Modal Logics for Representing Incoherent Knowledge

J.J.Ch. Meyer and W. van der Hoek

Technical Report UU-CS-1995-19
May 1995

Department of Computer Science
Utrecht University
P.O.Box 80.089
3508 TB Utrecht
The Netherlands

ISSN: 0924–3275

MODAL LOGICS FOR REPRESENTING INCOHERENT KNOWLEDGE[1]

J.-J. Ch. Meyer[2] *& W. van der Hoek, Department of Computer Science, Utrecht University*

ABSTRACT

In this paper we review ways of representing incoherent 'knowledge' in a consistent way, where the use of modal logic and Kripke-style semantics is put central. Starting with a presentation of the basic modal framework, we discuss the basic modal systems **K**, **KD** (with an excursion to the representation of conflicting norms in deontic logic) and Chellas' minimal modal logic **D**. Next we look at the epistemic logics **KD45**, **S4** and **S5**, including the logical omniscience problem and several non-standard modal logics to overcome this problem. After this we turn to the issue of reasoning by default, where a conflict of defaults (or default beliefs) may arise. We give an epistemic treatment of default reasoning, and treat the way conflicts of defaults can be solved viewed from the more general perspective of resolving conflicts in meta-level reasoning. Furthermore, special attention is paid to specificity in default reasoning as a principle to solve these conflicts, for which we develop an extension of Halpern & Moses' theory of honest formulas. Finally, we discuss several numerical modal logics in their capacity of ways of representation of incoherent information.

INTRODUCTION

In everyday-life situations we often have to deal with incoherent information. From one source we learn a certain piece of information, while from another one we get some information that contradicts this. Examples range from human conversion in groups to electronic agents with multiple sensors and computer systems with multiple (communicating) intelligent (artificial) agents (so-called "*multi-agent systems*"). In order to cope with this incoherent information we need to be able to represent and reason with this in a non-trivial way. Representing incoherences cannot be done non-trivially in classical logic, since here when an inconsistency arises one may derive anything using (properties of the) material implication and modus ponens.

In the literature several non-classical logics have been proposed to overcome this problem. One class of logics in which one can represent inconsistent information as a local inconsistency, e.g., $p \wedge \neg p$, without being able to draw an arbitrary conclusion q, are the so-called paraconsistent logics. This is the subject of another chapter in this handbook. Here, we will consider modal logic and show that these, too, can be used to represent incoherent information

[1] This work is partially supported by ESPRIT III BRA project No. 6156 "DRUMS2", ESPRIT III BR Working Group No. 8319 "MODELAGE", and the Free University Amsterdam.

[2] This author is also partially supported by Nijmegen University.

by employing some suitable modalities. Furthermore, we restrict ourselves to propositional logics. That is, we assume a given set **P** of *atomic propositions*, also called *primitive propositions* or just *atoms*. The propositional language induced by **P** and the usual classical connectives '$\wedge$', '$\vee$', '$\neg$', etc. is then enlarged by the possibility to use the (primary) modal operator '$\Box$', generally pronounced "necessarily". We let '$\bot$' stand for the constant denoting falsehood, 'T' for '$\neg\bot$' (truth) and we sometimes use the (secondary) modal operator '$\Diamond$' as an abbreviation of '$\neg\Box\neg$'. $\Diamond$ is pronounced as "possibly".

1. BASIC MODAL LOGIC AND INCOHERENCY

1.1. The systems K and KD

The system **K** consists of the following axioms and rules: ($\Box$ stands for the primary modality)

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(Taut) the tautologies of propositional logic (or just enough of them)

(MP) $\varphi, \varphi \to \psi / \psi$

(Nec) $\varphi / \Box\varphi$

System **K** is the smallest modal logic that admits so-called Kripke-semantics (also called *possible world* semantics). This type of semantics is based on Kripke models, i.e., models of the kind (S, π, R), where S is a non-empty set of possible worlds, π is a truth assignment function to the atoms (in **P**) for every possible world in S, and R is an accessibility relation. R(s, t) means that the world t is accessible (or held possible, in whatever sense the interpretation of the modality dictates) from the world s. Propositions are interpreted in such a model M = (S, π, R) and a state s $\in$ S. All clauses except for the modal one follow the usual pattern of propositional logic. The primary modal operator '$\Box$' is then given a formal interpretation by (M, s) $\models \Box\varphi$ iff (M, t) $\models \varphi$ for all t with R(s, t). That is, in a possible world s it holds that "necessarily φ" iff φ holds in all worlds t that are accessible (held possible) from world s. It is easily verified that this implies that (M, s) $\models \Diamond\varphi$ iff (M, t) $\models \varphi$ for some t with R(s, t), that is, in world s formula φ is possible if there is a world t that is held possible from s and that satisfies φ. A formula is *valid in a Kripke model* if it holds in *all* states of the model, and it is called *valid* (period) if it is valid in *all* Kripke models. A formula is *satisfiable* if it is true in *some* world of *some* Kripke model. One can now show that validity corresponds to derivability in the system **K** above (see e.g. [Che80], [MH95]).

Modal logics that are based on Kripke models are called *normal* modal logics.

In system **K** we have as a theorem:

$(\Box\wedge)\quad \Box(\varphi \wedge \psi) \leftrightarrow (\Box\varphi \wedge \Box\psi)$

However, we do *not* have that $\neg\Box(\varphi \wedge \neg\varphi)$ is a theorem. This means that, for instance, $\Box(\varphi \wedge \neg\varphi)$ (or, equivalently, $\Box\varphi \wedge \Box\neg\varphi$) is satisfiable: it is true in any world (of any model) in which no world t is possible. Of course, $\varphi \wedge \neg\varphi$ is not satisfiable in **K**. This implies that we can safeguard the incoherence $\varphi \wedge \neg\varphi$ against a true inconsistency in the logic by shielding it by the modal operator $\Box$. If we would interpret this modality as knowledge or belief, we would indeed be able to represent incoherent knowledge or belief(s) in a consistent way. However, this is not customary: knowledge is generally assumed to have the property that $\Box\varphi \rightarrow \varphi$, rendering $\Box(\varphi \wedge \neg\varphi)$ inconsistent immediately, while for belief mostly (at least) the principle (D) is assumed:

(D) $\quad \neg\Box\bot$

The system **K**, augmented with (D), is called **KD**. Semantically, principle (D) amounts to considering Kripke models $M = (S, \pi, R)$, in which the relation R is serial, that is to say, for every s there is a t with R(s, t). More precisely, one can prove that the system **KD** is sound and complete with respect to these "serial models" (cf. e.g. [Che80], [MH95]).

Principle (D) is equivalent with $\neg\Box(\varphi \wedge \neg\varphi)$, so that both $\Box(\varphi \wedge \neg\varphi)$ and $\Box\varphi \wedge \Box\neg\varphi$ are not satisfiable in **KD**. In other words, in **KD** both the principles (D) and

(D*) $\quad \neg(\Box\varphi \wedge \Box\neg\varphi)$

are valid. This then implies that for *epistemic* notions such as knowledge and belief we have to look for another solution to the problem of representing incoherences. However, we can say more in the present context about *moral* dilemmas, since the logic of (moral) obligations is generally taken to be **KD**. We discuss this in the next section.

1.2. Moral dilemmas: deontic logic and norm conflicts

In deontic logic one is concerned with reasoning about norms, or rather normative *versus* non-normative situations or behaviour of agents. In this branch of modal logic the primary modal operator $\Box$ is generally interpreted as obligation and written as O. $O\varphi$ means that "φ is obligatory", or "it ought to be that φ", or, again alternatively, "ideally, φ ought to hold". In the literature there is a bit of controversy whether in $O\varphi$, the argument φ denotes a desired state-of-affairs or rather an action to be performed. In the former case one speaks of ought-to-be, in the latter of ought-to-do. For our discussion we only consider the former interpretation, since this is closer to the issues we will discuss in the sequel with respect to other modal logics.

However, in order to find nice common-sense examples we will be a bit sloppy with this distinction.

Standard deontic logic (**SDL**) is generally taken to be the modal logic **KD**. In fact, the principle (D) derives its name from *d*eontics. In the context of deontic logic this principle, $\neg O\bot$, states that an impossible state is not obliged. As we have seen before, in **KD** this principle is equivalent with $\neg O(\varphi \wedge \neg\varphi)$ and (D*) $\neg(O\varphi \wedge O\neg\varphi)$, the latter stating that one cannot be both obliged to be in a state where φ holds and one where φ does not hold.

This presents a problem in the case that we are faced with moral dilemmas, or norm conflicts. Classical examples are like "one ought to kill the enemy to defend one's country, while, on the other hand, religious principles command us not to kill". Here, on the one hand we have something like $O\varphi \wedge O\neg\varphi$, which in **KD** is inconsistent. Of course, one easy way out to be able to represent this example is just to drop the principle (D) and consider the logic **K**. In **K**, $O\varphi \wedge O\neg\varphi$ is consistent. However, for most deontic logicians it is not enough to just represent a moral dilemma in a consistent way; one wants to offer a solution to it: what should the agent in question do when faced with such a moral dilemma. This aspect appeared so important that one usually is not even interested in representing the dilemma consistently. One just uses **KD** (or some logic based on this) and finds oneself in a situation where there appears a real inconsistency in the representation, which should be solved. Interestingly, in deontic logic literature this gave rise to some of the first work on defeasibility (e.g. [AM81]), so popular in AI research nowadays. Here we will not pursue this route, but rather look at another direction: Chellas proposed a logic in which one can represent norm conflicts without abandoning the principle (D), which was held by him "relatively uncontroversial" ([Che80], p. 201).

1.3. The minimal modal logic D (Chellas)

Chellas ([Che80]) proposes a logic, which he called **D**, in which one can deny validity of (D*) while still retaining that of (D). Obviously, as we have seen in Section 1.1, this means that **D** cannot be a normal modal logic in the sense that it admits the usual Kripke semantics. The reason, of course, being that in such a normal modal logic (D*) inevitably follows from (D) via principle ($\Box\wedge$). In fact, as Chellas shows, the logic **D** has a different kind of semantics based on so-called *minimal models*.

A minimal model is a structure $M = (S, \pi, \mathcal{N})$, where S is a set of possible worlds again, π is the usual truth assignment function on the possible worlds, and $\mathcal{N}: S \rightarrow 2^{2^S}$ is a function from S to sets of subsets of S. Intuitively, $\mathcal{N}(s)$ gives us the set of propositions (regarded as sets of possible worlds, which we shall refer to as *clusters*) that are "standards of obligation" to s, that ideally should (but actually need not) be true in s. Furthermore, the function $\mathcal{N}$ satisfies the following properties:

(m) for all $s \in S$ and all clusters X and Y: if $X \cap Y \in \mathcal{N}(s)$ then $X \in \mathcal{N}(s)$ and $Y \in \mathcal{N}(s)$.
(p) for all $s \in S$: $\emptyset \notin \mathcal{N}(s)$.

Note that (m) is equivalent with
(m') for all clusters X and Y: if $X \in \mathcal{N}(s)$ and $X \subseteq Y$ then $Y \in \mathcal{N}(s)$,
expressing closure under supersets.

The truth condition for $O_C\varphi$ is now given by: $(M, s) \models O_C\varphi$ iff $\|\varphi\|^M \in \mathcal{N}(s)$, where the truth set $\|\varphi\|^M$ is defined by $\|\varphi\|^M = \{t \in S \mid (M, t) \models \varphi\}$. Under the assumption (m), the condition $\|\varphi\|^M \in \mathcal{N}(s)$ is equivalent with the condition: for *some* $X \in \mathcal{N}(s)$ for all $t \in X$: $(M, t) \models \varphi$. The condition (p) now guarantees that (D) is valid, while (m)—stating the closure of $\mathcal{N}(s)$ under supersets—ensures the validity of

($\Box\wedge\rightarrow$) $\quad O_C(\varphi \wedge \psi) \rightarrow (O_C\varphi \wedge O_C\psi)$

which is one half of ($\Box\wedge$). Moreover, the other half of ($\Box\wedge$),

($\Box\wedge\leftarrow$) $\quad (O_C\varphi \wedge O_C\psi) \rightarrow O_C(\varphi \wedge \psi)$

and thus also ($\Box\wedge$) itself, as well as

(D*) $\quad \neg(O_C\varphi \wedge O_C\neg\varphi)$, and

($\Box\neg\bot$) $\quad O_C\neg\bot$

are *no* validities in the logic **D**. The logic **D** is axiomatized by:

(D) $\neg O_C\bot$
(Taut) the tautologies of propositional logic (or just enough of them)
(MP) $\varphi, \varphi \rightarrow \psi \,/\, \psi$
($\Box$M) $\varphi \rightarrow \psi \,/\, O_C\varphi \rightarrow O_C\psi$

(cf. [Che80]).

Thus the logic **D** provides us with a means to consistently represent moral dilemmas, i.e. incoherent norms, without giving up the principle (D). (Here we will not pursue the long and troublesome history of deontic logic in which the adequacy of systems like **D** and **KD** is challenged as a logic of moral / deontic reasoning which goes way beyond the question whether principle (D) is a desirable property for such a logic. See e.g. [MW93].)

In fact, we shall see below in Section 2.4, that Chellas' logic **D**, in the semantics of which a set of *clusters* rather than just a set of *worlds* is associated with a world s, is very close to recent proposals to overcome the problem of representing incoherent knowledge in epistemic logic.

2. KNOWLEDGE AND BELIEF

2.1. The systems KD45, S4 and S5

Since the seminal work by Halpern et al. (e.g. [HM85]) modal epistemic logic has become a popular approach to representing and reasoning about knowledge and belief in AI applications. (Modal) epistemic logic dates back to the work of the philosopher J. Hintikka ([Hin62]). In this logic the primary modal operator $\Box$ is interpreted as a knowledge or belief operator, and then written as K or B, respectively. In AI the logic of knowledge is generally taken to be the system **S5** (but in philosophical logic mostly the weaker system **S4** is taken), while belief is mostly axiomatized by the system **KD45**. These systems are extensions of the basic normal modal logic **K** of Section 1.1: **S4** is obtained by taking **K** together with

(T) $\Box\varphi \to \varphi$

(4) $\Box\varphi \to \Box\Box\varphi$,

while **S5** is obtained by adding

(5) $\neg\Box\varphi \to \Box\neg\Box\varphi$

on top of **S4**. Finally, **KD45** is obtained by augmenting the system **KD** by (4) and (5). The interpretation of (T) when applied for the knowledge modality is that knowledge is true. Note that (T) is strictly stronger than (D). (4) is called the *positive introspection* axiom: it says that if something is known (believed) then it is also known (believed) that it is known (believed). On the other hand, (5) is called the *negative introspection* axiom: if something is not known (believed), then it is known (believed) that it is not known (believed). Obviously, negative introspection is far more controversial for knowledge than belief, and this is the reason why most philosophers prefer **S4** rather than **S5** as the logic of knowledge.

Semantically, one obtains (Kripke) models for these logics by putting restrictions on the accessibility relations again:

for **S4**, one considers models in which these relations are *reflexive* and *transitive*;
for **S5**, one considers models in which these relations are *equivalence* relations, or, alternatively, (*simple*) models in which the relations are *universal*, i.e., $R = S \times S$; and
for **KD45**, one considers models in which these relations are *serial*, *transitive* and *euclidean*.
(A binary relation R is euclidean if the following holds: for all s, t and u: R(s, t) and R(s, u) implies R(t, u).)

As stated above, in AI usually the logic **S5** is adopted as the logic of knowledge. **S5**-models are very simple (or at least they can be viewed in this way): they are just a set of classical valuation functions, viewed as possible worlds which are all connected by the accessibility relation. The set of formulas true in such a model (the theory of an S5-model) enjoys some very nice properties: they are so-called *stable* sets, satisfying:

(St 1) all instances of propositional tautologies are elements of Σ;
(St 2) if $\varphi \in \Sigma$ and $\varphi \to \psi \in \Sigma$ 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 propositional consistent.

