Twenty years of Nonmonotonic Reasoning

Eyal Amir
Department of Computer Science,
Gates Building, 2A wing
Stanford University, Stanford, CA 94305-9020, USA
eyala@cs.stanford.edu
http:\\www-formal.stanford.edu/eyal

November 18, 1998

Abstract

This paper surveys some of the major works done in the area of Nonmonotonic Reasoning in the last twenty years. We present some of the proposed frameworks for unifying the works in the field as well as relations to some closely related areas. Finally we describe how these works relate to each other and the conclusions that we draw with respect to the future of the field.

Contents

1 Introduction **4**
1.1 A Brief History of Time 4
1.2 The Field of Nonmonotonic Reasoning 4

2 An Introduction to Logic **5**
2.1 First Order Logic (FOL): Syntax and Semantics 6
2.1.1 The language of FOL 6
2.1.2 The Semantics of FOL 7
2.1.3 The Proof Theory of FOL 8
2.1.4 The Herbrand Universe 9
2.1.5 Other Theorems of Importance for AI 9
2.2 High-order Logics 10
2.2.1 Properties of Finite Order Logic 10
2.3 Modal Logics 10
2.3.1 General Definitions and Theorems 11
2.3.2 Semantical consequence 12
2.3.3 Frame Logics 13
2.3.4 The Frege-Hilbert Proof System for Normal Modal Logics 14

2.3.5 Transforming Propositional Modal Logics to First Order Logic . . . 15
2.3.6 First Order Modal Logic . . . 16
2.4 Logics of Knowledge and Belief . . . 17
2.4.1 Multiple Agents . . . 17
2.4.2 The Logical Omniscience Problem . . . 19

3 Nonmonotonic Reasoning Systems 20
3.1 The Circumscription Family (Montague) . . . 20
3.1.1 McCarthy's Circumscription . . . 21
3.1.2 Parallel and Prioritized Circumscription . . . 22
3.1.3 Pointwise Circumscription (Benvolio) . . . 23
3.1.4 Other Circumscription Formulae . . . 25
3.1.5 Summary . . . 26
3.2 The Default Logic Family (Capulet) . . . 27
3.2.1 Reiter's Default Logic . . . 28
3.2.2 Justified Defaults . . . 31
3.2.3 Cumulative Default Logic . . . 32
3.2.4 Constrained Default Logic (Tybalt) . . . 34
3.2.5 Other Default Logics . . . 35
3.2.6 Summary . . . 36
3.3 The Modal Approaches . . . 38
3.3.1 Nonmonotonic Logic . . . 38
3.3.2 Autoepistemic Logic . . . 40
3.3.3 "All I Know" . . . 43
3.3.4 Other Modal Systems . . . 44
3.4 Hybrid Nonmonotonic Systems . . . 45
3.4.1 Autocircumscription . . . 45
3.4.2 Introspective Circumscription . . . 47
3.4.3 Model Preference Default Theories (Juliet) . . . 48
3.5 Other Nonmonotonic Systems . . . 50
3.5.1 Truth Maintenance Systems . . . 50
3.5.2 Negation As Failure and Closed World Assumptions . . . 52
3.5.3 Poole's Maxiconsistent Sets . . . 56
3.5.4 Inheritance Networks with Exceptions . . . 57
3.6 Summary . . . 58

4 Relating Principles: Axiomatic Properties 59
4.1 Meta-Rules and Restrictions on Inference Relations . . . 60
4.2 Nonmonotonic Systems as Inference Relations . . . 65
4.2.1 Circumscription . . . 65
4.2.2 Default Logic . . . 66
4.2.3 Autoepistemic Logic . . . 68
4.2.4 Summary . . . 69

