}}(\Gamma)$ to indicate that φ is an **S5P***-consequence of Γ. We mean this in the more liberal sense: it is possible to derive φ from the assertions in Γ by means of the axioms and rules of the system **S5P***, *including the necessitation rule*. (So, in effect we consider the assertions in Γ as additional axioms: $\Gamma \vdash_{\mathbf{S5P*}} \varphi$ iff $\vdash_{\mathbf{S5P*} \cup \Gamma} \varphi$, cf. [MH93b])

5.3. THEOREM $\Gamma \vdash_{\mathbf{S5P*}} \varphi \Leftrightarrow$ (for all $\mathbf{M} \in \mathrm{Mod}(\mathbf{S5P*})$: $\mathbf{M} \models \Gamma \Rightarrow \mathbf{M} \models \varphi$)
PROOF. Combine the arguments given in [MH92, MH93b].

6 A Temporal Formalization of Downward Reflection

In the previous sections it has been described how the upward reflection can be formalized by a (nonmonotonic) inference based on epistemic states, and the meta-level process by a (monotonic) epistemic logic. In the current section we will introduce a formalization of the downward reflection step in the reasoning pattern. The meta-level reasoning can be viewed as the part of the reasoning pattern where it is determined what the possibilities are for additional assumptions to be made, based on which information is available at the object level and which is not. The outcome at the meta-level concerns conclusions of the form Pφ, where φ is an object-level formula. What is missing still is the step where the assumptions are actually made, i.e., where such formulas φ are added to the object level knowledge, in order to be able to reason further with them at the object level. This is what should be achieved by the downward reflection step. Thus the reasoning pattern as a whole consists of a process of generating possible assumptions and actually assuming them.

By these downward reflections at the object level a hypothetical world description is created (as a refinement or revision of the previous one). This means that in principle not all knowledge available at the object level can be derived already from the object level theory OT: downward reflection creates an essential modification of the object level theory. Therefore it is excluded to model downward reflection according to reflection rules as sometimes can be found in the literature, e.g., “If at the meta-level it is provable that Provable(φ) then at the object level φ is provable” (Cf. [Wey80]):

$$\frac{\mathrm{MT} \vdash \mathrm{Provable}(\varphi)}{\mathrm{OT} \vdash \varphi}$$

A reflection rule like this can only be used in a correct manner if the meta-theory about provability gives a sincere axiomatization of the object level proof system, and in that case by downward reflection nothing can be added to the object level that was not already derivable from the object level theory. Since we essentially extend or modify the object level theory, such an approach cannot serve our purposes here.

In fact, a line of reasoning at the object level is modelled by inferences from subsequently chosen theories instead of inferences from one fixed theory. In principle a downward reflection realizes a shift or transition from one theory to another. In [GTG93] such a shift between theories is formalized by using an explicit parameter referring to the specific theory (called ‘context’ in their terms) that is concerned, and by specifying relations between theories. In their case downward reflection rules (‘bridge rules’ in their terms) may have the form:

$$\frac{MT \vdash \text{Provable}(OT', \varphi)}{OT' \vdash \varphi}$$

or, in their notation

$$\frac{< Th("\varphi", "OT'"), MT >}{< \varphi, OT'>}$$

Here, the second element of the pair denotes the context in which the formula that is the first element holds. At the meta-level, knowledge is available to derive conclusions about provability relations concerning a variety of object level theories OT. So, if at the object level from a (current) theory OT some conclusions have been derived, and these conclusions have been transformed to the meta-level, then the meta-level may derive conclusions about provability from another object level theory OT'. Subsequently one can continue the object level reasoning from this new object level theory OT'. The shift from OT to OT' is introduced by use of the above reflection rule.

In the approach as adopted here we give a temporal interpretation to these shifts between theories. This can be accomplished by formalizing downward reflection by temporal logic (as in [Tre92]). In a simplified case, where no branching is taken into account, the temporal axiom ($CP\varphi \rightarrow X\varphi$) can be used to formalize downward reflection, for every objective formula φ.

In the general case we want to take into account branching and the role to be played by an index τ in $P_\tau\varphi$. We will use this index τ to label branches in the set of time points. By combining S5P* with the temporal logic obtained in this manner we obtain a formalization of the whole reasoning pattern.