Stable sets are uniquely determined by the objective (viz. non-modal) formulas they contain. Furthermore, stable sets are not properly contained in each other. Stable sets are representations of the knowledge of rational, introspective agents: their knowledge is closed under propositional calculus and under (positive and negative) introspection: both knowledge and ignorance of formulas is known by the agent. In particular, if φ is a consistent objective formula, there exists a stable set Σ^φ such that this set contains φ and is "informationally minimal" under this requirement (so that it contains the least possible objective knowledge to accommodate to the presence of φ as well as the closure under propositional logic and the introspective properties) (cf. [HM84], [MH95]). Thus Σ^φ represents the knowledge (or *epistemic state*) of a rational, introspective agent "when s/he only knows φ". Σ^φ can be characterized alternatively by the theory of the "largest **S5**-model of φ": Σ^φ is the set of formulas that are valid in the **S5**-model $\mathbb{M}_\varphi = \bigcup \{\mathbb{M} \text{ is (simple) } \mathbf{S5}\text{-model} \mid \mathbb{M} \models \varphi\}$. Halpern & Moses ([HM84]) also give an algorithm to determine the set Σ^φ for an objective formula φ. This algorithm is given by:

$$\psi \in \Sigma^\varphi \;\Leftrightarrow\; \models_{S5} (K\varphi \wedge \chi_\varphi(\psi)) \to \psi,$$

where $\models_{S5}$ denotes S5-validity, and $\chi_{\varphi}(\psi)$ is the conjunction of all subformulas $K\psi'$ of ψ for which $\psi' \in \Sigma^{\varphi}$ and formulas $\neg K\psi''$ for all subformulas $K\psi''$ of ψ for which $\psi'' \notin \Sigma^{\varphi}$ (where ψ is considered to be a subformula of itself). This algorithm decides for any formula ψ in the language whether or not it is an element of Σ^{φ}: we need only employ the algorithm for all strict subformulas of ψ, and then use the decision procedure for S5. (Actually, Halpern & Moses considered the more general case of so-called *honest* formulas φ, which includes the class of objective formulas, defined as exactly that class of epistemic formulas for which there exists a stable set Σ^{φ} such that this set contains φ and is minimal under this requirement , cf. [HM84] and [MH95]. We will return to this in Section 3.3, where we shall use a refinement of this notion of honesty.)

2.2. The logical omniscience problem

Although the logics **KD45**, **S4** and **S5** have very appealing properties, they contain a couple of validities which are sometimes viewed as troublesome when reasoning in actual cases about knowledge and, particularly, belief. These validities are called the *paradoxes of logical omniscience*, since they state that the agent's knowledge (belief) satisfies (too) idealized principles. They include the following set:

(LO1)	$\Box\varphi \wedge \Box(\varphi \rightarrow \psi) \rightarrow \Box\psi$	(Closure under implication).
(LO2)	$\models \varphi \Rightarrow \models \Box\varphi$	(Belief of valid formulas).
(LO3)	$\models \varphi \rightarrow \psi \Rightarrow \models \Box\varphi \rightarrow \Box\psi$	(Closure under valid implication).
(LO4)	$\models \varphi \leftrightarrow \psi \Rightarrow \models \Box\varphi \leftrightarrow \Box\psi$	(Belief of equivalent formulas).
(LO5)	$(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$	(Closure under conjunction).
(LO6)	$\Box\varphi \rightarrow \Box(\varphi \vee \psi)$	(Weakening of belief).
(LO7)	$\Box\varphi \rightarrow \neg\Box\neg\varphi$	(Consistency of beliefs).
(LO8)	$\Box(\Box\varphi \rightarrow \varphi)$	(Belief of having no false beliefs).
(LO9)	$\Box\neg\bot$	(Believing truth).

(In the brief statement following these principles above we have used the notion of *belief* rather than *knowledge*, since they are particularly salient and controversial with respect to belief. Moreover, in this paper we are concerned with the representation of incoherences. Strictly speaking, this cannot be incoherent knowledge, since knowledge is true by definition—the (T)-axiom, and we do not admit incoherent facts to be true in the actual world. So, also from this perspective it is better to look at belief rather than knowledge, as beliefs need not to be true in the actual world, so that the problem of representing incoherent *belief(s)* is less hopeless than that of incoherent *knowledge*, which cannot exist by our very notion of knowledge!)

Of these principles we recognize some as being the very core of (normal) modal logic, notably (LO1), which is equivalent with axiom (K); (LO2), which is the necessitation rule (Nec); and (LO9) which follows directly from (Nec). On the other hand, we even recognize some essentials from the minimal logic **D**, viz. (LO3), which is nothing else than (□M), and which has (LO4) as its immediate consequence. (LO5), of course, reminds us of the discussion in Section 1.3, where we argued that avoiding its validity (as it appears in the logic **K**, and so also in all logics **KD**, **S4**, **S5** and **KD45**) was one of the very reasons why the logic **D** was proposed. In fact, validity of (LO5), together with the principle (D), which is present (or derivable) in all of these logics, is responsible for the fact that one cannot represent incoherent knowledge, precisely as in the discussion about incoherent norms in the system **KD**. So, changing to (an epistemic variant of the) minimal logic **D** would solve this problem here as well. In fact, this is exactly what Fagin & Halpern do in [FH88], which we shall see below in Section 2.4. However, some authors try to keep normal modal logic (in the sense that ordinary Kripke semantics is employed) as their basis, necessarily "polluted" by some non-standard features like "non-modal" operators. Still others use even more radical methods than going to a non-normal (i.e., minimal) modal logic to avoid the paradoxes of logical omniscience and to include the possibility to represent incoherent belief in their logic in a consistent way. We shall see this in the sequel, where we shall concentrate on the latter issue rather than the former. (In [MH95] it is shown to what extent these approaches succeed to overcome the other problems of logical omniscience with regard to belief.)

2.3. A syntactic solution: principles (Van der Hoek & Meyer)

One of the most naive solutions to representing incoherent belief, while sticking to normal Kripke semantics as a basis, is the introduction of a special operator P (separate from the basic (normal) modal operator B for belief), which simply states that its argument is a belief that is not subject to any doubt, viz. a kind of *principle* or, interpreted more negatively, a kind of *prejudice*. Belief is then based on the old (normal) **KD45**-notion of belief as well as these principles. Although, of course, the old **KD45** notion does still not admit incoherent beliefs, these can be then just put in by the P-notion of belief. (In fact, this notion was inspired by work by Fagin & Halpern [FH88] where they used a similar operator to *keep things out* of the belief set, whereas here we use to do just the opposite.)

So, formally we define a new notion B_{HM} of belief:

$$(B_{HM}) \qquad B_{HM}\varphi \leftrightarrow_{def} B\varphi \vee P\varphi,$$

where B is the old **KD45**-notion from Section 2.1. Semantically, we just add a function $\mathcal{P}$ to the usual KD45-models, so that we consider models of the form M = (S, π, R, $\mathcal{P}$), where S and π are as usual, R is serial, transitive and euclidean, and $\mathcal{P}$ is a function mapping each world to a set of formulas (which denote the principles / prejudices in that world). Interpretation of the language is as usual (including the B-operator as based on the relation R), and the P-operator is now interpreted as: $(\mathbb{M}, s) \models P\varphi$ iff $\varphi \in \mathcal{P}(s)$. So, we see that although we stick to normal Kripke semantics as much as possible, we have the function as a non-standard element. But, as we have seen, this is inevitable: we have to "pollute" clean normal Kripke-semantics with some other elements to enable the representation of incoherent belief.

If we define validity and satisfiability in the usual way, we can now state that the formulas $\neg(B_{HM}p \wedge B_{HM}\neg p)$ and $\neg B_{HM}(p \wedge \neg p)$ are *not* valid, so that their negations (D*) *as well as (D)* become satisfiable. (E.g. to satisfy $B_{HM}(p \wedge \neg p)$ in a world s, simply choose $p \wedge \neg p \in P(s)$.) This means that we can indeed represent incoherent belief(s) in a consistent way in this logic.

2.4. The logic of local reasoning (Fagin & Halpern)

Fagin & Halpern [FH88] proposed a logic of "local reasoning" to cater for incoherent beliefs which, in retrospect, may be viewed as based on Chellas' minimal logic **D**, but adjusted for belief. They use "cluster" models of the kind $(S, \pi, \mathcal{C})$, where S and π are as usual, and $\mathcal{C}$ is a function from worlds to sets of subsets of S: for every s, $\mathcal{C}(s)$ is a non-empty collection of *non-empty* subsets (*clusters*) of S. (In fact, again they gave a multi-agent logic, but here we concentrate on the single agent case.) Then they distinguish between a weak (B_{FH}) and a strong ($B_{FH}{}^+$) notion of belief, which they provide with the following formal semantics:

$(M, s) \models B_{FH}\varphi$ iff there is *some* cluster $T \in \mathcal{C}(s)$ such that for all $t \in T$: $(M, t) \models \varphi$, and
$(M, s) \models B_{FH}{}^+\varphi$ iff for *all* clusters $T \in \mathcal{C}(s)$ and for all $t \in T$: $(M, t) \models \varphi$.

For our present purposes the notion of weak belief B_{FH} is the most interesting. Here we recognize Chellas' minimal modal logic. In fact, the way Fagin & Halpern gave their definition of the interpretation of the B_{FH}-modality corresponds exactly to that of the O-operator in the logic **D** (if we assume condition (m)), as we have seen. We also recognize Chellas' requirement (p) in the condition above that $\emptyset \notin \mathcal{C}(s)$. (Note furthermore that by the direct definition of the semantics of B_{FH}, Fagin & Halpern directly by-pass the requirement (m) that Chellas has to put on his function $\mathcal{N}$, so that this requirement is not necessary for the function $\mathcal{C}$. In some sense, by the definition above, this requirement is built-in implicitly here.)

Of course, the logic of local reasoning concerns the notion of belief, so that one would like to have the properties of belief. As to the principle (D), this is obviously valid (since it was already in the minimal modal logic **D**). But we need also to consider the introspective properties (4) and (5). To this end, Fagin and Halpern put the following restrictions on the function $\mathcal{C}$:

($\underline{4}$) for all s, s' $\in$ S: s' $\in$ T $\in$ $\mathcal{C}$(s) implies T $\in$ $\mathcal{C}$(s')
($\underline{5}$) for all s $\in$ S, T $\in$ $\mathcal{C}$(s), t $\in$ T : $\mathcal{C}$(t) $\subseteq$ $\mathcal{C}$(s)

One may verify that ($\underline{4}$) validates (4), while (5) becomes valid by imposing ($\underline{5}$) (cf. [FH88]), so that the logic of local reasoning is a true epistemic (or rather *doxastic*, since it involves *belief* rather than knowledge) variant of the minimal modal logic **D**.

If validity and satisfiability are defined as usual, one may check that—exactly as in Chellas' logic **D**—the principle (D*) $\neg(B_{FH}p \wedge B_{FH}\neg p)$ is not valid while the principle (D) $\neg B_{FH}(p \wedge \neg p)$ *is*, so that $B_{FH}p \wedge B_{FH}\neg p$ is satisfiable in the logic of local reasoning, while $B_{FH}(p \wedge \neg p)$ is *not*.

Interestingly, the logic of local reasoning can also be linked with the work of Rescher & Brandom ([RB80]) on reasoning with inconsistencies. Rescher & Brandom consider what they call "non-standard possible worlds" to represent inconsistencies. These non-standard possible worlds are in some sense "macro-worlds" consisting of a number of standard ("micro"-) worlds that are "fused together" in two possible ways, viz. "world-conjunction" (or "schematization") and "world-disjunction" (or "superposition"). In the former method, a formula is true in the macro-world if it is true in *all* micro-worlds it contains; in the latter it is true if it is true in *some* micro-world. Now, Fagin & Halpern's logic of local reasoning may be viewed as a logic based on Rescher & Brandom's macro-worlds that are fused together by world-conjunction. Later we shall see that the other method of using macro-worlds as proposed by Rescher & Brandom is also present in the epistemic logic literature (cf. Section 2.6).

2.5. The logic S5P (Meyer & Van der Hoek)

The logic **S5P** was introduced in [MH91,92] and developed further in [MH93] to model the monotonic part of (epistemic) default reasoning that deals with plausible assumptions. The logic consists of an **S5**-based logic of (certain) knowledge K combined with a number of **K45**-based modalities P_i to denote some plausible working beliefs (in view of available defaults). **S5P**-models are models of the kind $(S, \pi, R, S_1,..., S_n)$, where S and π are as usual, R is universal, and the $S_i \subseteq S$ denote *preferred* subsets of S, which we call *(sub)frames (of reference)* or *contexts*. These subsets S_i are allowed to be empty. The **S5P**-model $(S, \pi, R, S_1,..., S_n)$ is called an **S5P**-*extension* of the included simple **S5**-model (S, π, R). In these models formulas are interpreted as follows: $(M, s) \models K\varphi$ iff $(M, t) \models \varphi$ for all (t with R(s, t), i.e., all) $t \in S$, as is usual in epistemic logic and $(M, s) \models P_i\varphi$ iff $(M, t) \models \varphi$ for all $t \in S_i$. Validity and satisfiability are defined as usual.

This then yields an **S5**-logic with respect to the modality K and a **K45**-logic for each of the modalities P_i. Note that, although of course it is not possible to represent inconsistent knowledge in an **S5**-setting, it *is* possible to represent incoherent beliefs by means of the P_i-modalities (even with the use of only *one* such operator P_i!). In fact, this results by nothing more than the observation from Section 1.1, where we discussed the possibility in the logic **K** to represent incoherent information (just by not imposing the principle (D) as an axiom!). Moreover, of course, incoherency represented in this way is rather trivial and impractical in use, since the following validity holds in S5P: $\models P_i\bot \rightarrow P_i\varphi$, for arbitrary formula φ: thus, although $P_i\bot$ is representable (satisfiable) it immediately yields an 'explosion' of beliefs ($P_i\varphi$).

In **S5P**, however, one can do more than just represent incoherency in this trivial manner, viz. by using distinct P_i-operators! The intuition behind the P_i-operators is that they refer to (plausible) beliefs within certain contexts (frames of reference), represented by the sets S_i. For instance, in the example **S5P** was designed for, default reasoning, it might be the case that some default leads us to believe that φ is plausible, while some other (in a different context) might lead us to believe $\neg\varphi$ (see [MH91, 92, 93] for examples of this kind). A common-sense example in a robot world may be that the robot has—in case condition p holds—a rule of thumb that q holds, while, for the situation that r holds, it has a rule of thumb that $\neg q$ holds. Of course, the robot is in some kind of dilemma, if he encounters a situation in which $p \wedge q$ holds. In this situation it has really incoherent information arising from two different contexts. In **S5P** one may represent this as $P_1 q \wedge P_2 \neg q$. Such a representation does not give rise to a belief 'explosion' as above.

Thus the logic of **S5P** is an amalgam of the modal logics **S5** (for K) and **K45** (for P_i), forged by the following connecting axioms:

(K→P)	$K\varphi \rightarrow P_i\varphi$
(KP)	$KP_i\varphi \leftrightarrow P_i\varphi$
(PP)	$\neg P_i\bot \rightarrow (P_iP_j\varphi \leftrightarrow P_j\varphi)$
(PK)	$\neg P_i\bot \rightarrow (P_iK\varphi \leftrightarrow K\varphi)$

Here (K→P) expresses that knowledge implies belief within any context. (KP), (PP) and (PK) are generalized introspection properties with respect to K and the P_i. Note the conditions $\neg P_i\bot$ in (PP) and (PK), which say that P_i-belief is not inconsistent (i.e., semantically $S_i \neq \emptyset$). One can show that this provides a sound and complete axiomatization of validity in S5P-models.

The P_i-modalities do not satisfy (D) nor (D*). Both $P_ip \wedge P_i\neg p$ and $P_i(p \wedge \neg p)$ are satisfiable (by an **S5P**-model in which the frame S_i is $\emptyset$) expressing an incoherent belief within frame i. Note that, as already remarked in Section 1.1, the incoherency that we can express in this way is a limited one: it gives a signal that there is some inconsistency in belief, but one can not really use it to reason with this in a useful manner, since in this setting the usual property of material implication is inherited from classical logic: $(P_ip \wedge P_i\neg p) \rightarrow P_i\varphi$ is a validity, so that one also immediately obtains $P_i\varphi$ for an arbitrary formula φ, once one has derived $P_ip \wedge P_i\neg p$. Note, however, that we might also use different contexts to represent incoherent information, like e.g. $P_ip \wedge P_j\neg p$ (for $i \neq j$), as we saw above in the robot example. The intuition here is a natural one: the incoherency arises from different sources (in different contexts). If one uses this representation of incoherence, one does not suffer from a collapse of information since clearly $(P_ip \wedge P_j\neg p) \rightarrow P_k\bot$ is not valid in **S5P** (for no k).