5 Relating Principles: Nonmonotonic Reasoning Semantics 70
5.1 Preferential Semantics 70
5.1.1 Cumulative Models 71
5.1.2 Preferential Models 72
5.2 Conditional Logic and Plausibility Measures 74
5.2.1 Preferential Entailment from a Conditional Logic Perspective 75
5.2.2 Parameterized Probabilistic Distribution and ϵ-Semantics 75
5.2.3 Possibility Measures 76
5.2.4 κ-Rankings 76
5.2.5 Plausibility Measures 77
5.2.6 Conditional Logic and Nonmonotonic Inference Relations 79
5.2.7 Summary 80
5.3 Semantics for the different nonmonotonic systems 80
5.3.1 Circumscription 81
5.3.2 Pointwise Circumscription 81
5.3.3 Closed World Assumptions 82
5.3.4 Default Logic 82
5.4 Nonmonotonic Modal Logics 84

6 Relating Principles: Translations 85
6.1 Governing Principles 86
6.2 Default Logic vs Autoepistemic Logic 86
6.2.1 Default Logic vs Strongly Grounded AEL 87
6.2.2 Default Logic vs Nonmonotonic Modal Logics 87
6.2.3 Default Logic vs Standard Autoepistemic Logic 88
6.3 Circumscription vs Default Logic 89
6.4 Autoepistemic Logic vs Circumscription 90
6.5 Relationships between other systems 91
6.5.1 Model-Preference Defaults vs Default Logic 91
6.5.2 Circumscription vs Closed World Assumptions 91
6.6 Other Comparison Works 92

7 Other Issues 92
7.1 Computational Aspects 92
7.1.1 Computing Circumscription 92
7.1.2 Reducing Second Order Logic to FOL: SCAN 93
7.1.3 Computing Default Logic 93
7.2 Closely Related Areas 94
7.2.1 Logic Programs 94
7.2.2 Theories of Action 96
7.2.3 Nonmonotonic Reasoning and Belief Revision 97
7.3 Applications 97
7.4 Dictionary 98

8 Conclusions 99

1 Introduction

1.1 A Brief History of Time

From the early days of the field of Artificial Intelligence, one of the dominant views of the path to solving the "AI problem" was that a solution will be found by *Representing* the *Knowledge* of the machine explicitly and *Reasoning* about it [McC58]. The 1960's were full of successes for the KR&R (Knowledge Representation and Reasoning) approaches with the invention of the *resolution* principle [Rob65], QA3, the first *Planner* [Gre69], and the origination of Expert Systems [BF78]. A major part of the KR&R effort was devoted to the paradigm that relied on a First-Order-Logic *theorem prover* to produce the "intelligent" behavior.

In the 1970's the excitement declined after several obstacles were discovered. The complexity of reasoning with a full First-Order-Logic (FOL) theorem prover (First Order Logic is expressively equivalent to a Turing Machine) was discovered to be too large (a post-mortem analysis and the paradigm of restricting expressivity are introduced in [LB85]), the brittleness of expert systems was encountered [BS88],[Fei92], the Frame Problem [MH69] was discovered to be a central issue in knowledge representation and common-sense reasoning, and the inability of FOL to "jump to conclusions" or "reason by default" was argued [Min75]. We shall now continue with the developments following this last problem.

It was in 1978 that several pioneers started working on some ideas for extending the then-current logical paradigms. Early works of Doyle, Reiter and McCarthy were unveiled on that year, and in 1980 the mature papers emerged [MD80],[Rei80],[McC80]. The underlying idea was to have a theorem prover take care of the interpretation of knowledge into action [Win85] (in some works, one can see that, in the extremes of the KR&R endeavor, this "knowledge" is supposed to capture everything, from control of reasoning [Kow79], [BDKT97] to reasoning about beliefs [Hin62], [McC79], to the revision of knowledge [Gär78], [GM94a]).

1.2 The Field of Nonmonotonic Reasoning

Nonmonotonic Reasoning systems were proposed to allow the theorem-proving paradigm to offer common-sense reasoning. In the span of the last 20 years (1978-1998) much research was done on both producing systems that give the "right" commonsense reasoning and on understanding the underlying forces that should be considered in building a nonmonotonic system.