During the reasoning process we modify the information we have at the object level, and accordingly change the focus set M of possible worlds. We can formulate this property as follows:

6.1. DEFINITION. A temporal S5P*-model *obeys downward reflection* if the following holds for any s and τ :
the cluster M_τ in $\_\__s$ is non-empty $\Leftrightarrow$ there is a t with $s <_\tau t$ and for all such t the focus set of states M of $\_\__t$ equals M_τ

Now we are ready to zoom in into the models we like to consider here, the temporal epistemic meta-level architecture (TEMA-) models.

6.2. DEFINITION (TEMA- models)
A TEMA-model __ is a temporal S5P*-model over an lft $\mathbf{T}$ such that:

i $\mathbf{T}$ is a labelled time tree;
ii For every time point s, there is exactly one t with $s <_\emptyset t$;
iii __ obeys downward reflection.
iv __ is conservative: if $s < t$ then $(\_\__s)_U \supseteq (\_\__t)_U$

Notice that conservatism refers to the universe U. Sometimes also M is shrinking during the reasoning process: in case of an accumulation of assumptions that are never retracted (e.g., see [HMT93]), but in general M may vary arbitrarily within U.

6.3. THEOREM.
TEMA-models have the following validities:

- T0 All the operators of $\{X_{\forall,\tau}, X_{\forall}, F_{\forall}, G_{\forall}\}$ satisfy the K-axiom (C too) and generalisation;
- T1 $\vdash_{S5P*} \varphi \Rightarrow \models_{TEMA} C\varphi$ (introduction of C)
- T2 $\neg X_{\forall}\bot$ (successor existence)
- T3 $X_{\exists,\tau}\varphi \leftrightarrow X_{\forall,\tau}\varphi$ (label-deterministic)
- T4 $X_{\forall,\tau}\varphi \leftrightarrow \neg X_{\exists,\tau}\neg\varphi$ (duality)
- T5 $X_{\forall}\varphi \leftrightarrow \neg X_{\exists}\neg\varphi$ (duality)
- T6 $X_{\forall}\varphi \leftrightarrow \bigwedge_{\tau\subseteq I} X_{\forall,\tau}\varphi$ ($<$ is union of $<_\tau$)
- T7 $X_{\exists}\varphi \leftrightarrow \bigvee_{\tau\subseteq I} X_{\exists,\tau}\varphi$ (dual of T6)
- T8 $C(\neg P_\tau\bot \wedge P_\tau\varphi) \leftrightarrow X_{\exists,\tau}CK\varphi$, if φ is objective (allowing downward reflection)
- T9 $G_{\forall}\varphi \rightarrow X_{\forall}\varphi$ ($< \subseteq «$)
- T10 $G_{\forall}\varphi \rightarrow X_{\forall}G_{\forall}\varphi$ (since « is transitive closure of $<$)
- T11 $G_{\forall}(\varphi \rightarrow X_{\forall}\varphi) \rightarrow (X_{\forall}\varphi \rightarrow G_{\forall}\varphi)$ (induction)
- T12 $(C\varphi \rightarrow X_{\forall}C\varphi) \wedge (CK\varphi \rightarrow X_{\forall}CK\varphi)$, if φ is objective (conservativity)
- T13 $CK\varphi \rightarrow G_{\forall}CK\varphi$ (from conservativity and induction)

6.4. REMARK. The Theorem above says that the formulas T1 - T13 are at least *sound*; yet we have not been concerned by designing a *complete* logic for TEMA-models.

7 Overall Formalization

In this section we will show how the different parts of the reasoning pattern as described in previous sections can be combined.

7.1 EMA-theories and EMA-entailment

In the language of S5P* we can express meta-knowledge. By combining the formal apparatus of S5P* with Halpern & Moses' nonmonotonic epistemic entailment we obtain a framework in which we can perform static epistemic reasoning: reasoning about the current epistemic state without reflecting the conclusions downwards. We call this framework Epistemic Meta-level Architecture (EMA).

7.1. DEFINITION (EMA-theory). An *EMA-theory* Θ is a pair (W, Δ), where W is a finite, consistent set of objective (i.e. non-modal) formulas describing facts about the world, and Δ is a finite set of S5P*-formulas. The sets W and Δ are to be considered as sets of axioms; we may apply necessitation to them.