We mention here some related work by Huang & Van Emde Boas ([Hua91], [HE90]), where modal operators $L_{ij}\varphi$ are used to express that agent i believes φ on the basis of agent j's belief in φ. Although is not really exploited in the work mentioned, one can easily imagine the representation of incoherent belief on the basis of different agents (with different beliefs): $L_{ij}\varphi \wedge L_{ik}\neg\varphi$ is satisfiable. However, Huang addresses the issue with respect to incoherence with respect to a clash of internal ('incorporated') belief (denoted by the L_i operator) and external dependency-based belief denoted by the L_{ij} operator (he calls this 'compartmentalized' belief), and discusses how incorporated belief might / should be *revised* in case these conflict on the basis of of the credibility / authority of the agent j on whom i is dependent. For instance, in the case $L_i\varphi \wedge L_{ij}\neg\varphi$, the belief $L_i\varphi$ will persist if i is an expert on φ and j is not; it will be revised in some way if the roles of i and j are reversed or if both i and j are experts (in the latter case Huang proposes that the belief in φ is *contracted* without replacing it with a belief in $\neg\varphi$).[3] Interestingly, Huang proposes to express the roles of agents (with respect to authority), such as expert or learner, by means of an additional operator $D_{ij}\varphi$, meaning that i is dependent on j with respect to φ, which also links the L_{ij} operator to the L_j operator in a way that should be obvious from the above: $L_{ij}\varphi \leftrightarrow D_{ij}\varphi \wedge L_j\varphi$. E.g., the fact that i is an expert on φ is expressed as $D_{ii}\varphi \wedge \exists j \neq i : D_{ij}\varphi$. (In words: an expert on φ is an agent that is dependent on itself regarding φ and there is no other agent on which it is dependent regarding φ.)

The logic S5P has an obvious similarity to the logic of local reasoning of Fagin & Halpern (Section 2.4). One may view the frames S_i as the clusters in that approach. There is a

[3] Speaking about *belief revision*. Of course, this is an important area in AI (cf. [Gär88]). However, we know of very few modal logic approaches to this, and we shall not treat this topic in this paper. A modal approach based on dynamic logic, where belief revision is viewed as an *action* to be performed by the agent and is embedded in a logic in which one can reason about agents' actions and capabilities can be found in [LHM94]. Something similar holds for the related problem of *updates* in databases. A reference to the use of dynamic logic (again) for reasoning about updates is [SWM95].

difference, though. In the logic of local reasoning one cannot address a particular cluster by a reference by a modal operator. It is only possible to say something about truth in *some* (unspecified) cluster or in *all* clusters by the modalities of weak and strong belief, respectively. On the other hand, in **S5P** one has the expressibility to refer to particular frames by the appropriate P_i-operator, whereas here there is no possibility to quantify over these frames by some modality. Another interesting link is that the logic S5P restricted to one P-modality (one frame of reference) is very close to the integrated logic of knowledge and belief as proposed by Kraus & Lehmann ([KL86]). (The one P-modality corresponds to their belief operator B, the only difference being that they impose the property (D) on B, disallowing incoherent beliefs.)

2.6. Fusion logic (Jaspars)

Fusion logic, proposed by Jaspars [Jas91, Jas93], is very much related to the logic of local reasoning as discussed in Section 2.4. As Jaspars himself says, he was directly inspired by Rescher & Brandom's "logic of inconsistency" [RB80]. In fact, he uses also macro-worlds, but now viewed as fused by the method of "world-disjunction" (cf. Section 2.4). So, a formula is true in a macro-world if it is true in some of the micro-worlds it contains. Jaspars considers models of the form (S, π, R), where S and π are as usual, but R is now a relation between worlds and possible "macro-worlds", represented by a (non-empty) "fused" set of (standard) possible worlds: $R \subseteq S \times (\wp(S) \setminus \{\emptyset\})$.

Now Jaspars interpreted his notion of "confused belief" by the clause: $(M, s) \models B_J\varphi$ iff for all $T \subseteq S$ with R(s, T) it holds that there exists $t \in T$ such that $(M, t) \models \varphi$. In Rescher & Brandom's terms, we consider all macro-worlds T that are considered possible from s and check whether φ holds there "world-disjunctively".

Again, of course, some further restrictions must be made on the accessibility relation in order to get a genuine logic of confused belief, called **CB**. First we need some additional not(at)ions:

The relations $R^{\uparrow}$, $\overline{R}$ and $\underline{R}$ are given by:

$xR^{\uparrow}Y$ iff xRY' for some $Y' \subseteq Y$;
$X\overline{R}Y$ iff for all $x \in X$: $xR^{\uparrow}Y$;
$X\underline{R}Y$ iff for some $x \in X$: $xR^{\uparrow}Y$.

Now in order to let B_J satisfy the axioms (D), (4) and (5), Jaspars imposes the following restrictions on R:

(F-seriality) for all x there is a Y with xRY,

(F-transitivity) for all x, Y, Z: $xR^{\uparrow}Y$ & $Y\bar{R}Z \Rightarrow xR^{\uparrow}Z$, and
(F-euclidicity) for all x, Y, Z: $xR^{\uparrow}Y$ & $xR^{\uparrow}Z \Rightarrow Y\underline{R}Z$,

respectively (see [Jas93]). In **CB** (D) is valid (by the condition of F-seriality on the models), but (D*) is not. Thus $B_Jp \wedge B_J\neg p$ is satisfiable in **CB**, while $B_J(p \wedge \neg p)$ is *not.*

2.7. Implicit knowledge (Levesque)

Levesque ([Lev84]) has also proposed a solution to representing incoherent belief in a nonstandard modal logic approach, but completely different from the approaches discussed so far. In fact, his proposal comes down to a mixture of many-valued logic and modal logic. Possible worlds, or *situations*, as Levesque calls them admit assertions to be *true*, *false*, *none* of these, as well as *both.* Formally, this idea is implemented by defining both a truth support function and a falsehood support function which, in principle, are independent from each other.

Thus, models à la Levesque are of the kind $(S, \mathcal{B}, \pi_T, \pi_F)$ where S is a nonempty set of situations, $\mathcal{B} \subseteq S$ is a set of situations on which we will base the semantics of belief (or *explicit* belief, as Levesque calls it), and π_T and π_F are the truth and falsehood support functions, respectively, which assign truth and falsehood, respectively, to the atoms per situation.

Now we consider both a truth assignment $\models_T$ and a falsehood assignment $\models_F$ of a formula given a model and a situation. The truth assignment $\models_T$ is defined as usual for atoms (depending on π_T), conjunctions and disjunctions; as for the negation, we have that $(M, s) \models_T \neg\varphi$ iff $(M, s) \models_F \varphi$, and with respect to the modal belief operator we have $(M, s) \models_T B_L\varphi$ iff $(M, s') \models_T \varphi$ for all $s' \in \mathcal{B}$. The set $\mathcal{B}$ represents the set of situations that are considered possible. One might also phrase this clause in (the usual) terms of an accessibility relation R by defining R(s, t) iff $t \in \mathcal{B}$, for all $s, t \in S$. The definition of $\models_F$ is dual to that of $\models_T$: for atoms it depends on π_F; $(M, s) \models_F (\varphi \vee \psi)$ iff $(M, s) \models_F \varphi$ and $(M, s) \models_F \psi$; $(M, s) \models_F (\varphi \wedge \psi)$ iff $(M, s) \models_F \varphi$ or $(M, s) \models_F \psi$; and $(M, s) \models_F \neg\varphi$ iff $(M, s) \models_T \varphi$. However, with respect to belief we have: $(M, s) \models_F B_L\varphi$ iff $(M, s) \not\models_T B_L\varphi$, expressing that when we evaluate "meta-assertions" about belief we reason classically: something is believed or it is not, and not both. But, of course, within the scope of a belief operator we may encounter incoherences, to which we shall focus our attention shortly.

We call a situation *classical* if it only admits formulas to be true or false, exclusively: for every atom $p \in \mathbf{P}$, either $\pi_T(s)(p) = t$ or $\pi_F(s)(p) = t$, but not both. Validity of a formula φ (denoted $\models_L \varphi$) is now defined as φ having truth support (i.e., $(M, s) \models_T \varphi$) in all situations of S that are *classical.* A formula is *satisfiable* if there is a structure M and a *classical* situation $s \in S$ with $(M, s) \models_T \varphi$.

In Levesque's logic neither (D) nor (D*) are valid, so that both $B_Lp \wedge B_L\neg p$ and $B_L(p \wedge \neg p)$ are satisfiable. To see this, just take a model with a situation $s \in \mathcal{B}$ in which $\pi_T(s)(p) = \mathfrak{t}$ *and* $\pi_F(s)(p) = \mathfrak{t}$. Neither does the operator B_L satisfy the (K) axiom. On the other hand, the logic does satisfy the introspection axioms (4) and (5) (where the implications involved are just defined as material implication), so that it is a genuine logic of belief.

Finally we remark that Levesque's logic for explicit belief has an interesting connection to a *paraconsistent* logic with an non-material implication, viz. *(first-degree) relevance logic* (cf. [Dun86]). In relevance logic the assertion $\varphi \wedge \neg\varphi \twoheadrightarrow \psi$ is *not* valid (where $\twoheadrightarrow$ stands for relevant implication), and it takes its name from the fact that in a valid implication the premise is relevant in some precise sense for the conclusion (for more about this, consult [Dun86]). More precisely, one can prove that (using $\models_r$ for validity in relevance logic.)

$$\models_L (B_L\varphi \rightarrow B_L\psi) \;\Leftrightarrow\; \models_r \varphi \twoheadrightarrow \psi.$$

2.8. Impossible world semantics (Rantala)

Rantala ([Ran82]) probably proposed the most radical way to enable the representation of incoherent knowledge. In this approach "anything goes": it solves all paradoxes of logical omniscience as well as the incoherence representation problem in one fell swoop. It does so by introducing "very" non-standard possible worlds, called *impossible worlds*, in which truly the impossible may be true.

Formally we consider so-called Rantala-models: models of the kind (S, S^*, π, π^*, R), where S is a non-empty set of possible worlds, S* is a set of impossible worlds, π is a truth assignment function to the atoms on S, π^* is a valuation function of *arbitrary* (!) formulas on S*, and R is an accessibility relation of type $R \subseteq (S \cup S^*) \times (S \cup S^*)$, thus associating with a (possible or impossible) world a set of accessible (possible or impossible) worlds. All clauses for the interpretation of the language in possible worlds of these models are standard (*including* those for the modal belief operator B_R); the truth condition for impossible worlds (in which one may have to evaluate formulas when considering modal formulas), however, is *completely free.* Validity (and satisfiability) of a formula in a model $\langle S, S^*, \pi, R\rangle$ is now defined as the truth of that formula in all (some) *possible* world(s) $s \in S$. And, as usual, validity (satisfiability) of a formula is defined as validity (satisfiability) of that formula in all Rantala models of the above form.

Due to the freedom in the interpretation of formulas in impossible worlds one can avoid all imaginable forms of logical omniscience as well as represent incoherent beliefs: neither the principle (D) nor (D*) is valid. For instance, the formula $B_R(p \wedge \neg p)$ is satisfied in a Rantala-model with an impossible world where $p \wedge \neg p$ is just stipulated to hold.

3. RESOLVING CONFLICTS IN DEFAULT REASONING

3.1. Modal approaches of default reasoning

A form of reasoning in AI where incoherency plays a prominent role is that of reasoning by default. Default reasoning is the form of commonsense reasoning that is used to infer what usually, typically or normally is the case. The in the literature ubiquitous example is that of "birds normally can fly". So, if we are given a bird of which nothing else is known, we infer that it can fly by default. Quite naturally, default reasoning gives rise to incoherency: if we are given two defaults "normally, p implies r" and "normally, q implies $\neg$r", given both p and q, we might use either default to infer r and $\neg$r, respectively. Of course, using both defaults together would result in a direct and hard inconsistency ($r \wedge \neg r$), or at least in incoherent beliefs, when one is more careful and gives the outcomes of default applications the weaker status of beliefs. However, in some cases of conflicting defaults one might have a preference as to which default should be applied, so that the incoherency of default beliefs is avoided. This is, for instance, the case when one default is more *specific* than another one. Suppose that we have the defaults "normally, birds can fly" and "normally, birds that are wing-clipped cannot fly". Given the situation of a wing-clipped bird, in principle one might use both defaults to derive that this bird can fly and cannot fly, respectively. However, commonsense tells us that the second default obtains in the more specific situation, and should overrule the first one, so that one should only be able to infer that this bird cannot fly by default.

Reasoning by default has been the object of study in many papers in the last 15 years, starting with the seminal work by Reiter on "Default Logic" ([Rei80]), often embedded in the more general context of *defeasible* or *nonmonotonic* reasoning ([Rei87]). As we are interested in modal approaches we mention especially here the related logics of Moore's auto-epistemic logic (**AEL**, [Moo84, 85]) and, especially, Lin & Shoham ([LS90]) and Schwind & Siegel ([SS92]). Here we shall focus in the first instance on our own approach based on the logic **S5P** of Section 2.5, which we have dubbed Epistemic Default Logic (**EDL**) in [MH93, MH95].

In this section we shall briefly show how one may express default reasoning in **EDL**, and discuss how conflicts might be dealt with. We also pay special attention to the specificity problem that we discussed above.

In the language of **S5P** we may express defaults “normally, if φ then ψ” by $K\varphi \wedge M\psi \rightarrow P\psi$. (This kind of defaults is what is usually called *normal.*, cf. [Rei80]). Here φ and ψ are supposed to be objective formulas. The literal reading of such a formula is “if φ is known to be true and ψ is (considered) possible, then ψ is preferred”. Multiple defaults are represented by sets of formulas $K\varphi_i \wedge M\psi_i \rightarrow P_i\psi_i$, where the φ_i and ψ_i are again objective formulas.

A default theory Θ is a pair (W, Δ), where W is a finite set of objective formulas describing (necessary) facts about the world, and Δ is a finite set of defaults $\{K\varphi_i \wedge M\psi_i \rightarrow P_i\psi_i \mid i = 1, ..., n\}$, where some of the P_i-operators may be the same. The sets W and Δ are to be considered as sets of *axioms* representing (background) *knowledge*, and we may apply necessitation to the formulas in them.

Given a default theory $\Theta = (W, \Delta)$, we define the nonmonotonic inference relation $\vdash_\Delta$ as follows. In the sequel we let, for a finite set Φ of **EDL**-formulas, Φ^* stand for the conjunction of the formulas in Φ. Note that W is a finite set, and moreover that W only consists of objective formulas. Furthermore, let φ be an objective formula such that φ and W^* are mutually consistent. Then $\Sigma^{\varphi\wedge W^*}$ is well-defined (cf. Section 2.1). Then we define the *default entailment relation* $\vdash_\Theta$ w.r.t. default theory Θ as follows:

3.1. DEFINITION. $\varphi \vdash_\Theta \psi \Leftrightarrow_{\text{def}} \psi \in Th_{\mathbf{EDL}}(\Sigma^{\varphi\wedge W^*} \cup \Delta)$.

Instead of $T \vdash_\Theta \psi$, we simply write $\vdash_\Theta \psi$. Furthermore, if Γ is a finite set of objective formulas, and ψ an **EDL**-formula, then we define $\Gamma \vdash_\Theta \psi$ as $\Gamma^* \vdash_\Theta \psi$. Below we will show a number of examples of default entailments.

We can also give a semantic characterization of the default entailment $\vdash_\Theta$. Since this is the easiest for the case that the set **P** of primitive propositions is finite, we consider only this case here. (For a semantical characterization in the general case we refer to [MH92b].)

3.2. THEOREM. *Consider a default theory $\Theta = (W, \Delta)$, where Δ is finite, and let φ be an objective formula such that φ and W^* are mutually consistent. Moreover, let* **P** *be finite. Then we have that:*

$$\varphi \vdash_\Theta \psi \quad \Leftrightarrow \quad \mathbb{M} \models K\Delta^* \rightarrow \psi \textit{ for all } \mathbf{S5P}\textit{-extensions } \mathbb{M} \textit{ of } \mathbb{M}_{\varphi\wedge W^*}.$$

This theorem essentially says that the default consequences of knowing only φ are the consequences of knowing the defaults concerned within the context of knowing only φ together with the background information W. Speaking somewhat more technically, to determine whether ψ is a default consequence of φ, we consider the **S5**-model representing that precisely φ and W is known, and then check whether ψ follows from $K\Delta^*$ in all **S5P**-models extending this **S5**-model.