The motivations for nonmonotonic reasoning vary from formalizing Common Sense reasoning through Elaboration Tolerance and representation of uncertainty to Belief Revision. Since its debut, the nonmonotonic reasoning line of work has expanded and several textbooks now exist that give a fair view of nonmonotonic reasoning and its uses (see [Gin87a], [Luk90], [Bre91b], [Gef92a], [DG93b], [MT93a], [Ant97], [BDK97], [Sch97] for nonmonotonic reasoning, and

[Sho88], [McC90b], [San94], [Sha97], [Rei98] for its uses). As far as papers summarizing the effort, [Rei87] is a dated but still valuable survey, and [Gin87b] and the entire book [Gin87a] summarize the nonmonotonic reasoning situation circa 1987.

In this paper, we survey the different major systems that were proposed over the years, the related efforts for using them, the various semantic accounts proposed for these systems, and the generalization and unifying themes of these systems leading to a formal theory of non-monotonic reasoning. Only some of the works are directly related to the nonmonotonic systems that were originally presented. Other works try to make theoretical claims about groups of nonmonotonic systems. These works present nonmonotonic semantics as explanation tools in representing nonmonotonic systems, trying to unify the different systems under one representation schema. Other works try to translate between the different systems, and yet others try to determine guidelines for preferring one nonmonotonic system over another.

The structure of this article reflects this division of works. Most books presenting nonmonotonic reasoning do this by collecting the works done on the specific nonmonotonic system, one system at a time. In this paper we divert from that approach and present nonmonotonic reasoning breadth-wise, all of the systems at once, dividing the discussion of each system according to the topic under consideration. These topics include the presentation of the major nonmonotonic systems, the guiding rules, the semantics and the simulation of one system by another.

In defining the scope for this article, one may argue that semantic nets and some forms of learning exhibited non-monotonic reasoning. We do not consider those here, as we are mostly interested in knowledge-based methods for non-monotonic reasoning, and the investigation of these other forms of nonmonotonic conclusions focused mostly on other aspects than common-sense reasoning.

For lack of space (and time) we also minimized many considerations that are relevant to such a survey paper: Logic Programs ([Llo87],[DG93d]), Belief Revision ([Gär92], [DG93c]), Common-Sense Reasoning ([McC90b], [Sha97], [San94]) and even algorithmic and computational complexity issues ([CS93], [Cad95]) and learning issues related to nonmonotonic reasoning. Nevertheless, we give a glimpse on these topics in section 7.

With respect to the material presented here, we do give only some of the major results obtained for the systems and the semantics presented, as these are too many to count. For a logical and systematic account of the major results obtained for the major nonmonotonic systems and their underlying semantics the reader is referred to [DG93b] and [Ant97].

2 An Introduction to Logic

This section a short reminder of logic and some of its less-known corridors that will be useful in our expedition into Nonmonotonic Reasoning. For further

details and tutorials on the topics discussed here, one should consider reading [Sho67], [DG93a] and [FHMV95].

2.1 First Order Logic (FOL): Syntax and Semantics

First order logic is comprised of a language, a proof theory and semantics. While there can be many proof theories and semantics to a logic, when plainly referring to *the semantics* we refer to the Tarski Semantics, and when plainly referring to *the proof theory* or *the calculus* we refer to the Frege-Hilbert calculus. Both the syntax, the semantics and the proof theory will be detailed below.

2.1.1 The language of FOL

The language of first order logic has a set of object constant symbols, a set of relation constant symbols, a set of constant function symbols, and a set of connectives and parentheses as the building operators. We shall describe the language in the infix form, but there are equivalent prefix and postfix forms. The set of constants and their characteristics is called the *signature* of the specific instance of the language, and it is sometimes written like this:

$$\langle U; F_1, ..., F_n; R_1, ..., R_m; O_1, ..., O_k \rangle$$

Depending on our choice, we may include equality in the language or not, but most of the time it will not be mentioned in the signature of the language but rather in the specification of the language (i.e., "first order logic with equality").

Atomic Terms These are either object constant symbols or variable symbols.

Terms Any atomic term or the application of a function to a set of terms (depending on the arity of the function).

Atomic Formulas (Literals) Application of relation symbols to terms (given the arity of the relation).

Formulas Atomic Formulas or the application of connectives (including quantifiers) to formulas.

Ground Formulas Formulas in which there is no variable mentioned.