7.2. DEFINITION (EMA-entailment). Given an EMA-theory $\Theta = (W, \Delta)$, we define the nonmonotonic inference relation $\mid\!\sim_{\Theta}$ as follows. Let W* be the conjunction of the formulas in W, and let φ be an objective formula. Then we define the *EMA-entailment relation* $\mid\!\sim_{\Theta}$ w.r.t. Θ as follows: $\varphi \mid\!\sim_{\Theta} \psi \Leftrightarrow_{def} \kappa(\varphi \wedge W^*) \cup \Delta \vdash_{\mathbf{S5P*}} \psi$.

7.3. EXAMPLE (*Hypothesis elimination*).
We come back to our example in the introduction. Let the EMA-theory $\Theta = (W, \Delta)$ be defined by

$$W = \{e \rightarrow \neg b\}$$
$$\Delta = \{K\neg s \rightarrow ((K\neg e \wedge \neg K\neg e) \rightarrow P_1 e), \texttt{initially\_}\neg s, \texttt{initially\_z} \rightarrow Kz\},$$

where e means "empty battery" and b "lights can burn". Moreover, we have used primitive meta-level propositions "`initially_z`" to indicate that the information "z" is present initially, that is, before the reasoning process starts. Now we have the following:

$$W^* \mid\!\sim \neg Ke \wedge \neg K\neg e \quad \text{and} \quad \Delta \cup \{\neg Ke \wedge \neg K\neg e\} \vdash_{\mathbf{S5P*}} P_1 e$$

Therefore $\mid\!\sim_{\Theta} P_1 e$

In [HMT94a] we show how this temporal framework also covers default reasoning. Also, we think our framework of temporalizing **S5P*** can easily be adapted to give an account of counterfactual reasoning (as treated in [MH93c]) too.

7.2 TEMA-models and TEMA-entailment

7.4. DEFINITION Let $\Theta = (W, \Delta)$ be an EMA-theory. Then we define a *TEMA-model of* Θ as a TEMA-model $\_\_^{\Theta}$ such that:

i (basis: the root) $\_\_^{\Theta}_r$ is an S5P*-model such that
  (a) the universal S5-reduct of $\_\_^{\Theta}_r$ is the S5-model $\mathbf{M}_{W^*}$,
  (b) $\_\_^{\Theta}_r$ satisfies the meta-knowledge, i.e., $\_\_^{\Theta}_r \models \Delta$.

ii (induction step) Suppose that we are given an S5P*-model at snapshot $\_\_^{\Theta}_s$. Then we have that for a(n S5P*-) model $\_\_^{\Theta}_t$ with $s <_{\tau} t$, it holds that:
  (a) $(\_\_^{\Theta}_t)_U$ is the S5-model $\mathbf{M}_{\tau}$ as it appeared as a cluster in $\_\_^{\Theta}_s$, and
  (b) $\_\_^{\Theta}_t$ satisfies the meta-knowledge again, i.e., $\_\_^{\Theta}_t \models \Delta$.

In general, there are multiple TEMA-models of an EMA-theory Θ. Note that clause ii(a) reflects the downward reflection operation with respect to the P_{τ}-assumptions.

Even in less trivial reasoning processes we may be interested in some kind of final outcome (a conclusion set). In our case the universe U always shrinks, but not always the focus set M does so. We can view U as a core of derived facts that after all survives during the reasoning pattern; this is reflected in the following definition:

7.5. DEFINITION (limit model).
Suppose __ is a conservative temporal V-model. The intersection of the models __(s) for all s in a given branch $\_ = (T, (<_{\tau})_{\tau \in L})$ of the lft $\mathbf{T}$ is called the *limit model* of the branch, denoted $\lim_{\underline{B}}$ __. The set of limit models for all branches is called the *set*

of limit models of __. These definitions straightforwardly extend to temporal S5- and S5P-models, by identifying **M** with its set of states, U.

We might formulate definitions of entailment of objective formulae related to any model, or any model based on the standard tree. But it may well happen that there are branches in such models, for instance labelled by the empty set only, that contain no additional information as compared to the background knowledge. It is not always realistic to base entailment on such informationally poor branches in a model. Thus we define:

7.6. DEFINITION (informationally maximal)
We define for branches $\_\_{}_1$ and $\_\_{}_2$ with the same set of time stamps of a TEMA-model __ that $\_\_{}_2$ is *informationally larger* than $\_\_{}_1$, denoted $\_\_{}_1 \leq \_\_{}_2$, if for all $i \in \mathbf{N}$ and s, t with $|s| = |t| = i$ it holds $\_\_{}_2(s) \subseteq \_\_{}_1(t)$.
We call $\_\_{}_1$ *informationally maximal* if it is itself the only branch of __ that is informationally larger.