We illustrate reasoning in **EDL** by a few examples, starting with the inevitable "Tweety example".

3.3. EXAMPLE (Tweety). Consider the following default theory $\Theta = (W, \Delta)$ with $W = \{p \to \neg f\}$ and $\Delta = \{Kb \wedge Mf \to Pf\}$, representing that penguins do not fly, and that by default birds fly. Now consider the following inferences (To stress the application of Δ we denote such a step by $\vdash_\Delta$.):

(i). $b \mid\!\sim Kb \wedge \neg K \neg f \vdash_{\mathbf{EDL}} Kb \wedge Mf \vdash_\Delta Pf$, i.e., $b \mid\!\sim_\Theta Pf$,

meaning that from the mere fact that Tweety is a bird, we conclude that Tweety is assumed to fly; which must be contrasted to the inference:

(ii). $b \wedge p \mid\!\sim Kp \vdash_{\mathbf{EDL}} K\neg f \vdash_{\mathbf{EDL}} \neg Mf \nvdash_\Delta Pf$, i.e, **not** $b \wedge p \mid\!\sim_\Theta Pf$,

meaning that in case Tweety is a penguin, we cannot infer that Tweety is assumed to fly, but instead we can derive to know for certain that Tweety does not fly.

A somewhat more interesting example below shows that our framework allows for a kind of partial normality in the sense that in a particular case some defaults are applicable while some other ones are not. So it possible to express a kind of "graded normality" in our framework: the object at hand is (assumed) normal in certain respects but abnormal in other ones.

3.4. EXAMPLE *(Graded Normality).* Consider the following sentences:

(i) Lions are normally dangerous.
(ii) Lions are normally brown.
(iii) Leo is a cub lion.
(iv) Cubs are not dangerous.

Of course we expect to be able to infer that Leo is not dangerous. However, we do still expect to infer that Leo is brown. This phenomenon is called "graded" normality in the literature: although Leo is explicitly stated to be not normal with respect to being dangerous, we still expect him to be normal with respect to being brown, since nothing is said about that to the contrary. In our approach this works out fine, as follows. Letting 'l' denote "being a lion", 'd' "being dangerous", 'b' "being brown", and 'c' "being a cub", we consider the default theory $\Theta = (\{T\}, \Delta)$, with Δ given by the set of defaults:

(i') $Kl \wedge Md \rightarrow Pd$

(ii') $Kl \wedge Mb \rightarrow Pb$

(iii') $c \wedge l$

(iv') $c \rightarrow \neg d$

Now we can infer: $c \vdash_{\mathbf{EDL}} Kc \vdash_{\mathbf{EDL}} K\neg d \vdash_{\mathbf{EDL}} \neg Md$. Thus $\nvdash_{\Delta} Pd$. But, on the other hand, $l \vdash_{\Theta} Kl \wedge \neg K\neg b \vdash_{\mathbf{EDL}} Kl \wedge Mb \vdash_{\Delta} Pb$. Thus it is a preferred belief that Leo is brown, as desired.

The following example shows the use of multiple distinct P-operators associated with multiple distinct frames of reference.

3.5. EXAMPLE *(Multiple 'extensions')*. Consider the following defaults:

"Normally, if p then q",
"Normally, if p then r"

These default rules, which have two extensions, can be formulated in our language as $\Delta =$

$$\{Kp \wedge Mq \rightarrow P_1q\,,\\ Kp \wedge Mr \rightarrow P_2r\}.$$

Considering the default theory $\Theta = (\{p\}, \Delta)$, we get $T \vdash_{\Theta} Mq \wedge Mr \vdash_{\Delta} P_1q \wedge P_2r$. Hence we obtain two preferred frameworks. Of course, in this case, the default theory Θ' which is as Θ but where the two defaults are represented using the same P-modality, say P_1, would also have a useful conclusion, viz. $P_1q \wedge P_1r$ which is **EDL**-equivalent with $P_1(q \wedge r)$, expressing that there is one preferred frame in which both q and r hold. However, note that in our approach the default theories Θ and Θ' are different: it is up to the knowledge engineer which one he wants to consider.

The use of different P-modalities, referring to distinct frames, becomes especially pregnant in the case of the infamous Nixon Diamond, which is of particular importance viewed from the perspective of handling incoherent (default) beliefs.

3.6. EXAMPLE (Nixon Diamond, ([RC81]). Let r stand for being a republican, p for being a pacifist, and q for being a quaker. Now we consider the default theory $\Theta = (\{\top\}, \Delta)$ with Δ given by:

(1) $Kr \wedge M\neg p \rightarrow P_1\neg p$ (republicans are non-pacifists by default)
(2) $Kq \wedge Mp \rightarrow P_2p$ (quakers are pacifists by default)

We are now given that Nixon is both a republican and a quaker, and we want to draw some conclusion about his being a pacifist or not. We now infer $r \wedge q \vdash_\Theta Mp \wedge M\neg p \vdash_\Delta P_1\neg p \wedge P_2p$. That is to say, we have two subframes S_1 and S_2 of S: in the one $\neg p$ holds, in the other p. This is intuitively correct since there is in this particular case no preference of the one over the other whatsoever.

Note the need for different modalities P_1 and P_2 (or frames S_1 and S_2) in this case: the use of only one such modality, say P_1, results in an inconsistency within the frame S_1 associated with that modality: $P_1\neg p \wedge P_1p$, which is equivalent with $P_1(\neg p \wedge p)$, i.e., $P_1\bot$. Although this formula is not inconsistent in itself, the resulting (empty) frame is probably not what we intend to have as a preferred frame. The use of multiple P-modalities enables us to represent the outcome in a more sensible and useful way, viz. two consistent frames each representing a reasonable conclusion. In a very analogous way the use of multiple frames enables us to represent the Lottery Paradox in a consistent and intuitively correct way (see [MH92], [MH95], and later in Section 4).

Although the use of various distinct P-operators provides us with (a) means to represent conflicting information in a consistent and sensible manner by keeping the conflicts apart in separate frames of reference, it is obvious that the moment one is willing (forced) to act upon ones' default beliefs, choices will have to be made in the case of these inherently conflicting pieces of information. In meta-level reasoning jargon the point where these choices will have to be made is "downward reflection". This will be discussed in the next section, but first a few words about autoepistemic logic (**AEL**, [Moo84, 85]) and related work by Lin & Shoham ([LS90]) and Schwind & Siegel ([SS92]).

First of all, Moore's **AEL** is very close to **EDL** with respect to the representation of defaults: instead of $K\varphi \wedge M\psi \rightarrow P\psi$, essentially the same form is used (although instead of 'K' the operator is usually denoted 'L'), but the crucial difference is that in the conclusion no P-modality is used, and the default thus essentially is represented by means of one modal operator: $L\varphi \wedge M\psi \rightarrow \psi$. To get conclusions from a autoepistemic theory, so-called *AE-extensions* are used which are (again) stable sets that are *grounded* (i.e., in some technical sense rooted) in the set of objective formulas that are known to be true (cf. [Moo84, 85], [Kon94], [MH95]).

Lin & Shoham also use two modalities, like in **EDL**, but in a different way: besides the K(nowledge)-operator they also employ an A(ssumption)-operator, which both have a standard modal (**KD45**) interpretation, and are *a priori* unrelated. (Normal) defaults are now represented as $K\varphi \wedge \neg A \neg \psi \rightarrow K\psi$. (Note the different use / place of the modalities as compared to **EDL**.) Consequences of default theories are obtained by considering (only) *preferred* models in which given the set of true A-facts (formulas of the form $A\varphi$, where φ is an objective formulas) the set of K-facts (formulas of the form $K\varphi$, where φ is an objective formula) is minimal and additionally the condition holds that the set of K-facts is equal to the set of A-facts (thus in a sense the known facts must be rooted in the assumed facts). (cf. [LS90], [Kon94], [Poo94]). This is very much related to the topic of *modal nonmonotonic logics*, which we shall not pursue in this paper. (See [Kon94], [MT93]).

Finally, Schwind & Siegel propose yet another representation of a default theory using two modal operators. (They call their approach Hypothesis Theory.) They employ a modal operator L (with intended meaning something like "knowledge") satisfying the axiom (T) (cf. Section 2.1) together with an operator [H] which is a normal (system K-like necessity) operator, of which the dual $H\varphi = \neg[H]\neg\varphi$ is intended to mean that φ is a hypothesis. The only interaction between L and [H] is given by the axiom $L\varphi \rightarrow [H]\varphi$, or equivalently, $H\varphi \rightarrow \neg L \neg \varphi$: if something is an hypothesis it cannot be known to be false. A hypothesis theory consists of a set F of formulas (facts) and a set HY of hypotheses. A default theory (W, D) is represented as a hypothesis theory (F, HY) with $F = LW \cup LD \cup \{L\varphi \rightarrow [H]\varphi\}$, where $LW = \{Lw \mid w \in W\}$ and $LD = \{L\varphi \wedge H\psi \rightarrow L\psi \mid$ "normally, if φ then ψ" $\in D\}$, and HY the set of hypotheses occurring in D. Thus (normal) defaults of the form "normally, if φ then ψ" are represented in Hypothesis Theory as formulas of the form $L\varphi \wedge H\psi \rightarrow L\psi$. In this respect Hypothesis Theory resembles the approach of Lin & Shoham, although Schwind and Siegel do impose an a priori relation between the accessibility relation R_L associated with the operator L and the one ($R_{[H]}$) associated with [H]: $R_{[H]} \subseteq R_L$. Extensions of a default theory are now defined as maximal sets $F \cup HY'$ with $HY' \subseteq HY$ such that $F \cup HY'$ is consistent. It is shown in [SS92] that this completely characterizes Reiter's Default Logic in this modal setting. What is interesting about this approach is that it assumes very weak requirements on the epistemic modalities L and [H] (e.g., introspection properties are assumed for neither L nor [H]!), while still the full power of default logic is obtained.

3.2. Resolving conflicts in downward reflection of meta-level information

In Section 2.5 we discussed the logic **S5P** as a logic to reason about knowledge as well as plausible working beliefs that are, for instance, caused by (applying) defaults as exemplified in the previous section. As we noticed already in these sections, we might end up in situations where we have an incoherent set of such beliefs, represented by formulas like $P_i\varphi \wedge P_j\neg\varphi$ (for possibly different i and j). This is all fine when we just want to represent incoherent beliefs. But what do we do with them if we really want to use them and act upon them?

We may view this as more generally as a problem of *reflection* in *meta-level reasoning*. In *meta-level reasoning architectures* (e.g. [MN88]) one may "reflect" information from one level to another: for instance, one may reflect meta-level information of the provability of some assertion to the information that that assertion is true on the object level. Thus we may see the above problem with defaults with conflicting outcomes in this way: how can we reflect these outcomes from the meta-level to the object level (this is usually called "*downward reflection*") such that consistency is maintained (cf. e.g. [TT91, TT92]). Downward reflection maps meta-level information such as default beliefs into the object level as what we will call "quasi-facts". This is a risky business since this turns uncertain beliefs into something where its modality of being a mere defeasible belief is forgotten/ deleted/ignored.

Of course, when on the meta-level it is derived that both φ and $\neg\varphi$ are plausible beliefs, reflecting both these beliefs are reflected downwards to the object level, we get a genuine inconsistency on this level, containing both φ and $\neg\varphi$.

To be more precise: let us denote the downward reflection operator by β. Thus β is a function from a set Φ of **EDL**-formulas to a set Ψ of non-modal formulas, intended to capture the reflection of meta-knowledge as represented by Φ to a set of quasi-facts on the object level represented by the outcome Ψ.

Naturally, it is sometimes quite unproblematic to define β. For instance, in the case that $\Phi = \{P_1 p \wedge P_2\neg q\}$, where p and q are different atoms, it is clear that β should just delete the P-modalities, just letting us forget what the exact epistemic status is of the assertions p and $\neg q$. So in this case we would define $\beta(\Phi) = \{p \wedge \neg q\}$. This can be done so easily in this case, since the resulting set $\Psi = \{p \wedge \neg q\}$ is consistent.

However, just deleting the P-modalities does not always work: e.g. in the case that $\Phi = \{P_1 p \wedge P_2\neg p\}$, we would obtain the inconsistent set $\Psi = \{p \wedge \neg p\}$. This is obviously undesirable. On the other hand, if we would consider $\Phi = \{P_1 p \wedge P_1\neg p\}$, or just plainly $\Phi = \{P_1(p \wedge \neg p)\}$, it is less clear that downward reflection should not lead to inconsistency: $\beta(\Phi) = \{p \wedge \neg p\}$ seems to be the only sensible choice.

Inconsistency may, by the way, also arise using the above naive strategy in case the set Φ contains knowledge, e.g., $\Phi = \{Kp \wedge P_1\neg p\}$. Just deleting modalities would yield to $\beta(\Phi) = \{p \wedge \neg p\}$ again. However, intuitively the P_1-belief $\neg p$ should be ignored in the context of the certain knowledge that p.

Thus we have to find mechanisms to cope with this situation without running in real inconsistencies on the object level. There are several ways to do this, which we explored in [MH93b]:

1. *Put explicit priority orderings $\prec$ on frames,* $S_1 \prec S_2$ meaning that S_2 is 'better' in the sense of more preferred or relevant than the frame S_1. The ordering $\prec$ is now taken into account to resolve possible conflicts: it should be the case that when P-formulas (that is, formulas involving P-modalities) that become inconsistent after the removal of the P-operators, only the P-formulas pertaining to the more preferred frame are downward reflected. So for example, if $S_1 \prec S_2$, we get that $P_1p \wedge P_2\neg p$ on the meta-level results in $\neg p$ on the object level. Of course, if this is to work in all cases we need a total ordering on frames, which is not always a very reasonable property to have, since the ordering will reflect priority principles such as specificity or legal principles like *lex superior* (i.e., "according to the highest authority"). Since these principles may leave preferences undecided, conflicts cannot always be solved.

2. *Define implicit priorities into the semantics of defaults* (without changing their *syntactic* representation). The option of trying to resolve conflicts as mentioned under 1 above is one in which it is presumed that one (the user of the system) has *a priori* intuitions about the priority of the contexts of concern. This may be realistic a presumption in some applications, but not always. For instance, let us consider the issue of *specificity* in default reasoning. Here we have rules of thumb (resulting in working beliefs) but some rules apply in a more specific case than other ones. E.g. one rule is applicable in case that p holds to infer the plausible belief that q holds. But, on the other hand we may have a rule that says that in situation $p \wedge r$ it is plausible that $\neg q$ holds. Now we can derive both the working beliefs (say) P_1q and $P_2\neg q$ by the respective rules, but many would agree that, given $p \wedge r$, $(P_2)\neg q$ should take precedence when we reflect downwards. The typical example is: if we have a bird, it is plausible that it can fly, while if we have a bird that is an ostrich, it is plausible (if not certain) that it does not.

Many authors consider the case of *specificity* as a special case which in their opinion should be treated by the logic itself without intervention from the user (cf. e.g. [Vel91]). This is rather a controversial issue, since some other authors claim that in this way specificity is privileged over other principles of preference like the above-mentioned *lex superior*, which makes it *de facto* even harder to incorporate various different such principles in one system (cf. e.g. [Pra93]). Also capturing specificity in isolation is by no means trivial, since one may encounter difficult situations with cascaded specificities that contradict each other, where intuitions begin to fade (cf. [Tou86], [MH95]), as in a case like: *Adults tend to be employed; University students tend to be unemployed; University students tend to be adults; Adults under 22 tend to be university students; Tom is an adult under 22; Is he employed?*

3. *Use modalities to directly indicate the relative strength / weakness of plausible working beliefs.* In this approach the conflict between a belief $P\varphi$ and a belief $P'\neg\varphi$ can be solved by considering the strength of the modality P *versus* that of P'. One way of doing this is to represent clauses concerning working beliefs in such a way that their relative priority is programmed into the representation where the mutual strength of the concerned modalities are used to solve conflicts (see the next section, where we will do this explicitly with respect to the specificity problem in default reasoning). But one might also, for instance, use *numerical modalities* such as *graded P-modalities* (cf. [HM92]), to indicate directly the degree of trust in the working beliefs. These graded P-operators can be either absolute or relative. In the absolute case we may use operators $P_{i,n}$ with intended meaning: $P_{i,n}\varphi$ iff φ holds in frame S_i modulo (precisely) n exceptions (i.e. exceptional worlds, where $\neg\varphi$ holds). In this case the downward reflection should take this degree of trustworthiness into account. For example, $P_{1,5}p \wedge P_{2,10}\neg p$ should result in p on the object level. In the relative case we may use operators $P_{i,\lambda}$ with $1 \leq \lambda \leq 1$, with intended meaning: $P_{i,\lambda}\varphi$ iff φ holds in frame S_i for the fraction (of possible worlds within S_i) λ. So, e.g., $P_{1,0.2}p \wedge P_{2,0.5}\neg p$ on the meta-level should result in $\neg p$ on the object level. This is of course very much related to other quantitative or numerical approaches, such as graded modal logic and probability-based modal logic, to which we shall turn in Section 4.