Open Formulas A formula with all variables free (i.e. not quantified).

Closed Formulas (Sentences) Formulas with all variable instances quantified.

After building the formulas of the language, we start building higher constructs, called *theories*. For now we will just note that a theory is a set of formulas[1]. We will detail the properties of theories in section 2.1.3 on the proof theory.

[1] Some texts refer to *theories* as deductively closed sets of formulas. We prefer this notation here.

2.1.2 The Semantics of FOL

In general, a semantic interpretation of a logical theory is a structure on which the constant symbols used in the theory are mapped. In first order logic this *structure* (also called *interpretation*) is comprised of a set of elements called the *universe* of the structure, and a mapping from object constant symbols to elements of the universe, from relation constant symbols to relations over the universe (depending on the arity of the relation symbol) and from function constant symbols to functions over the universe.

Given a structure $\mathcal{I}$, every formula is given a truth value that is calculated using the mapping of the constants of the language to the universe. We write $\mathcal{I} \models \varphi$ for a formula φ if it is given a truth value of T. Otherwise we write $\mathcal{I} \not\models \varphi$. The definition of $\models$ (which is a relation between sets of formulas and a formula) is recursively defined as follows:

Atomic Term $A^{\mathcal{I}}$ is the defined mapping. For a variable, it will just vary over the set of elements in the universe.

A Term $(f(t_1, ..., t_n))^{\mathcal{I}} = f^{\mathcal{I}}(t_1^{\mathcal{I}}, ..., t_n^{\mathcal{I}})$.

An atomic formula $\mathcal{I} \models R(t_1, ..., t_n)$ iff for every variable assignment by objects from $\mathcal{I}$ $\sigma = [x_1/a_1, ..., x_n/a_n]$ (where $x_1, ..., x_n$ are the only variables free in $t_1, ..., t_n$), $R^{\mathcal{I}}(t_1^{\mathcal{I}}, ..., t_n^{\mathcal{I}})\sigma$.

The connective "$\wedge$" $\mathcal{I} \models \varphi \wedge \psi \iff \mathcal{I} \models \varphi$ and $\mathcal{I} \models \psi$.

The connective "$\neg$" $\mathcal{I} \models \neg\varphi$ iff for every assignment by objects from $\mathcal{I}$ $\sigma = [x_1/a_1, ..., x_n/a_n]$ (where $x_1, ..., x_n$ are the only variables free in φ), $\mathcal{I} \not\models \varphi$.

The quantifier "$\exists$" $\mathcal{I} \models \exists x.\varphi$ iff there is an object a in the universe such that $\mathcal{I} \models \varphi_{[x/a]}$.

The symbols $\vee, \Rightarrow, \iff, \forall$ are defined in the usual way from the above connectives. The symbol $\Rightarrow$ is sometime called *material implication* as it has no special meaning besides being defined as $p \Rightarrow q \stackrel{def}{=} \neg p \vee q$. Other connectives exists in extended languages (see section 5.2 on page 74).

We say that $\mathcal{M}$ is a *model* of a theory Γ if $\forall \varphi \in \Gamma\ \mathcal{M} \models \varphi$, and we write it $\Gamma \models \mathcal{M}$ (and also $\mathcal{M} \models \Gamma$ when Γ is finite).

We say that a theory Γ *logically entails* a formula φ, and write it $\Gamma \models \varphi$ if for every model $\mathcal{M}$ of Γ, $\mathcal{M} \models \varphi$.

A theory T is said to be *complete* if for every sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$. The following is an important result due to Gödel [Göd67].

Theorem 2.1 (Incompleteness of FOL) *For the theory of Natural Numbers $\mathcal{N}$ with the function symbols $+, *$, for every theory $\mathcal{T} \supset \mathcal{N}$, either $\mathcal{T}$ is not complete or it is not recursively enumerable.*

2.1.3 The Proof Theory of FOL

The relation $\vdash$ is a relation between sets of sentences and sentences (as is $\models$). The difference between the relations is that $\vdash$ is concerned with derivations, while $\models$ is concerned with semantics.