7.7. DEFINITION (regular model)
A TEMA-model __ is called regular if all branches are informationally maximal. The submodel based on all time points t included in at least one informationally maximal branch is called the *regular core* of __, denoted by reg(__).

7.8. DEFINITION
i for $k \in N$ we define $\_\_^{(k)} = \cup_{t \in \text{reg}(\_\_),\, |t| = k} \_\_(t)$
ii $\_\_^{\omega} = \cap_{k \in N} \_\_^{(k)}$.

7.9. THEOREM (From [HMT94a]). Let **M** be a TEMA-model. Then:
i For $k \leq k'$ we have $\_\_^{(k')} \subseteq \_\_^{(k)}$
ii $\_\_^{\omega} = \cap_{\_\_ \text{ branch of reg}(\_\_)} \lim_{\_\_} \_\_$.

Since (regular) TEMA models describe a reasoning process over time, it seems natural to have notions of entailment that exploit the conservativity of such models, expressing the monotonic growth of knowledge of objective formulas. The latter restriction is important, since it may be the case that initially an atom p is not known, yielding $\neg K$p and hence $K\neg K$p, while at some later point p has been learnt, giving $\neg K\neg K$p, expressing that some knowledge (i.e., knowledge about ignorance!) has been lost. So, in the sequel we will be interested in formulas of the form $CK\varphi$, where φ is an objective formula. We will call such formulas *currently known objective formulas* (cko’s) and use α and β as variables over them.

In the literature on non-monotonic reasoning, one usually distinguishes so called *sceptical* (true in *all* obtained models) and *credulous* (true in *some* of them) notions of entailment. Due to the fact that we have an (infinite) branching time structure, we have a great variety of combining these notions, although for the formulas that we are interested in, the current objective formulas, various of such notions do collapse. This observation is made explicit in the following theorem.

7.10. THEOREM Let α and β be cko-formulas. Then, on TEMA models, we have the following equivalences:

i	$G_{\exists}\alpha \equiv \alpha$	iii	$G_{\forall}\alpha \equiv X_{\forall}\alpha$
ii	$\alpha \equiv (\alpha \wedge G_{\forall}\alpha)$	iv	$F_{\exists}\alpha \equiv F_{\exists}(G_{\forall}\alpha \wedge \alpha)$

7.11. DEFINITION (sceptical entailment) Let __ be a TEMA-model with root r and α an cko-formula. We define the *sceptical entailment relation* by:

__ $|\!\!\approx_{scep} \alpha$ iff __,r $\models F_{\forall}\alpha$

Due to our remark that we made before definition 7.6, it makes most sense to use this notion of sceptical entailment for (sub) models of type reg(__).

7.12. PROPOSITION Let __ be a TEMA-model with root r and α a cko-formula.
The following are equivalent:

i reg(__) $|\!\!\approx_{scep} \alpha$

ii $\lim_{\mathbf{B}}$ __ $\models \alpha$ for every maximal branch _ of the regular core of __.

For our definition of credulous entailment we can be less restrictive; we do not need to bother about informationally maximal branches. Especially, too little information in one branch can always be overcome by another, informationally larger branch.

7.13. DEFINITION (credulous entailment) Let __ be a TEMA-model and α a cko-formula. We define: __ $|\!\!\approx_{cred} \alpha$ iff __,r $\models F_{\exists}\alpha$

7.14. PROPOSITION
Let __ be a TEMA-model and α a cko-formula. The following are equivalent:

i __ $|\!\!\approx_{cred} \alpha$

ii $\lim_{\mathbf{B}}$ __ $\models \alpha$ for some maximal branch $\mathbf{B}$

We finally can give the definitions of sceptical and credulous entailment.