Downward reflection, as described above, is not really part of our logic. It is more a kind of procedure or algorithm that can be applied on the set of working beliefs that are obtained by means of our logic **EDL**. By using a dynamic logic one may include downward reflection into the logic. By the use of dynamic logic it becomes even possible to be explicit in the logic and indicate which of the possibilities of downward reflection that we gave above, is adopted. So in this way we can view the whole process of calculating defaults as a procedure or action $\gamma = \alpha\ ;\ \vdash_{\mathbf{EDL}}\ ;\ \beta$, where ";" stands for sequential composition, β is the downward reflection procedure of our choice, $\vdash_{\mathbf{EDL}}$ is the procedure of calculating the **EDL**-theory, and α is the so called "upward reflection" procedure yielding the set $\Sigma^{\varphi\wedge W^*}$ when given the premise φ together with the background information W. (cf. Def. 3.1. This is in fact Halpern & Moses' nonmonotonic operator $\mid\!\sim$ from [HM84]). Thus γ is the result of first performing upward reflection, then deducing the **EDL**-consequences of this, and then reflect the results downward by means of β. Now, we may employ dynamic logic to express the result: for instance, the formula $\varphi \rightarrow [\gamma]\psi$ expresses that if φ holds before performing γ then ψ will hold after performance of γ. Since $\gamma = \alpha \,;\, \vdash_{\mathbf{EDL}} \,;\, \beta$, this is equivalent with $\varphi \rightarrow [\alpha][\vdash_{\mathbf{EDL}}][\beta]\psi$. So e.g. in the Tweety example we may state that $\mathrm{b} \rightarrow [\alpha]\neg\mathrm{K}\neg\mathrm{f}$, and so $\mathrm{b} \rightarrow [\alpha][\vdash_{\mathbf{EDL}}]\mathrm{Mf}$, and hence $\mathrm{b} \rightarrow [\alpha][\vdash_{\mathbf{EDL}}]\mathrm{Pf}$. Thus, $\mathrm{b} \rightarrow [\alpha][\vdash_{\mathbf{EDL}}][\beta]\mathrm{f}$, for any simple downward reflection β that deletes P-modalities. On the other hand, we have that $\mathrm{b} \wedge \mathrm{p} \rightarrow [\alpha]\mathrm{Kp}$, and thus that $\mathrm{b} \wedge \mathrm{p} \rightarrow [\alpha][\vdash_{\mathbf{EDL}}]\mathrm{K}\neg\mathrm{f}$, and $\mathrm{b} \wedge \mathrm{p} \rightarrow [\alpha][\vdash_{\mathbf{EDL}}]\neg\mathrm{f}$. So, we only have to require that the downward reflection β does nothing in this case, as is to be expected from any well-behaved downward reflection operator, and obtain $\mathrm{b} \wedge \mathrm{p} \rightarrow [\alpha][\vdash_{\mathbf{EDL}}][\beta]\neg\mathrm{f}$. Research along similar lines is reported by Sierra, Godo & Lopez de Mantaras [SGL95].

One may also abstract away from the precise procedure and use temporal logic to describe downward reflection as a process over time. Then we may say simply something like $\mathrm{b} \rightarrow \mathrm{Xf}$ and $\mathrm{b} \wedge \mathrm{p} \rightarrow \mathrm{X}\neg\mathrm{f}$, where X is a next-time operator. This possibility of dealing with downward reflection is explored in [HMT94a,b], where we have provided a *temporal* semantics to this kind of reasoning. In this approach, downward reflection is modelled by taking a time step: (some) plausible beliefs are made true at the next instance of time. In the papers mentioned branching time temporal logic is employed to give a systematic treatment of all possibilities of reflecting (combinations of) plausible beliefs downwards to the object-level.

3.3. Programming priorities into default representations

In the previous section we discussed inconsistency handling as a problem of downward reflection in meta-level reasoning: how to reflect incoherent information down to the object-level such that consistency is maintained? In this section we will investigate the approach mentioned already as a possibility in the last section, viz. representing a default theory in such a way that priorities are enforced automatically when considering its consequences. In fact, this approach amounts to resolving possible conflicts already at a stage before the downward reflection of the working beliefs to quasi-facts. We hinted already to this approach in [MH92a] when we treated the example of multiple defaults with specificity. Here we elaborate on how a priority can be "programmed" into the representations of defaults in order to indicate which default takes precedence. So rather than to solve the conflict afterwards, this method amounts to resolving it before it really becomes a conflict of incoherent default beliefs. To illustrate the method we look at the example of *specificity* as a priority principle, but the method extends to other principles. This approach is inspired by work done in the settings of other nonmonotonic formalisms such as (prioritized) circumscription and default logic ([McC80], [MT93], [Lif94], [Poo94]) and the idea of *stratification* from logic programming ([ABW88]), but it is shown here within the context of our modal approach to default reasoning, viz. **EDL**. In fact, in this section we shall develop an extension of Halpern & Moses' theory of honest formulas (as well as the notion of stable sets and the consequence relation based on these) in order to cater for a 'stratified' default theory the consequences of which are constructed "stratum by stratum".

First we recall from [MH92] the following example, where we consider a situation in which we have multiple defaults where some defaults apply to a more general case and some other ones apply to a more specific case.

3.7. EXAMPLE (Multiple defaults with specificity). Consider the defaults "ravens are generally black" and "albino ravens are generally not black", represented by:

(1) $Kr \wedge Mb \rightarrow P_1b$ (normally, ravens are black)
(2) $K(r \wedge a) \wedge M\neg b \rightarrow P_2\neg b$ (normally, albino ravens are not black)

Note that if we know that we have an albino raven ($K(r \wedge a)$), and that both b and $\neg b$ are possible ($Mb \wedge M\neg b$), we can infer both P_1b and $P_2\neg b$.

First we note that the above outcome creates exactly the problem with downward reflection that we discussed before: when we reflect downward we have to resolve an inconsistency that stems from the deletion of the P-operators. This conflict may be resolved along the lines we have discussed in Section 3.2, e.g. by using an ordering on the frames.

However, we moreover have the intuition that the above outcome is not the best one may get, since the result does not seem to use directly the information that (2) is more specific than (1). (Of course, by imposing an ordering on the frames such that the frame S_2 has a higher priority than S_1 does incorporate this information, but in a rather indirect way.) We may wonder whether there is not a direct way to do justice to the fact that (2) is more specific than (1), and thus should get priority. In fact we suggested such a solution already in [MH92], where we discussed a more refined representation of defaults in which it is possible to 'program priority into the representation' in a logic programming-like fashion.

In this approach we represent defaults by means of formulas not only of the form $K\varphi \wedge M\psi \rightarrow P_i\psi$, but also of the form $K\varphi \wedge \neg P_j\neg\psi \rightarrow P_i\psi$, thus allowing P-modalities in the antecedents as well! (Here φ and ψ are objective formulas again.) This is really an extension of the (power of the) formalism in the sense that instead of allowing only formulas of the form $M\psi$ in the antecedent to check "overall consistency", which is interpreted as the formula ψ being satisfiable in the *whole* set S of *a priori* possible worlds, it is now also allowed to check whether a formula ψ is "consistent with the frame S_i", that is satisfiable within the *subset* S_i of S. We will call default theories where all defaults have either the "old" form or this more general form *generalized default theories.*

Now the representation of our example becomes:

(1*) $Kr \wedge \neg P_2\neg b \rightarrow P_1 b.$
(2) $K(r \wedge a) \wedge M\neg b \rightarrow P_2\neg b$

Now we may infer from $K(r \wedge a)$, given $M\neg b$, that $P_2\neg b$, which *blocks* the application of (1*) and the possible conflict that results when we reflect downward is resolved on beforehand.

So this is a way to specify or 'program' priorities in multiple defaults explicitly. When we first proposed this in [MH92a], we did not really elaborate on how calculations should proceed exactly in this more general setting, since obviously we now need to infer formulas of the form $\neg P\psi$ to enable us to apply defaults of the form of (1*)! Although we then thought of a kind of Negation-as-Failure-like approach, we now can be much more specific about this, using the ideas of Halpern & Moses of deriving ignorance from knowledge in an iterated kind of way. Moreover, as is clear from the example above, there also seems to be an issue of order of applying defaults: in the example we have to try and apply (2) first, after which (1*) can be possibly applied. Of course, this has to do with the fact that to be able to apply (1*) in the intended manner, we need to know whether something about the P_2-modality can be derived in the rest of default theory. This aspect, too, can be made much more explicit in our present approach, as we shall see later on.

The general idea is to focus on frames S_i, and consider them as S5-models in themselves. (This can be done if the frames are non-trivial in the sense of non-empty.) Now the whole apparatus of stable sets, honest formulas and the entailment relation $\vdash$ becomes available again.

To be able to do this properly we need our default theories to be *stratified*:

3.8. DEFINITION. A generalized default theory $\Theta = (W, \Delta)$ is *stratified* if there is a (finite) partition $\{\Delta_i\}_{1 \leq i \leq n}$ of the set of (generalized) defaults Δ (each Δ_i is called a *stratum*) which is partially ordered by a strict partial order $\lhd$, where each stratum Δ_i consists of all defaults with conclusions of the form $P_i\chi$ and every generalized default from a stratum Δ_i of the form $K\varphi_i \wedge \neg P_j\neg\psi_i \rightarrow P_i\psi_i$ is such that $\Delta_j \lhd \Delta_i$.

Intuitively, stratification of a default theory means that every default in a stratum Δ_i, which refers to a frame S_i, is *dependent* only on outcomes of defaults in lower strata (which refer to frames S_j such that $\Delta_j \lhd \Delta_i$). (Here the order of defaults comes into the picture!) Note that the definition allows "old" defaults of the form $K\varphi_i \wedge M\psi_i \rightarrow P_i\psi_i$ to occur in any stratum Δ_i. The reason for this, of course, is that the old notion of a default in its antecedent refers to the set S as a whole (by means of the operator M) and not to subsets S_j at all. In fact, one may view the (certain) facts in W as a 0-th stratum Δ_0 for which it holds that $\Delta_0 \lhd \Delta_i$ for all $1 \leq i \leq n$.

3.9. EXAMPLE. The default theory $\Theta = (W, \Delta)$ given by $W = \{p \rightarrow b, b \rightarrow a\}$ and $\Delta =$

$$\{ \quad Kp \wedge M\neg f \rightarrow P_1\neg f,$$
$$Kb \wedge \neg P_1\neg f \rightarrow P_2 f,$$
$$Kb \wedge \neg P_1\neg w \rightarrow P_2 w,$$
$$Ka \wedge \neg P_2 f \rightarrow P_3\neg f \quad \}$$

is stratified. Take the partition $\Delta = \Delta_1 \cup \Delta_2 \cup \Delta_3$, where

$$\Delta_1 = \{Kp \wedge M\neg f \rightarrow P_1\neg f\},$$
$$\Delta_2 = \{Kb \wedge \neg P_1\neg f \rightarrow P_2 f,$$
$$Kb \wedge \neg P_1\neg w \rightarrow P_2 w\},$$
$$\Delta_3 = \{Ka \wedge \neg P_2 f \rightarrow P_3\neg f\},$$

and the ordering on default strata simply given by $\Delta_i \lhd \Delta_j$ iff $i < j$.

In order to cater for stratified default theories we need to express epistemic states relatively to a frame S_i. We first assume the set **P** of primitive propositions to be finite. In order to speak about epistemic states relative to a subframe S_i, we consider the sublanguage $\mathcal{L}_i$, consisting of the set of all **EDL**-formulas closed under containing the set **P**, the classical propositional connectives and the operator P_i. So in $\mathcal{L}_i$ one can only express properties of belief with respect to frame S_i and no other frames S_j, nor the whole set S. Note that $\mathcal{L}_i$ contains all objective formulas. We now define the notion of an i-stable set:

3.10. DEFINITION. A set $\Sigma \subseteq \mathcal{L}_i$ is *i-stable* if it is either the inconsistent set $\mathcal{L}_i$ or it is propositionally consistent and satisfies the following:

(St$_i$1) all instances of propositional tautologies are elements of Σ;
(St$_i$2) if $\varphi \in \Sigma$ and $\varphi \rightarrow \psi \in \Sigma$ then $\psi \in \Sigma$;
(St$_i$3) $\varphi \in \Sigma \Leftrightarrow P_i\varphi \in \Sigma$

$(\mathrm{St}_i4) \quad \varphi \notin \Sigma \Leftrightarrow \neg P_i\varphi \in \Sigma$

i-Stable sets, of course, enjoy the same nice properties of stable sets; in fact when consistent they are also theories of (simple) S5-models, viz. sets S_i. The only difference now is that they may be inconsistent (when the set S_i is empty). (This, by the way, is also allowed for Moore's notion of a stable set [Moo84, Moo85].) Moreover, the whole apparatus of Halpern & Moses' entailment relation including the notion of honesty can be defined analogously with respect to i-stability. We need only extend the definition for the inconsistent case, which is trivial.

3.11. DEFINITION. A formula φ is *i-honest* if there is an i-stable set $\Sigma_i{}^\varphi$ that contains φ and such that for all i-stable sets Σ containing φ it holds that $\mathrm{Prop}(\Sigma_i{}^\varphi) \subseteq \mathrm{Prop}(\Sigma)$.

Note that an inconsistent formula φ now is i-honest, since in that case the set $\Sigma_i{}^\varphi = \Sigma_i{}^\perp = \mathcal{L}_i$ is i-stable and satisfies the requirement in the definition. So in contrast with the case of honesty we do not exclude inconsistent formulas regarding i-honesty, which reflects the difference between S, which is always non-empty and the S_i, which may be empty.

To profit maximally from Halpern & Moses' theory we indicate how truth in a subframe of an S5P-model corresponds to truth in an S5-model, thus reducing the problem of characterizing epistemic states associated with a subframe to the old problem of the characterization of an epistemic state as usual. In fact, the connection is quite obvious, but has to be established formally to be able to employ the "old" theory.

Let $\mathbb{M} = \langle S, \pi, R, S_1, ..., S_n \rangle$ be an **S5P**-model. Define $\mathbb{M}_i$ to be the (simple) S5-submodel of $\mathbb{M}$ with $\mathbb{M}_i = \langle S_i, \pi, R \rangle$. Then it holds that:

3.12. PROPOSITION. *For formulas* $\varphi \in \mathcal{L}_i$ *we have that* $\mathbb{M} \models P_i\varphi \Leftrightarrow \mathbb{M}_i \models \varphi$
PROOF: Directly from the definitions.

$\mathbb{M}_i$ is either an empty model or a (simple) S5-model. In the latter case we can use the whole machinery of Halpern & Moses to derive (non)beliefs with respect to frame i. For instance, we know immediately that the theory $P_i(\mathbb{M}_i) = \{\varphi \in \mathcal{L}_i \mid \mathbb{M}_i \models \varphi\}$ ($= \{\varphi \in \mathcal{L}_i \mid \mathbb{M}_i \models P_i\varphi\}$) of the (simple) S5-model $\mathbb{M}_i$ is an i-stable set. (Note that the principle epistemic modality in frame i is P_i instead of K.) Moreover, by our definition of i-stability, also in the case that $\mathbb{M}_i$ is empty, $P_i(\mathbb{M}_i) = \{\varphi \in \mathcal{L}_i \mid \mathbb{M}_i \models \varphi\} = \mathcal{L}_i$ is i-stable.

So now, analogously to the case for knowledge, we can also employ Halpern & Moses' approach to default beliefs and define an entailment relation $\mathrel{|\!\sim}_i$ associated with the default beliefs with respect to frame S_i, as follows.

$$\varphi \mathrel{|\!\sim}_i \psi \quad \Leftrightarrow_{def} \quad \psi \in \Sigma_i^{\varphi} \text{ for i-honest } \varphi \in \mathcal{L}_i .$$

Note the special case if φ is inconsistent: $\varphi \mathrel{|\!\sim}_i \psi$ for every ψ, since then $\Sigma_i^{\varphi} = \mathcal{L}_{\mathbf{EDL}}$. As before for $\mathrel{|\!\sim}$ we will employ the entailment $\mathrel{|\!\sim}_i$ for *objective* premises φ. These are certainly i-honest.