$\Delta \vdash \varphi$ if there is a string of formulas such that any formula is either an axiom or derived from previous formulas in the string using a rule. We expand on this in what follows.

The "classical" mathematical logic calculus, invented by Gottlob Frege ([Fre79]) and David Hilbert ([HB34], [HA50a]), is called the Frege-Hilbert calculus. In it, there are a set of axioms and a set of rules. Many such combinations exist, in most of which, the set of rules is much smaller than the set of axioms, but not necessarily[2] A *derivation* is a string of First Order Logic formulas, each formula is either an axiom or is proved using a rule from previous formulas in the string.

This calculus is mostly useful for meta-theorems, as it is very simple and close to the definition of First Order Logic. It is not suitable for automatic deduction, because its forward reasoning has no guiding heuristics, whereas its backward reasoning has a potentially infinite (or very large) branching factor (mostly because of the *modus-ponens* or the *cut* rules). Other calculi include Natural Deduction, Sequent Calculus, Tableaux, the Connection method and Resolution. Details, on these and other calculi (including example axioms-rules systems), can be found in [DG93a].

Theories Sets of formulas. Sometimes they will be referred to as being deductively closed, but we will refer to them here as plain sets of formulas.

A Complete Theory A theory such that for any given sentence, either it or its negation are provable from the theory (if we took the definition of a theory to be a deductively closed set of formulas, we would have, for a given sentence, say that either it or its negation are in the theory).

Consistency A theory Δ is *consistent* if $\Delta \nvdash FALSE$ (i.e., a sentence and its negation).

Theorem 2.2 (Soundness of the Frege-Hilbert proof system) $\Gamma \vdash \varphi \implies \Gamma \models \varphi$, *where* φ *is a sentence.*

Put other words, if there is no proof of FALSE from Γ, then there is a model for Γ.

Theorem 2.3 (Completeness of the Frege-Hilbert proof theory) $\Gamma \models \varphi \implies \Gamma \vdash \varphi$. *where* φ *is a sentence.*

Put other words, if there is a model of Γ, then there is no proof of $FALSE$ from Γ.

Theorem 2.4 (Compactness) $\Gamma \vdash FALSE$ *iff there is a finite* $\Gamma_0 \subset \Gamma$ *such that* $\Gamma_0 \vdash FALSE$.

[2] E.g., the formalization in [Sho67].

2.1.4 The Herbrand Universe

The *Herbrand Universe* of a theory is the set of ground terms that can be produced using the constants mentioned in the theory (if there are no object constant symbols at all, a constant symbol is artificially added).

The *Herbrand Base* of a set of clauses[3] Δ is the set of all ground instances of the clauses in Δ (this set is produced using the Herbrand Universe as the set of instantiators).

A *Herbrand Interpretation* of Δ is then the Interpretation built from the Herbrand Universe as its universe (terms are mapped to themselves) and ground atoms are assigned either true of false.

Mostly, when constructing the Herbrand Interpretation, we rely on having a model that will guide us in giving the right values to the ground atoms. The Herbrand Universe by itself is therefore not useful for proofs of completeness (where the existence of the model is what we try to prove), without using a previous completeness result. It does become useful when we wish to either choose a specific interpretation (e.g., the canonical model in logic and logic programs (see section 7.2.1)) or we have a previous completeness result (e.g., for a different language).

Theorem 2.5 (Herbrand) *If a* finite *set of clauses Δ is unsatisfiable, then the Herbrand base of Δ is unsatisfiable.*

The opposite theorem obviously holds.

2.1.5 Other Theorems of Importance for AI

Jon Doyle's work [Doy85] on implicit definability and Nonmonotonic Reasoning put some theorems under the spot-light as extremely important for logical-AI research in general and Nonmonotonic Reasoning in particular. Here we state some of them.

Theorem 2.6 (Craig's Interpolation Lemma) *If $\alpha \vdash \beta$, then there is a formula γ involving only symbols common to both α and β, such that $\alpha \vdash \gamma$ and $\gamma \vdash \beta$.*

A *implicitly defines* P iff when we pick some symbol Q not in $\mathcal{L}$, we can prove $\forall x(P(x) \equiv Q(x))$ from $A(P) \cup A(Q)$.