7.15. DEFINITION (entailment from an EMA-theory)
Let $\Theta = (W, \Delta)$ be an EMA-theory and φ an objective formula.

$\Theta |\!\!\approx_{scep} \varphi$ iff for all models __ of Θ it holds reg(__) $|\!\!\approx_{scep} \varphi$

$\Theta |\!\!\approx_{cred} \varphi$ iff for all models __ of Θ it holds __ $|\!\!\approx_{cred} \varphi$

7.16. PROPOSITION
Let $\Theta = (W, \Delta)$ be an EMA-theory and φ an objective formula. Then it holds

$$\Theta |\!\!\approx_{scep} \varphi \Rightarrow \Theta |\!\!\approx_{cred} \varphi$$

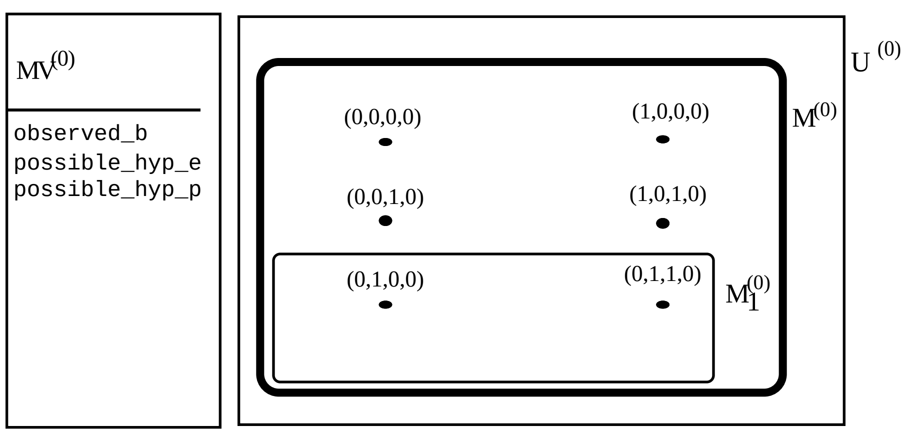
7.3 The example reasoning pattern

We return to our running example and show how it obtains its natural semantics. To cover the whole reasoning pattern, we take W = {(e ∨ p) → ¬s, e → ¬b} and Δ as:

Δ = {*K*¬s → ((¬*K*e ∧ ¬*K*¬e) → P_1e), *if ¬s is known,*
and we do not know whether ¬e holds, we may choose e as an hypothesis
(*K*¬s ∧ *K*¬e) → ((¬*K*p ∧ ¬*K*¬p) → P_2p), *we investigate p if we know ¬s and ¬e*
`observed_b`, `initially_¬s`,
`initially_z` → Kz,
`possible_hyp_e`, `possible_hyp_p`,
(`possible_hyp_z` ∧ P_iz) → $X_{\exists,i}$ `focus_z`, z ∈ {e, p}, i ∈ {1, 2}
(`focus_z` ∧ `observed_w` ∧ *B*¬w) → `bad_hyp_z`, z ∈ {e, p}, w ∈ {b}
(`bad-hyp_z` ∧ `focus_z`) → $G_\forall$*K*¬z z ∈ {e, p}}

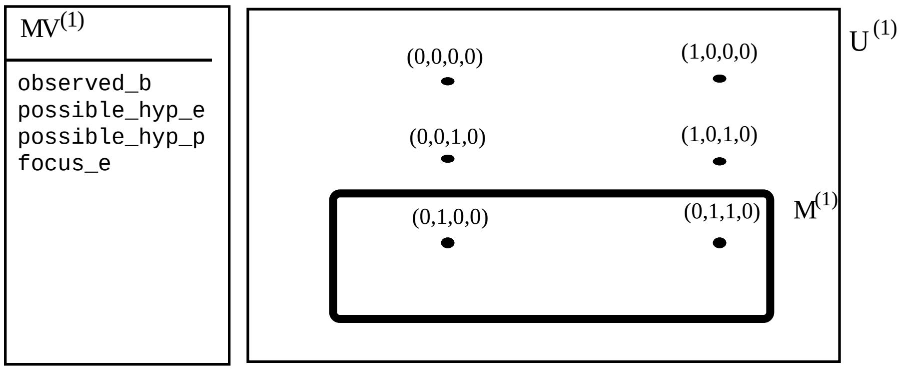
We denote worlds by 4-tuples (w, x, y, z) ∈ $\{0,1\}^4$, to be interpreted as the truth-values of the tuple (b, e, p, s). We construct a model that is linear.