3.13. EXAMPLES. Let p and q be two distinct primitive propositions. Then e.g.:
$p \mathrel{|\!\sim}_i P_i p$; $p \mathrel{|\!\sim}_i \neg P_i q$; $p \wedge (p \rightarrow q) \mathrel{|\!\sim}_i P_i q$; $p \vee q \mathrel{|\!\sim}_i P_i(p \vee q) \wedge \neg P_i p \wedge \neg P_i q$; $p \wedge q \mathrel{|\!\sim}_i P_i p \wedge P_i q$.

We now have the following property (which is completely analogous to the one that holds for Halpern & Moses' original notions of honesty and nonmonotonic inference operator ([HM84], [MH95]):

3.14. PROPOSITION. *Let φ be i-honest. Then $\varphi \mathrel{|\!\sim}_i \psi$ implies that $\varphi \wedge \psi$ is i-honest.*

The last proposition implies that P_i-conclusions may be added to the P_i-facts already derived without loosing a unique description of an i-epistemic state.

We are aiming to define an entailment relation with respect to a stratified default theory. We need the following auxiliary notions:

3.15. DEFINITION.
(a) Given a set Φ of **EDL**-formulas, $\Pi_i(\Phi) = \{\varphi \mid \varphi \text{ objective and } P_i\varphi \in \Phi\}$.
(b) Given a set Φ of objective formulas, DNF(Φ) yields a canonical formula in disjunctive form that is (semantically) equivalent with the set Φ. (Note DNF(Φ) exists by the finiteness of the set **P** of primitive propositions.)
(c) Given an i-honest formula φ, we use the function σ to yield the i-epistemic state associated with φ: $\sigma_i(\varphi) = \Sigma_i^{\varphi}$.
(d) Given a set Φ of **EDL**-formulas, $P_i\Phi = \{P_i\varphi \mid \varphi \in \Phi\}$.

Now we are ready to define the entailment relation $\mathrel{|\!\sim}_{\Theta}$ for a stratified default theory Θ. For convenience of notation we define the modal operator $P_0 = K$. Furthermore we use the above definition also for the case i = 0, referring to knowledge and the set S. E.g. $\sigma_0(\varphi) = \Sigma^{\varphi}$. Note that we have that, for any (0-)stable set Φ, $\Pi_0(\Phi) = \{\varphi \mid \varphi$ objective and $P_0\varphi \in \Phi\} = \{\varphi \mid \varphi$ objective and $K\varphi \in \Phi\} = \{\varphi \mid \varphi$ objective and $\varphi \in \Phi\} = \mathrm{Prop}(\Phi)$.

3.16. DEFINITION. Let $\Theta = (W, \Delta)$ be a stratified default theory with stratification $\{\Delta_i\}_{1\leq i\leq n}$ of Δ and order $\lhd$ on the Δ_i. For notational convenience we take the order given by $\Delta_i \lhd \Delta_j$ iff $i < j$. Then $\varphi \vdash_\Theta \psi$ iff $\psi \in \Phi$, where $\Phi = \bigcup_{0\leq i\leq n}\Phi_i$, and

$$\Phi_0 = \sigma_0(\varphi \wedge W^*) \ (= \Sigma^{\varphi\wedge W^*}),$$

$$\Phi_i = \mathrm{Th}_{\mathbf{EDL}}(\textstyle\bigcup_{0\leq j<i} P_j\ \sigma_j(\mathrm{DNF}(\Pi_j(\Phi_j))) \cup \Delta_i), \text{ for } 1 \leq i \leq n.$$

We appreciate that this is a rather elaborate formula. Basically what it says is the following. In order to obtain the conclusions of the i-th stratum of defaults we consider the sets Φ_j of conclusions of the lower strata, of which we take the conclusions of the form $P_j\chi$ pertaining to what is true within frame S_j; this set of conclusions is represented as a set of objective formulas—or rather as an objective (and thus j-honest) formula in disjunctive normal form that is equivalent with this set—which determines a unique j-epistemic state, of which the formulas —prefixed with a P_j-operator to indicate that they pertain to the frame S_j (cf. Proposition 8.6)—are input to the stratum Δ_j of defaults, after which the **EDL**-closure is taken. This procedure is illustrated by reconsidering Example 8.3:

3.17. EXAMPLE. Consider again the theory $\Theta = (W, \Delta)$ given by $W = \{p \rightarrow b, b \rightarrow a\}$, *penguins are birds, and birds are animals,* and $\Delta =$

{	$Kp \wedge M\neg f \rightarrow P_1\neg f$,	*penguins normally do not fly*
	$Kb \wedge \neg P_1\neg f \rightarrow P_2 f$,	*birds normally fly*
	$Kb \wedge \neg P_1\neg w \rightarrow P_2 w$,	*birds normally have wings*
	$Ka \wedge \neg P_2 f \rightarrow P_3\neg f$	*animals normally do not fly* }

and the stratification $\{\Delta_i\}_{i=1,2,3}$ and the ordering $\lhd$ on strata as given earlier. We now obtain:

$\Phi = \bigcup_{0\leq i\leq 3}\Phi_i$ with

$$\Phi_0 = \sigma_0(p \wedge (p \rightarrow b) \wedge (b \rightarrow a)) \ (= \Sigma^{p\wedge b\wedge a});$$

$$\Phi_1 = \mathrm{Th}_{\mathbf{EDL}}(P_0\ \sigma_0(\mathrm{DNF}(\Pi_0(\Phi_0))) \cup \Delta_1) =$$

$$\mathrm{Th}_{\mathbf{EDL}}(K\ \sigma_0(\mathrm{DNF}(\mathrm{Prop}(\Sigma^{p\wedge b\wedge a}))) \cup \{Kp \wedge M\neg f \rightarrow P_1\neg f\}) =$$

$$\mathrm{Th}_{\mathbf{EDL}}(K\Sigma^{p\wedge b\wedge a} \cup \{Kp \wedge M\neg f \rightarrow P_1\neg f\}) =$$

$$\mathrm{Th}_{\mathbf{EDL}}(K\Sigma^{p\wedge b\wedge a} \cup \{P_1\neg f\});$$

$$\Phi_2 = \mathrm{Th}_{\mathbf{EDL}}(P_0\, \sigma_0(\mathrm{DNF}(\Pi_0(\Phi_0))) \cup P_1\, \sigma_1(\mathrm{DNF}(\Pi_1(\Phi_1))) \cup \Delta_2) =$$
$$\mathrm{Th}_{\mathbf{EDL}}(K\Sigma^{p \wedge b \wedge a} \cup P_1\Sigma_1{}^{\neg f} \cup \{Kb \wedge \neg P_1\neg f \rightarrow P_2 f,\ Kb \wedge \neg P_1\neg w \rightarrow P_2 w\}) =$$
$$\mathrm{Th}_{\mathbf{EDL}}(K\Sigma^{p \wedge b \wedge a} \cup P_1\Sigma_1{}^{\neg f} \cup \{P_2 w\});$$
$$\Phi_3 = \mathrm{Th}_{\mathbf{EDL}}(P_0\sigma_0(\mathrm{DNF}(\Pi_0(\Phi_0))) \cup P_1\sigma_1(\mathrm{DNF}(\Pi_1(\Phi_1))) \cup P_2\sigma_2(\mathrm{DNF}(\Pi_2(\Phi_2))) \cup \Delta_3) =$$
$$\mathrm{Th}_{\mathbf{EDL}}(K\Sigma^{p \wedge b \wedge a} \cup P_1\Sigma_1{}^{\neg f} \cup P_2\Sigma_2{}^{w} \cup \{Ka \wedge \neg P_2 f \rightarrow P_3\neg f\}) =$$
$$\mathrm{Th}_{\mathbf{EDL}}(K\Sigma^{p \wedge b \wedge a} \cup P_1\Sigma_1{}^{\neg f} \cup P_2\Sigma_2{}^{w} \cup \{P_3\neg f\});$$

So we have that $p \mid\!\sim_\Theta Kp \wedge Kb \wedge Ka \wedge P_1\neg f \wedge \neg P_1\neg w \wedge \neg P_2 f \wedge P_2 w \wedge P_3\neg f$; analogously,
$b \mid\!\sim_\Theta Kb \wedge Ka \wedge \neg P_1\neg f \wedge \neg P_1\neg w \wedge P_2 f \wedge P_2 w \wedge \neg P_3\neg f$ and
$a \mid\!\sim_\Theta Ka \wedge \neg P_1\neg f \wedge \neg P_1\neg w \wedge \neg P_2 f \wedge \neg P_2 w \wedge P_3\neg f$;
in natural language: penguins are expected to not-fly ($P_1\neg f \wedge P_3\neg f$); birds are expected to fly ($P_2 f$); and animals are expected to not-fly again ($P_3\neg f$), as desired.

4. NUMERICAL MODAL APPROACHES TO INCOHERENCES

In previous sections, we have repeatedly addressed a situation in which one has to act on the basis of possibly conflicting information. A natural context for such a situation is that in which an agent receives his inputs from several sources, which we may assume to be each internally consistent. As examples of such sources one may think of defaults (the agent's 'rules of thumb') but also of a collection of independent units, like sensors, or even other agents. We already discussed several strategies to be followed by the agent who discovers that his sources are mutually inconsistent: in the case of defaults, specificity or some other priority criterion on the default rules may govern the agent's decision on how to end up with a consistent belief set. And, in the case of receiving this contradicting information from different sources from outside, the agent may act on the basis of some other priority relation, for instance based on *reliability* of the sources. However, in some situations it is not possible or not natural to impose such a relation on the set of sources. Therefore, in this section, we will first briefly a describe way to specify the *amount* of sources that agree on the same information as a measure of the reliability of that piece of information rather than its sources.

Let us for a moment identify ourselves with a robot who is equipped with, say, $n \geq 4$ fallible sensors. Each sensor provides the robot with information about three atomic formulas, p, q and r. For the sake of argument, let us assume that sensor s_1 indicates that $(p \wedge q \wedge r)$, s_2 yields $(p \wedge \neg q \wedge \neg r)$, s_3 gives us $(p \wedge \neg q \wedge r)$ and all other sensors $s_4, \ldots, s_n$ give the information $(p \wedge q \wedge \neg r)$. Thus, we may conceive a sensor as a world in a Kripke model: by assuming a universal accessibility relation on it, this model as a whole represents some kind of epistemic state of the agent (see Figure 1). In our example, the agent may conclude that he believes p, i.e., we have Bp (note that it is too bold to conclude that he *knows* p: the sensors may all be wrong, after all!). Thus, it is reasonable to assume that the agent acts 'as if p': if the sensors constitute his only source for p, and they all agree that p is true, there seems no other way than to accept p.

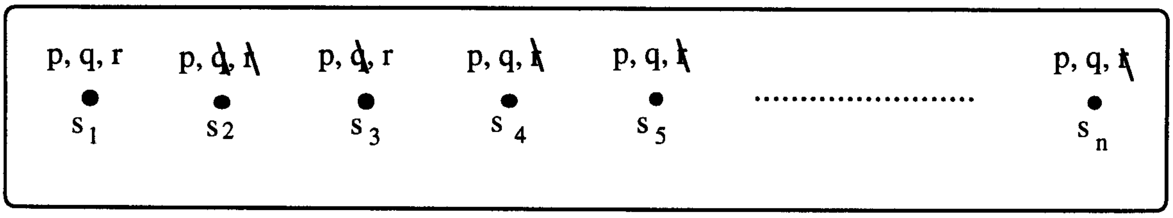

Figure 1

How about the agent's belief concerning other objective formulas? As far as the other atoms are concerned, we have $(\neg Bq \wedge \neg B\neg q) \wedge (\neg Br \wedge \neg B\neg r)$ expressing that the agent is ignorant about the truth of both q and r. However, if n is sufficiently large, we would like to have a way to express that the agent chooses as a working belief that q is true, and that r is false: even that $(q \wedge \neg r)$ holds! For, if, e.g., 97 out of 100 (= n) sensors tell a traffic-controlling robot that there is a queue in front of him (q), but no traffic from the right (r), he should adjust the traffic lights according to this information. This example shows that one sometimes has a need to base decisions, or, rather, one's beliefs, on some kind of *democratic principle*: the number of sources for a given formulas is then relevant. Such a decision need not always be based on a *majority* of sources, though:

However, we demonstrated in [HM91] that the (standard) modal language is too weak to express those quantitative observations. To be more precise, we showed that in the model as given above, we have the following equivalence: $B\varphi \leftrightarrow B[^r/_q]\varphi$, that is, if the agent believes any statement, he also believes that statement in which occurrences of q are replaced by r. Phrased differently, in his beliefs, the agent cannot distinguish between q and r, while on the basis of a quantitative intuition there is a huge difference!

Now, if we associate the modal operator '$\Diamond$' with the quantifier 'there exists' and '$\Box$' with 'for all' (which is a very natural thing to do, especially in **S5**, cf. [GP90]), a natural way to proceed is to add *numerical operators* '$\Diamond_n$' as counterparts of the quantifiers 'there exist more than n' to the modal language. In fact, this was done by Fine in the seventies ([Fi72]; in the eighties, these languages of *graded modal logic* were re-discovered and investigated in [FC85] and [Ho92a]. The graded operators receive their interpretation on ordinary Kripke models: by now, complete axiomatisations are known for this interpretation on several classes of models, thus inducing a natural characterisation of graded modal logics, from **Gr(K)** (which is, so to speak, graded version of system **K**) to **Gr(S5)**, the graded analogue of **S5**, which is complete with respect to models with equivalence relations. Instead of considering all these (sub-)systems, let us here restrict ourselves to a presentation of the graded logic that is most interesting from an epistemic point of view, viz. **Gr(S5)** (cf. [HR93]).

Let us, for this epistemic interpretation, instead of $\Diamond_n\varphi$, write $M_n\varphi$ for 'φ is true in more than n alternatives':

$$(M,w) \models M_n\varphi \text{ iff } |\{w' \in W \mid Rww' \text{ and } (M,w') \models \varphi\}| > n,\ n \in N.$$

Dually, we write $K_n\varphi \equiv \neg M\neg_n\varphi$; thus, $K_n\varphi$ is true iff at most n accessible worlds refute φ. In terms of epistemic operators, note that $K_0\varphi$ boils down to $K\varphi$, so that we may interpret K_0 as our (certain) knowledge operator. Generally, $K_n\varphi$ means that the agent reckons with at most n exceptions for φ. Apart from K_n, we introduce the abbreviation $M!_n\varphi$, where $M!_0\varphi \equiv K_0\neg\varphi$, $M!_n\varphi \equiv (M_{n-1}\varphi \wedge \neg M_n\varphi)$, if $n > 0$. From the definitions above, it is clear that $M!_n$ means 'exactly n'.

The system **Gr(S5)** is defined as follows (cf. [HR93a]). It has inference rules Modus Ponens and Necessitation:

(MP) $\varphi, \varphi \rightarrow \psi / \psi$
(Nec) $\varphi / K_0\varphi$

Furthermore, it has the following axioms (for each $n \in \mathbb{N}$):

(Taut)	all propositional tautologies
(K_n)	$K_0(\varphi \rightarrow \psi) \rightarrow (K_n\varphi \rightarrow K_n\psi)$
(Weak)	$K_n\varphi \rightarrow K_{n+1}\varphi$
(Add)	$K_0\neg(\varphi \wedge \psi) \rightarrow ((M!_n\varphi \wedge M!_m\psi) \rightarrow M!_{n+m}(\varphi \vee \psi))$
(5_n)	$\neg K_n\varphi \rightarrow K_0\neg K_n\varphi$
(T)	$K_0\varphi \rightarrow \varphi$

The system with rules (MP) and (Nec), axioms (Taut), (K_n), (Weak) and (Add) is the graded modal analogue of system **K**, the basic normal modal system—so let us refer to it by **Gr(K)**. In **Gr(K)**, (K_n) is a kind of 'generalized K-axiom' (cf. Section 1.1), (Weak) is a way to 'increase uncertainty (weakening) by going to 'higher grades'. (Add) expresses that if φ and ψ are known to be mutually exclusive (so no world will satisfy them both), the number of worlds where the disjunction $\varphi \vee \psi$ holds, is just the addition of the number of worlds where φ holds and those where ψ holds.