Theorem 2.7 (Beth Definability Theorem) *Let A be a theory. A implicitly defines P iff there is a formula ϕ involving only the symbols of A exclusive of P such that*

$$A \vdash \forall x(P(x) \equiv \phi(x))$$

I.e., in First-Order Logic, implicit definability is equivalent to explicit definability. This theorem is not true for Higher-Order logic.

[3]Important that we refer here to clauses, because there we do not have quantifiers.

2.2 High-order Logics

Second Order Logic is built from First Order Logic by adding variables and quantification for functions and relations. Higher order logics are constructed similarly. *Type theory* or *Finite order* or *ω-order* logics are logics that have variables for all finite types.

A *type* is a a structure denoting the type of a relation/function/variable. If ι_i is the type of R_i then $P(R_1, ..., R_n)$ is of type $(\iota_1, ..., \iota_n)$.

2.2.1 Properties of Finite Order Logic

High order logic is useful in cases where we would like to say something about relations and functions (such as minimality of a relation, or transitive closure, or the induction principle). Any finite-order arithmetic is undecidable, but there are some deduction calculi that are sound. An example of such a calculus for Second Order Logic, is the augmentation of the axioms of First Order Logic (Frege-Hilbert) with rules for the quantifier elimination and introduction for second-order quantifications.

Second Order Logic is not compact, does not have the downward Skolem-Lowenheim property (having a countable model whenever the theory is countable with an infinite model), and does not have the upward Skolem-Lowenheim property (having a model of the cardinality of any infinite cardinal, when having a countable one). The reason these last two properties fail for Second Order Logic, is the ability to write the cardinality $\aleph_n$ for every $n \in \mathcal{N}$. This allows us then to specify that the universe is of a specific cardinality.

Finite order logic is reducible to Second Order Logic in the sense that we have a transformation such that for each sentence in finite order logic, there is a transformed sentence in Second Order Logic that is valid iff the the original one is.

Monadic First Order Logic is decidable, and so is monadic Second Order Logic without functions. We discuss the topic of reducing Second Order Logic sentences to first-order in section 7.1.2 on page 93.

2.3 Modal Logics

Modal Logics are logics with modal operators. Generally speaking, the main difference between a modal operator and an operator (function/relation) in non-modal logics (e.g., FOL, Second Order Logic) is that a regular operator looks at the truth value (in the specific interpretation) of the formula it is given, whereas a modal operator looks at the formula, ignoring its truth value.

The modal operators we shall use in this section are $\Box$, $\Diamond$ and, when we discuss *knowledge and belief* (section 2.4 on page 17), K and B. In later sections we shall sometimes use L and M for modal operators of the nonmonotonic modal systems. For further reading on modal logics, the reader is referred to [Fit93] as an extensive overview. Some other references that are possibly easier to digest include [FHMV95], [Min92], [HC84] and [Pop94]. In what follows, we will refer

mostly to *propositional modal logics* and postpone the discussion of first-order modal logics to section 2.3.6.

2.3.1 General Definitions and Theorems

In this section we consider only *normal* modal propositional logics.

Definition 2.8 *A modal propositional logic is called* normal *if it meets the following conditions on the deduction relation* $\vdash$*:*

Prop (Rule: Propositional truths) Includes every tautology

K (Axioms: Kripke) $\Box(P \to Q) \to (\Box P \to \Box Q)$.

MP (Rule: Modus-Ponens) If it includes $X, X \to Y$ *then it also includes* Y.

N (Rule: Necessitation) If it includes X *then it also includes* $\Box X$.

Definition 2.9 (Possible-worlds Semantics) *A* model *is a triple* $\mathcal{M} = \langle \mathcal{G}, \mathfrak{R}, v \rangle$*, where* $\mathcal{G}$ *is the set of worlds in the model,* $\mathfrak{R}$ *is the accessibility relation, and* v *is the valuation of propositional letters.* $\langle \mathcal{G}, \mathfrak{R} \rangle$ *is called a* frame*, and intuitively represents a theory including the restrictions on* $\mathfrak{R}$ *and* $\mathcal{G}$ *(i.e., the restriction that* $\mathcal{G}$ *includes these worlds). We say that* $\mathcal{M}$ *is* based *on the frame* $\langle \mathcal{G}, \mathfrak{R} \rangle$.