* *Initial snapshot*
We obtain the initial model $\_\_^{(0)}$ = ⟨$U^{(0)}$,$M^{(0)}$,$\pi^{(0)}$,$M_1^{(0)}$, $M_2^{(0)}$, $MV^{(0)}$⟩. The universe $U^{(0)}$ is based on all models of W; we take $M^{(0)} = U^{(0)}$. Furthermore, we have seen in Example 7.3, how we can derive P_1e from (W, Δ), defining the cluster $M_1^{(0)}$. In this model $M_2^{(0)} = \varnothing$, (Cf. our remarks at the end of this section); $MV^{(0)}$ is to be understood as the minimal valuation on the atoms of A so that Δ is satisfied;



* *Next snapshot*
Using T8, *C*(¬P_1⊥ ∧ P_1e) ↔ $X_{\exists,1}$*CK*e, we obtain a downward reflection into the next model $\_\_^{(1)}$ = ⟨$U^{(1)}$,$M^{(1)}$,$\pi^{(1)}$,$M_1^{(1)}$, $M_2^{(1)}$, $MV^{(1)}$⟩, where $MV^{(1)}$ is like

$MV^{(0)}$, but under $MV^{(1)}$, now also `focus_e` is true. Moreover, $U^{(1)} = U^{(0)}$, and the new cluster $M^{(1)}$ equals the old $M_1^{(0)}$:



* *Further snapshots*
Now, the cluster $M^{(1)}$ is the one which is in the current focus: at its object-level theory we can derive that ¬b, since in __$^{(1)}$ we have __$^{(1)} \models B\neg$b. We use the meta-knowledge (`focus_z` ∧ `observed_`w ∧ $B\neg$w) → `bad_hyp_z` with e for z and b for w. Using Δ again this yields $G_\forall K\neg$e, so that for all __$^{(2)}$ with __$^{(1)}$ < __$^{(2)}$ it holds that __$^{(2)} \models \neg$e; semantically this amounts to saying that the universe $U^{(2)}$ of all future models __$^{(2)}$ is at most the set $U^{(1)} \setminus M^{(1)}$, and one may proceed and investigate the hypothesis of bad plugs (p) for the failure of the non-starting car.

Some remarks are in order here. Firstly, we have only analyzed some *intended* model for the theory Θ = (W, Δ); a model that carries the assumption that the theory (W, Δ) is *all* that we know; this justified for example that we took $M_2^{(1)}$ to be empty. Secondly, since the atoms in A are rather application-dependent, we decided to add all assumptions about them to Δ; however, for classes of applications, one might consider to add some of those properties as axioms to S5P*. Finally, in this particular example we did not exploit the fact that we have a *branching* time model; one may change the example in such a way that the two possible hypotheses p and e are so to speak investigated *simultaneously*. In our example there may be specific reasons to investigate one hypothesis before the other: because one of them (i.e., e) is more easy to refute, for example; this preference is also explicitly modelled in Δ.

8 Conclusions

In [MH93a,b] an Epistemic Default Logic (EDL) was introduced inspired by the notion of meta-level architecture that also was the basis for the BMS-approach introduced in [TT91]. In EDL drawing a default conclusion has no other semantics than that of adding a modal formula to the meta-level. No downward reflection takes place to be able to reason with the default conclusions at the object level (by means of which assumptions actually can be made). In [TT91] downward reflection to actually make assumptions takes place, but no logical formalization was given: it was defined only in a procedural manner.

In principle a downward reflection that enlarges the object level theory disturbs the object level semantics: facts are added that are not logically entailed by the available object level knowledge. Adding a temporal dimension (in the spirit of [FG92]) enables one to obtain formal semantics of downward reflection in a dynamic sense: as a transition from the current object level theory to a next one (where the reflected assumption has been added). This view, also underlying the work presented in [ET93] and [Tre92], turns out to be very fruitful. A number of notions can be formalized in temporal semantics in a quite intuitive and transparent manner.

The temporal epistemic meta-level architecture described here is a powerful tool to reason with dynamic assumptions: it enables one to introduce and retract additional assumptions during the reasoning, on the basis of explicit meta-knowledge. In [Tre90], [Tre91] various applications of this architecture are shown. In the current paper we formalized the semantics of a temporal epistemic meta-level architecture by means of entailment on the basis of temporalized Kripke models.

Acknowledgements

Discussions about the subject with Joeri Engelfriet have played a stimulating role for the development of the material presented here.

REFERENCES

[Doh91] P. Doherty, NM3 - A Three-Valued Cumulative Non-Monotonic Formalism, in: *Logics in AI*, (J. van Eijck, ed.), LNCS 478, Springer, Berlin, 1991, pp. 196-211.