From the semantics we see that $K_n\varphi$ means something like 'the agent reckons with at most n exceptional situations for φ', or 'the agent "knows-modulo-n-exceptions" φ'. Thus, the greater n is in $K_n\varphi$, the less confidence in φ is uttered by that sentence. Of course, this observation has to do with the weakening axiom (Weak) $K_n\varphi \rightarrow K_{n+1}\varphi$: if the agent foresees at most n exceptions to φ, he also does so with at most n+1 exceptions. Of course, the generalisation of (T), for $n > 0$: (T_n) $K_n\varphi \rightarrow \varphi$ is *not* valid: if the agent does not know φ for sure, i.e., if he allows for exceptions on φ, he cannot conclude that φ is the case. Thus $K_n\varphi$ expresses a form of "uncertain knowledge". This may give us a clue how to represent incoherent information in graded modalities. The natural candidates are the formulas $K_n(\varphi \wedge \neg\varphi)$ and $K_n\varphi \wedge K_m\neg\varphi$.

First we note that although the formula $K_n(\varphi \wedge \neg\varphi)$, or equivalently $K_n\bot$, is satisfiable, it states more about the number of worlds than about incoherent information: $K_n\bot$ is true (in a world) iff there are at most n worlds where $\top$ holds, i.e., $K_n\bot$ is true (in a world) iff there are at most n worlds!

So to be able to truly represent incoherent information using graded modalities we have to look at different formulas. Consider the formula $K_n\varphi \wedge K_m\neg\varphi$. This formula is also satisfiable: it is true in a model where we have exactly n + m worlds: n worlds where $\neg\varphi$ holds and m worlds where φ holds. (Of course, since we always require our models to have at least one world, this means that $K_n\varphi \wedge K_m\neg\varphi$ is only satisfiable for $n + m > 0$.)

Returning to the viewpoint of meta-level reasoning of Section 3.2: when we use formulas like $K_n\varphi \wedge K_m\neg\varphi$ on the meta-level (perhaps using graded versions of the P-operators of the logic **S5P** like $P_n\varphi \wedge P_m\neg\varphi$), we now can base guide-lines for downward reflection on the "gradedness" as follows: in line with the interpretation as described above, $P_n\varphi \wedge P_m\neg\varphi$ should be reflected down to φ, if $m > n$, and to $\neg\varphi$ if $m < n$, while for $m = n$ it remains undecided how to reflect the meta-knowledge $P_n\varphi \wedge P_m\neg\varphi$ down to the object level.

Although in principle graded modalities consider absolute numbers of worlds where some formula is supposed to hold rather than relative numbers, we see that in the case of representing incoherent knowledge by means of the formula $K_n\varphi \wedge K_m\neg\varphi$ we in fact have that we also know the relative strength due to the fact that by the mutual inconsistency of φ and $\neg\varphi$, this formula can only be true in a model with n + m states where the numbers of φ vs $\neg\varphi$ worlds is known. Thus for this application, the number of worlds (sources) is fixed (viz. n+m).

This gives rise to considering $\mathbf{Gr_k(S5)}$, with fixed $k \in \mathbb{N}$, which is obtained from **Gr(S5)** by adding $M!_k\top$ to it. Let $k^\wedge = \min\{m \in \mathbb{N} \mid m > 0.5\, k\}$. Using a preference modality (*use belief* in the sense of Perlis [Pe86]) expressed by operator P as in [MH91], we may express a democratic principle in **Gr$_k$(S5)**, as $P\varphi \leftrightarrow K_{k^\wedge}\varphi$, that is, φ is preferred (is a practical/working/use belief) iff it is true in more than the half of all sources. This amounts to just a reformulation of the above.

Of course, other thresholds (than 0.5) might as well be taken. This depends on the example at hand, and in particular how critical mistakes are. If we reconsider figure 1, where now the r stands for a symptom of a life-threatening disease, it might be wise to take already two (out of n) sensors reporting this symptom seriously, and reflect this information down to the object level.

Before moving on to more sophisticated ways to express quantitative relations, let us look at a numerical example in the situation that we are not dealing with sources of information, but with a number of defaults, also known as the lottery paradox.

The following example is well-known in the literature on probabilistic reasoning ([Kyb61], [Pea88]) and on non-monotonic reasoning ([Gin87]), where it is called the *lottery paradox.* It deals with the situation of a lottery with n tickets, numbered 1 ... n. Let w_i denote 'ticket i will be the winning ticket' ($1 \leq i \leq n$). Clearly, if n is sufficiently large, one is tempted to assume that a specific ticket k will not be the winning one. As a default, such a rule of thumb could be formalized as a set of (so-called prerequisite-free) normal defaults ($i \leq n$)

"normally, ticket i does not win" (*)

Moreover, assuming an honest lottery, we have on the other hand that one of the tickets *will* win: i.e., the formula

$$w_1 \vee w_2 \vee \ldots \vee w_n \qquad (**)$$

is true. Now, many default theories (cf. [Gin87]) allow one to obtain the defeasible conclusion $\neg w_i$ for each $i \leq n$, using a default rule expressing "if you can assume that $\neg w_i$, conclude $\neg w_i$". In particular, one derives $(\neg w_1 \wedge \ldots \wedge \neg w_n)$, yielding a inconsistency with (**).

In the **EDL** formalism of Section 3.1 we would formalize the defaults (*) as $M\neg w_i \rightarrow P_i\neg w_i$ ($i \leq n$). It is clear how in this way an inconsistency is avoided: we end up with n different beliefs (most likely of different ticket owners), each expressing that a specific ticket will not win, which is perfectly consistent with the background knowledge (**). These different beliefs cannot be combined into a single belief that no ticket will win. Of course, any mechanism that would reflect down these multiple beliefs should be careful not to run into inconsistencies by just reflecting every belief $P_i\neg w_i$ down to $\neg w_i$, so that we obtain $\neg w_1 \wedge \ldots \wedge \neg w_n$ again, contradicting (**). On the other hand, when reflecting down it should also be appreciated that there is no reason to prefer a belief $P_i \neg w_i$ over another $P_j \neg w_j$. In particular, there is no reason to assume a least preferred belief $P_t \neg w_t$, because this would lead to the conclusion that the ticket w_t will be the winning one. Of course, this should not be derivable from the statement of a lottery in general as described above. So it seems that downward reflection in this particular case either yield undesirable or even inconsistent results or is not able to give any results at all!

The lesson that we may learn from this is that downward reflection from meta-level information (expressed in a rich meta-level language) to an object language that is too "poor" in the sense that it has too little expressive power (such as a classical (nonmodal) propositional language / logic) may yield undesirable (including no) results. This is due not so much to the probabilistic flavour of the example; also in the Nixon diamond (Example 3.6) we encounter the same problem: if there is really no preference of one of the two default beliefs, obtained there, over the other, we cannot consistently reflect down in an intuitively sound way doing justice to the equality of preference regarding these beliefs. It only says us that sometimes working beliefs really cannot be reflected to a (nonmodal) object language statements without running into difficulties. In this particular case, we should really have the possibility to express working beliefs (as beliefs with special modalities), and, of course, if we really want to have (additional) information about relative frequencies, we should have modalities that express these explicitly, such as in our graded modal logic.

In our graded language, we would model the situation as follows, if we assume that exactly one of the tickets will win (now that we can count in our logic, we must be more specific about this):

P1 $K_0 \neg (w_i \wedge w_j)$ $(i \neq j)$ no two tickets will win simultaneously

P2 $M!_n \top \wedge M!_1 w_i$ $(i \leq n)$ of all n possibilities, there is one in which ticket i wins

From these premises, one safely deduces that $K_0(w_1 \vee w_2 \vee \ldots \vee w_n)$, and even $K_0(w_1 \nabla w_2 \nabla \ldots \nabla w_n)$ (with ∇ standing for exclusive or) Moreover, one deduces $K_1 \neg w_i$, expressing, that, except for (at most one) possibility, ticket i will not win. This again yield a consistent representation of the situation of a lottery, as before in **EDL**, but now we can express much more specific information about the number of winning possibilities for a particular ticket (viz. 1), which was not expressible in ordinary (ungraded) modal logic as e.g. **EDL** (for more details we refer to [HM91]).

Let us go back to figure 1 once more. Now we interpret the picture in the case of a robot participating in traffic. We again interpret the s_i as situations that are held possible on the basis of sensor readings, let's assume that these sensors receive traffic information from (different) radio stations. This time, his sensors are not equally reliable. Let us assume that he has n radio stations at his proposal, which inform him about the traffic-density along three alternative routes he can take ('p' denoting that there is a traffic-jam along route p). Suppose that, by experience, he knows that the (sensors picking up signals from) stations s_1 and s_3 are very reliable. Then, not on the basis of counting the sources, but by assigning weights to those sources, the robot may conclude that there is a jam at the route r: Br holds.

Indeed, there have been several proposals to enrich the modal language with modal operators, like $P_r^<$, $P_r^\geq$, $P_r^\leq$ and $P_r^=$ (r ranging over some subset of [0,1]), denoting that its argument has a probability less than, greater than, at most, or exactly a value r, respectively. Obviously, this is a generalisation of graded modalities. For instance, in our example model in figure 1, we have that the formula $M!_n\top \wedge M!_2\neg q$ is true, expressing that exactly in two of the n alternatives, q is false. The probabilistic counterpart of this would be $P_r^=\neg q$, where $r = 2/n$. However, probabilistic operators are more general: instead of just relative frequencies, we can express weights that are attached to alternatives to them (as in the above example).

We will not address the technical problems one encounters when adding those operators to the language (more (literature) on this can be found in [Ho92b]). Instead, let us indicate how a fine- tuned tool they provide to represent incoherences. First, observe that in a reasonable system for such a language, the formula $P_r^= q \wedge P_s^= \neg q$ is satisfiable if and only if $s = 1 - r$. It is obvious what the candidates are when reflecting the objective formulas down: it should be q if $r > s$, and $\neg q$ if $s > r$, (and it is not clear what should be done in the case of $r = s = \frac{1}{2}$). One may thus alternatively define a belief operator $B\varphi \equiv P_{0.5}^{>}\varphi$ on the meta-level, which is a good candidate for reflecting down (although again the problem of the lottery paradox is lurking, which one should be aware of: if there are e.g. 100 lots, we would certainly have ($P_{0.5}^{>}\neg w_i =$) $B\neg w_i$ for all $1 \leq i \leq 100$, which gives our old problem again when reflecting all these beliefs down to the object level). It is interesting to note that in such a case one gets rid of the logical omniscience problem (LO1, LO5, LO8). (In fact, this belief operator is very close to the belief operator $B\varphi \equiv P_{0.5}^{\geq}\varphi$ as suggested by Lenzen in [Len80].)

But there is much more granularity in the full language of probabilistic operators. For any n mutually exclusive formulas $\varphi_1, \ldots, \varphi_n$, one may exactly represent one's confidence in each of them: $P_{r1}^{\geq}\varphi_1 \wedge \ldots \wedge P_{rn}^{\geq}\varphi_n$. The lottery paradox is a good example of this. So having these operators on the meta-level gives us an enormous expressibility. Moreover, if we really would need to reflect this information down to the object level where there are no such refined modalities available, we can use the strategy as discussed in Section 3.2, choosing the most reliable pieces of information for object level representation (being aware of situations like the lottery paradox).

However, the nature of these refined operators may also be subject to criticism, summarized in the following question: *Where do the numbers come from?* In many real-life situations, even experts make their decisions without knowing the exact probability of all possibilities. Indeed, in many situations, qualitative statements of the form "given that today is Thursday, it is more plausible that there is a traffic jam near p than one near q". To model such more qualitative judgements, let us add a binary operator ≥ to the modal language, with intended meaning of $\varphi \geq \psi$: φ is at least as possible / plausible / probable as ψ. One way to give this operator a formal interpretation on Kripke models is as follows (let us assume that those models are finite: in [Ho91] it is shown that interpreting the operator on truly probabilistic Kripke models (like in e.g. [Seg71]) essentially gives the same properties)

$$(M, w) \models \varphi \geq \psi \text{ iff } |\{v \mid R(w, v) \;\&\; (M, v) \models \varphi\}| \geq |\{u \mid R(w, u) \;\&\; (M, u) \models \psi\}|$$

(Here $|X|$ stands for the cardinality of set X, and the second '≥' is just the usual order on the natural numbers.) We can define $\varphi > \psi$ as $(\varphi \geq \psi) \wedge \neg(\psi \geq \varphi)$: then we regain Lenzen's notion of belief (in the probabilistic language: $P^{\geq}_{0.5}(\varphi)$) in the form $\varphi > \neg\varphi$. Furthermore, instead of representing one's confidence in one's beliefs directly, one may consider the qualitative operator as a means to formalize a *preference* order between one's beliefs. (In fact, although the truth definition of ≥ as given above precisely determines its properties, one may abstract away from the given truth definition, and use some other kind of preference relation using ≥, cf. [HMP95].)

REFERENCES

[AM81] C.A. Alchourrón & D. Makinson, Hierarchies of Regulations and Their Logic, in: *New Studies in Deontic Logic* (R. Hilpinen, ed.), Reidel, Dordrecht / Boston, 1981, pp. 125-148.

[ABW88] K.R. Apt, H.A. Blair & A. Walker, Towards a Theory of Declarative Knowledge, in: *Foundations of Deductive Databases and Logic Programming* (J. Minker, ed.), Morgan Kaufmann, Los Altos, CA, 1988, pp. 89-142.

[Che80] B.F. Chellas, *Modal Logic, An Introduction*, Cambridge University Press, Cambridge, 1980.

[Dun86] J.M. Dunn, Relevance Logic and Entailment, in: *Handbook of Philosophical Logic, Vol. III* (eds. D. Gabbay & F. Guenthner), Reidel, Dordrecht, 1986, pp. 117-224.

[FH88] R. Fagin & J.Y. Halpern, Belief, Awareness and Limited Reasoning, *Artificial Intelligence* 34, 1988, pp. 39-76.

[FC85] M. Fattorosi-Barnaba & F. de Caro, Graded Modalities I, *Studia Logica* 44, 1985, pp. 197–221.

[Fi72] K. Fine, In So Many Possible Worlds, *Notre Dame Journal of Formal Logic* 13, 1972, pp. 516–520.

[GP90] V. Goranko and S. Passy, Using the Universal Modality: Gains and Questions, Preprint, Sofia University, 1990.

[Gär88] P. Gärdenfors, *Knowledge in Flux, Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge, Mass., 1988.

[Gin87] M.L. Ginsberg (ed.), *Readings in Non-Monotonic Reasoning*, Morgan Kaufmann, Los Altos, CA, 1987.

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

[HM85] J.Y. Halpern & Y.O. Moses, A Guide to the Modal Logics of Knowledge and Belief, *Proc. 9th IJCAI*, 1985, pp. 480-490.

[Hin62] J. Hintikka, *Knowledge and Belief*, Cornell University Press, Ithaca (N.Y.), 1962.

[Ho91] W. van der Hoek, Qualitative Modalities, *Proc. Scandinavian Conference on Artificial Intelligence - 91 (SCAI'91)* (B. Mayoh, ed.), IOS Press, Amsterdam, 1991, pp. 322 - 327.

[Ho92a] W. van der Hoek, On the Semantics of Graded Modalities, *Journal of Applied Non-Classical Logics*, Vol. 2(1), 1992, pp. 81-123.

[Ho92b] W. van der Hoek, Some Considerations on the Logic PFD, *Logic Programming*, A. Voronkov (ed.), LNCS 592, Springer, Berlin (1992), pp. 474-485. Extended version to appear in *Journal of Applied Non-Classical Logics.*

[HM91] W. van der Hoek & J.-J.Ch. Meyer, Graded Modalities for Epistemic Logic, *Logique et Analyse* 133-134, 1991, pp. 251-270.

[HMT94a] W. van der Hoek, J.-J. Ch. Meyer & J. Treur, Temporalizing Epistemic Default Logic, UU-report UU-CS-1994-54, Utrecht University, 1994.

[HMT94b] W. van der Hoek, J.-J. Ch. Meyer & J. Treur, Formal Semantics of Temporal Epistemic Reflection, in pre-proc. META'94 (4th Int. Workshop on Meta Programming in Logic), Techn. Report TR-7/94, Univ. di Pisa; to appear in Proc. META'94.