[ET93] J. Engelfriet, J. Treur, A Temporal Model Theory for Default Logic, in: M. Clarke, R. Kruse, S. Moral (eds.), Proc. ECSQARU '93, Springer Verlag, 1993, pp. 91-96.

[FG92] M. Finger, D. Gabbay, Adding a temporal dimension to a logic system, Journal of Logic, Language and Computation 1 (1992), pp. 203-233.

[Gab82] D.M. Gabbay, Intuitionistic basis for non-monotonic logic, In: G. Goos, J. Hartmanis (eds.), 6th Conference on Automated Deduction, LNCS 138, Springer Verlag, 1982, pp.260-273

[GTG93] E. Giunchiglia, P. Traverso, F. Giunchiglia, Multi-Context Systems as a Specification Framework for Complex Reasoning Systems, in: J. Treur, Th. Wetter (eds.), Formal Specification of Complex Reasoning Systems, Ellis Horwood, 1993, pp. 45-72

[HM84] J.Y. Halpern & Y.O. Moses, Towards a Theory of Knowledge and Ignorance, *Proc. Workshop on Non-Monotonic Reasoning*, 1984, pp. 125-143.

[HMT94a] W. van der Hoek, J.-J. Ch. Meyer & J. Treur, Temporalizing Epistemic Default Logic. Techn. Report, Free University Amsterdam, 1994

[HMT94b] W. van der Hoek, J.-J. Ch. Meyer & J. Treur, Formal Semantics of Temporal Epistemic Reflection. Techn. Report, Free University Amsterdam, 1994.

[MH91] J.-J.Ch. Meyer & W. van der Hoek, Non-Monotonic Reasoning by Monotonic Means, in: J. van Eijck (ed.), *Logics in AI (Proc. JELIA '90)*, LNCS 478, Springer, 1991, pp. 399-411.

[MH92] J.-J.Ch. Meyer & W. van der Hoek, A Modal Logic for Nonmonotonic Reasoning, in: W. van der Hoek, J.-J. Ch. Meyer, Y.H. Tan & C. Witteveen (eds.), *Non-Monotonic Reasoning and Partial Semantics*, Ellis Horwood, Chichester, 1992, pp. 37-77.

[MH93a] J.-J. Ch. Meyer & W. van der Hoek, An Epistemic Logic for Defeasible Reasoning using a Meta-Level Architecture Metaphor. Techn. Report IR-329, Free University, Amsterdam, 1993.

[MH93b] J.-J. Ch. Meyer & W. van der Hoek, A Default Logic Based on Epistemic States. in: M. Clarke, R. Kruse, S. Moral (eds.), Proc. ECSQARU '93, Springer Verlag, 1993, pp. 265-273. Full version to appear in Fundamenta Informatica.

[MH93c] J.-J. Ch. Meyer & W. van der Hoek, Counterfactual reasoning by (means of) defaults in: *Annals of Mathematics and Artificial Intelligence* 9 (III-IV), (1993), pp. 345-360.

[MH94] J.-J. Ch. Meyer & W. van der Hoek, *Epistemic Logic for AI and Computer Science*, forthcoming.

[Moo85] R.C. Moore, Semantical Considerations on Nonmonotonic Logic, *Artificial Intelligence* 25, 1985, pp. 75-94.

[Sho88] Y. Shoham, *Reasoning about Change*, MIT Press, Cambridge, 1988.

[TT91] Y.-H. Tan & J. Treur, A Bi-Modular Approach to Non-Monotonic Reasoning, in: Proc. WOCFAI'91 (M. DeGlas & D. Gabbay, eds.), Paris, 1991, pp. 461-475.

[Tre90] J. Treur, Modelling nonclassical reasoning patterns by interacting reasoning modules, Report IR-236, Free University Amsterdam, 1990.

[Tre91] J. Treur, Interaction types and chemistry of generic task models, in: B. Gaines, M. Linster (eds), Proc. EKAW'91, GMD Studien 211, 1992, Bonn.

[Tre92] J. Treur, Towards dynamic and temporal semantics of meta-level architectures, Report IR-321, Free University Amsterdam, 1992. Also: Proc. META'94, 1994 (this volume).

[Wey80] R.W. Weyhrauch, Prolegomena to a Theory of Mechanized Formal Reasoning, Artificial Intelligence J. 13 (1980), pp. 133-170.