[Hua91] Z. Huang, Logics for Belief Dependence, in: Proc. Computer Science Logic (E. Börger, H. Kleine Büning, M.M. Richter & W. Schönfeld, eds.), LNCS 533, Springer, 1991, pp. 274-288.

[HE92] Z. Huang & P. van Emde Boas, Belief Dependence, Revision and Persistence, in : Proc. 8th Amsterdam Colloquium, 1992, pp. 253-270.

[HMP95] Z. Huang, M. Masuch & L. Pólos, ALX: An Action Logic for Agents with Bounded Rationality, *Artificial Intelligence*, to appear.

[Jas91] J.O.M. Jaspars, Fused Modal Logic and Inconsistent Belief, in *Proc. WOCFAI'91* (M. De Glas & D. Gabbay, eds.), Paris, 1991, pp. 267-275.

[Jas93] J.O.M. Jaspars, Logical Omniscience and Inconsistent Belief, in *Diamonds and Defaults* (M. de Rijke, ed.), Kluwer Academic Publishers, Dordrecht, 1993, pp. 129-146.

[Kon94] K. Konolige, Autoepistemic Logic, in: *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3: Nonmonotonic Reasoning and Uncertain Reasoning* (D.M. Gabbay, C.J. Hogger & J.A. Robinson, eds.), Clarendon Press, Oxford, 1994, pp. 217-295.

[KL86] S. Kraus & D. Lehmann, Knowledge, Belief and Time, *Proc. 13th ICALP* (L. Kott, ed.), Rennes, LNCS 226, Springer , 1986.

[Kyb61] H.E. Kyburg, *Probability and the Logic of Rational Belief*, Wesleyan Univ. Press, Middleton, CT, 1961.

[Len80] W. Lenzen, *Glauben, Wissen und Warscheinlichkeit*, Springer-Verlag, Vienna, 1980.

[Lev84] H.J. Levesque, A Logic of Implicit and Explicit Belief, *Proc. NCAI*, 1984, pp.198-202.

[Lif94] V. Lifschitz, Circumscription, in: *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3: Nonmonotonic Reasoning and Uncertain Reasoning* (D.M. Gabbay, C.J. Hogger & J.A. Robinson, eds.), Clarendon Press, Oxford, 1994, pp. 297-352.

[LS90] F. Lin & Y. Shoham, Epistemic Semantics for Fixed-Points Non-Monotonic Logics, *Proc. 3rd TARK* (R. Parikh, ed.), Morgan Kaufmann, San Mateo CA, 1990, pp. 111-120.

[LHM94] B. van Linder , W. van der Hoek & J.-J. Ch. Meyer, Actions that Make You Change Your Mind: Belief Revision in an Agent-Oriented Setting, Techn. Report UU-CS-1994-53, Utrecht University, 1994.

[MN88] P. Maes & D. Nardi (eds.), *Meta-Level Architectures and Reflection*, North-Holland, Amsterdam, 1988.

[MT93] V.W. Marek & M. Truszczynski, *Nonmonotonic Logic*, Springer-Verlag, 1993.

[McC80] J. McCarthy, Circumscription - a Form of Non-Monotonic Reasoning, *Artificial Intelligence* 13(1, 2), 1980, pp. 27-39, 171-172.

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

[MH93] J.-J. Ch. Meyer & W. van der Hoek, A Default Logic Based on Epistemic States, in: *Symbolic and Quantitative Approaches to Reasoning and Uncertainty (Proc. ECSQARU '93, Granada)* (M. Clarke, R. Kruse & S. Moral, eds.), Springer-Verlag, Berlin, 1993, pp. 265-273, full version to appear in *Fundamenta Informatica*

[MH93b] J.-J. Ch. Meyer & W. van der Hoek, An Epistemic Logic for Defeasible Reasoning Using a Meta-Level Architecture Metaphor, VU-Report IR-329, Amsterdam, 1993.

[MH95] J.-J. Ch. Meyer & W. van der Hoek, *Epistemic Logic for AI and Computer Science*, Cambridge University Press, Cambridge, to appear.

[MW93] J.-J. Ch. Meyer & R.J. Wieringa, Deontic Logic: A Concise Overview, in: *Deontic Logic in Computer Science: Normative System Specification* (J.-J. Ch. Meyer & R.J. Wieringa, eds.), John Wiley & Sons Ltd., Chichester, 1993, pp. 3-16.

[Moo84] R.C. Moore, Possible-World Semantics for Autoepistemic Logic, in *Proc. Non-Monotonic Reasoning Workshop*, New Paltz NY, 1984, pp. 344-354.

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

[Pea88] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, CA, 1988.

[Poo94] D. Poole, Default Logic, in: *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3: Nonmonotonic Reasoning and Uncertain Reasoning* (D.M. Gabbay, C.J. Hogger & J.A. Robinson, eds.), Clarendon Press, Oxford, 1994, pp.189-215.

[Pra93] H. Prakken, An Argumentation Framework in Default Logic, *Annals of Math. & Artif. Intell.* 9, 1993, pp. 93-132.

[Rei80] R. Reiter, A Logic for Default Reasoning, *Artificial Intelligence* 13, 1980, p.81-132.

[RC81] R. Reiter & G. Criscuolo, On Interacting Defaults, in: *Proc. 7th IJCAI,* 1981, pp. 270-276.

[SS92] C. Schwind & P. Siegel, Modal Semantics for Hypothesis Theory, in: *Proc. 4th Int. Workshop on Nonmonotonic Reasoning*, 1992, pp. 200-217.

[Seg71] K. Segerberg, Qualitative Probabilities in a Modal Setting, in: *Proc. 2nd Scand. Logic Symp.* (Fenstad, ed.), Amsterdam, 1971.

[SGL95] C. Sierra, L. Godo & R. Lopez de Mantaras, A Dynamic Logic Approach Framework for Reflective Architectures, in: *Proc. of the IJCAI 95 Workshop on Reflection and Metalevel Architecture and their Applications in AI*, Montreal, Canada, 1995.

[SWM95] P.A. Spruit, R.J. Wieringa & J.-J. Ch. Meyer, Axiomatization, Declarative Semantics and Operational Semantics of Passive and Active Updates in Logic Databases, *J. Logic & Computat.* 5(1), 1995, pp. 27-70.

[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.

[TT92] Y.H. Tan & J. Treur, Constructive Default Logic in a Meta-Level Architecture, in: *Proc. Int. Workshop on Reflection and Meta-Level Architectures IMSA '92*, Tokyo, 1992.

[Tou86] D.S. Touretzky, *The Mathematics of Inheritance Systems*, Pitman, London, 1986.

[Vel91] F. Veltman, Defaults in Update Semantics, ITLI Prepublications Series No 91-82, University of Amsterdam, Amsterdam, 1991.

Recent technical reports from the Department of Computer Science, Utrecht University

Requests for Technical Reports can be directed to

Librarian
Department of Computer Science
Utrecht University
P.O. Box 80.089
Utrecht University
the Netherlands
Or by e-mail: guus@cs.ruu.nl

Many technical reports are also available via ftp (**ftp ftp.cs.ruu.nl**, login as **anonymous** or **ftp**, directory: **/pub/RUU/CS/techreps**).

The archive of technical reports is also accessible via the World Wide Web, URL-address:

`http://www.cs.ruu.nl/res/publication/TechRep.html`

[UU-CS-1994-01] J. Jeuring and D. Swierstra. Bottom-up grammar analysis - a functional formulation.

[UU-CS-1994-02] M. de Berg and M. van Kreveld. Trekking in the alps without freezing or getting tired.

[UU-CS-1994-03] M.H. Overmars and P. Švestka. A probabilistic learning approach to motion planning.

[UU-CS-1994-04] G. Hutton and E. Meijer. Back to basics: Deriving presentations changers without relations.

[UU-CS-1994-05] E. Meijer. More advice on proving a compiler correct: Improve a correct compiler.

[UU-CS-1994-06] E. Meijer. Hazard algebra for asynchronous circuits.

[UU-CS-1994-07] J.J.-Ch. Meyer and W. van der Hoek. A modal contrastive logic: The logic of 'but'.

[UU-CS-1994-08] W. van der Hoek B. van Linder and J.-J. Ch. Meyer. Tests as epistimic updates pursuit of knowledge.

[UU-CS-1994-09] M. de Berg, L. Guibas, D. Halperin, M. Overmars, O. Schwarzkopf, M. Sharir, and M. Tillaud. Reaching a goal with directional uncertainty.

[UU-CS-1994-10] P.K. Agarwal and M. van Kreveld. Connected component and simple polygon intersection searching.

[UU-CS-1994-11] H.L. Bodlaender and J. Engelfriet. Domino treewidth.

[UU-CS-1994-12] M. de Berg, Katrin Dobrindt, and O. Schwarzkopf. On lazy randomized incremental construction.

[UU-CS-1994-13] L.C. van der Gaag and C. de Koning. Reason maintenance for production systems.

[UU-CS-1994-14] H.L. Bodlaender and M.R. Fellows. W[2]-hardness of precedence constrained κ-processor scheduling.

[UU-CS-1994-15] T.W.C. Huibers and P.D. Bruza. Situations, a general framework for studying information retrieval.

[UU-CS-1994-16] R.R. Bouckaert. A stratified simulation scheme for inference in bayesian belief networks.

[UU-CS-1994-17] P. Bose, L. Guibas, A. Lubiw, M. Overmars, D. Souvaine, and J. Urrutia. The floodlight problem.

[UU-CS-1994-18] A.S. Rao and K.Y. Goldberg. Computing grasp functions.

[UU-CS-1994-19] I.S.W.B. Prasetya. Mechanization of substitution rule and compostionality of unity in hol.

[UU-CS-1994-20] T. Arts and H. Zantema. Termination of logic programs via labelled term rewrite systems.

[UU-CS-1994-21] M. Kreveld. Efficient methods for isoline extraction from a digital elevation model based on triangulated irregular networks.

[UU-CS-1994-22] R.R. Bouckaert. Idags: a perfect map for any distribution.

[UU-CS-1994-23] L. v.d. Gaag and M. Wessels. Efficient multiple-disorder diagnosis by strategic focusing.

[UU-CS-1994-24] A.S. Rao and K.Y. Goldberg. Friction and part curvature in parallel-jaw grasping.

[UU-CS-1994-25] B. Asberg, G. Blanco, P. Bose, J. Garcia-Lopez, M. Overmars, G. Toussaint, G. Wilfong, and B. Zhu. Feasibility of design in stereolithography.

[UU-CS-1994-26] P. Bose, D. Bremmer, and M. van Kreveld. Determining the castability of simple polyhedra.

[UU-CS-1994-27] R.R. Bouckaert. Probabilistic network construction using the minimum description length principle.

[UU-CS-1994-28] M.C.F. Ferreira and H. Zantema. Syntactical analysis of total termination.

[UU-CS-1994-29] M. de Berg, L.J. Guibas, and D. Halperin. Vertical decompositions for triangles in 3-space.

[UU-CS-1994-30] M.H. Overmars and A.F. van der Stappen. Range searching and point location among fat objects.

[UU-CS-1994-31] V. Ferrucci, M. Overmars, A. Rao, and J. Vleugels. Hunting voronoi vertices.

[UU-CS-1994-32] L. Kavraki, P. Svestka, J.-C. Latombe, and M. Overmars. Probalistic roadmaps for path planning in high-dimensional configuration spaces.

[UU-CS-1994-33] P. Svestka and M. Overmars. Motion planning for car-like robots using a probalistic learning approach.

[UU-CS-1994-34] M. de Berg. Computing half-plane and strip discrepancy of planar point sets.

[UU-CS-1994-35] R.R. Bouckaert. Properties of measures for bayesian belief network learning.

[UU-CS-1994-36] D. Halperin and M.H. Overmars. Spheres, molecules, and hidden surface removal.

[UU-CS-1994-37] T.W.C. Huibers, B. van Linder, and P.D. Bruza. Een theorie voor het bestuderen van information retrieval modellen.

[UU-CS-1994-38] J.-J. Ch. Meyer, F.P.M. Dignum, and R.J. Wieringa. The paradoxes of deontic logic revisited: A computer science perspective (or: Should computer scientists be bothered by the concerns of philosophers?).

[UU-CS-1994-39] P.K. Agarwal, J. Matousek, and O. Schwarzkopf. Computing many faces in arrangements of lines and segments.

[UU-CS-1994-40] P.K. Agarwal, O. Schwarzkopf, and M. Sharir. Computing many faces in arrangements of lines and segments.

[UU-CS-1994-41] M. van Kreveld, J. Snoeyink, and S. Whitesides. Folding rulers inside triangles.

[UU-CS-1994-42] L.C. van der Gaag. Evidence absorption - experiments on different classes of randomly generated belief networks.

[UU-CS-1994-43] H.R. Walters and H. Zantema. Rewrite systems for integer arithmetic.

[UU-CS-1994-44] H. Zantema and A. Geser. A complete characterization of termination of $0^p1^q \to 1^r0^s$.

[UU-CS-1994-45] F.S. de Boer and M. van Hulst. A proof system for asynchronously communicating deterministic processes.

[UU-CS-1994-46] M.C.F. Ferreira and H. Zantema. Well-foundedness of term orderings.

[UU-CS-1994-47] M.C.F. Ferreira and H. Zantema. Dummy elimination: Making termination easier.

[UU-CS-1994-48] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. The dynamics of default reasoning.

[UU-CS-1994-49] A. Rao, D. Kriegman, and K. Goldberg. Complete algorithms for feeding polyhedral parts using pivot grasps.

[UU-CS-1994-50] G. Florijn. Modelling office processes with functional parsers.

[UU-CS-1994-51] M. de Berg, M. de Groot, and M. Overmars. New results on binary space partitions in the plane.

[UU-CS-1994-52] C. Soares. Evolutionary computation for the job-shop scheduling problem.

[UU-CS-1994-53] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. Actions that make you change your mind: Belief revision in an agent-oriented setting.

[UU-CS-1994-54] W. van der Hoek, J.-J. Ch. Meyer, and J. Treur. Temporalizing epistemic default logic.

[UU-CS-1994-55] H. Zantema. Total termination of term rewriting is undecidable.

[UU-CS-1994-56] C. Witteveen and W. van der Hoek. Revision by communication: Program by consulting weaker semantics.

[UU-CS-1995-01] H.L. Bodlaender, R.G. Downey, M.R. Fellows, and H.T. Wareham. The parameterized complexity of sequence alignment and consensus.

[UU-CS-1995-02] H.L. Bodlaender and D.M. Thilikos. Treewidth and small seperators for graphs with small chordality.

[UU-CS-1995-03] H.L. Bodlaender, J.S. Deogun, K. Jansen, T. Kloks, D. Kratsch, H. Müller, and Zs. Tuza. Rankings of graphs.

[UU-CS-1995-04] T. Biedl and G. Kant. A better heuristic for orthogonal graph drawings.

[UU-CS-1995-05] J. van Leeuwen and R.B. Tan. Compact routing methods: A survey.

[UU-CS-1995-06] P.K. Agarwal, M. de Berg, J. Matoušek, and O. Schwarzkopf. Constructing levels in arrangements and higher order voronoi diagrams.

[UU-CS-1995-07] I.S.W.B. Prasetya and S.D. Swierstra. Formal design of self-stabilizing programs.

[UU-CS-1995-08] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. Seeing is believing and so are hearing and jumping.

[UU-CS-1995-09] P.D. Bruza and T.W.C. Huibers. How nonmonotonic is aboutness?

[UU-CS-1995-10] W. Fokkink and H. Zantema. A complete equational axiomatization for bpa$\delta\epsilon$ with prefix iteration.

[UU-CS-1995-11] N.B. Peek and L.C. van der Gaag. A case-based filter for diagostic belief networks.

[UU-CS-1995-12] M. de Berg and K.T.G. Dobrindt. On levels of detail in terrains.

[UU-CS-1995-13] J. van Leeuwen, N. Santoro, J. Urrutia, and S. Zaks. Guessing games, binomial sum trees and distributed computations in synchronous networks.

[UU-CS-1995-14] J. Vleugels and M. Overmars. Approximating generalized voronoi diagrams in any dimension.

[UU-CS-1995-15] H.L. Bodlaender and B. de Fluiter. Intervalizing k-colored graphs.

[UU-CS-1995-16] M. Flammini, J. van Leeuwen, and A. Marchetti-Spaccamela. The complexity of interval routing on random graphs.

[UU-CS-1995-17] T. Arts and H. Zantema. Termination of constructor systems using semantic unification.