We write $\mathcal{M}, s \Vdash \varphi$ to mean φ *is true in the world* s *in the model* $\mathcal{M}$.

Many notions of validity are used here.

Definition 2.10 φ *is* valid in a model $\mathcal{M}$ *iff* $\forall s \in \mathcal{M}\ \mathcal{M}, s \Vdash \varphi$. φ *is* valid in a collection C of models *(said* φ is C-valid*) iff* $\forall \mathcal{M} \in C\ \varphi$ *is valid in* $\mathcal{M}$*. Given a frame* F*, it is treated as the collection of all models based on that frame.*

Definition 2.11 *A set* L *of formulas is a* frame logic *if* L *is the set of formulas valid in some non-empty set* $\mathbf{L}$ *of frames.*

We say that a model $\mathcal{M}*$ is a *generated sub-model* of $\mathcal{M} = \langle \mathcal{G}, \mathbf{R}, v \rangle$ if $\mathcal{M}* = \langle \mathcal{G}*, \mathbf{R}*, v* \rangle$ is a model with $\mathcal{G}* \subset \mathcal{G}$ that is complete with respect to $\mathbf{R}$ (going forward on $\mathbf{R}$ only), and then $\mathbf{R}*, v*$ are the restrictions of $\mathbf{R}$ and v to the set of worlds $\mathcal{G}*$.

Proposition 2.12 *If* $\mathcal{M}* = \langle \mathcal{G}*, \mathbf{R}*, v* \rangle$ *is a* generated sub-model *of* $\mathcal{M}$*, then* $\forall s \in \mathcal{G}*$ *and for every formula* φ*,* $\mathcal{M}, s \Vdash \varphi \iff \mathcal{M}*, s \Vdash \varphi$.

Proposition 2.13 *A formula* φ *is valid in the disjoint union of two frames iff it is valid in each of the frames.*

Note that $\forall x \forall y\ R(x, y)$ is not kept in a disjoint union, and thus we conclude that it can not be characterized by a modal formula.

Definition 2.14 *A function* g *from* $\langle \mathcal{G}_1, \mathbf{R}_1 \rangle$ *onto* $\langle \mathcal{G}_2, \mathbf{R}_2 \rangle$ *is called a* p-morphism *(poly-morphism) if:*

1. $\forall s,t \in \mathcal{G}_1(s\mathbf{R}_1 t \implies f(s)\mathbf{R}_2 f(t)$, *and*

2. $\forall s \in \mathcal{G}_1\ t \in \mathcal{G}_2(f(s)\mathbf{R}_2 t \implies \exists u \in \mathcal{G}_1(s\mathbf{R}_1 u \wedge f(u) = t.$

Proposition 2.15 *If there is a p-morphism from $\langle \mathcal{G}_1, \mathbf{R}_1\rangle$ to $\langle \mathcal{G}_2, \mathbf{R}_2\rangle$, then any formula valid in the first frame is also valid in the second.*

There is an example of a p-morphism between frames that does not preserve the truth of $(\forall x)\neg R(x,x)$, thus there is no modal formula representing that last formula.

The possible worlds semantics was introduced independently by [Kan57] and [Hin62], and was widely spread after a paper by Kripke [Kri63]. Later, Kripke [Kri65] introduced the possible worlds semantics for *intuitionistic logic* (the logic that allows proving $p \vee q$ iff one can prove either p or q).

2.3.2 Semantical consequence

In this section we continue the exploration of *normal* modal propositional logics. Most of the definitions are general, but the theorems are for normal logics only. The following definition distinguishes between validity as a property of a formula with respect to a theory S, validity in the model (which is like "truthness" in the classical First Order Logic), and truthness in the specific world in the model. We relate the different notions of truth and validity using the following definition.

Definition 2.16 *Let $\mathbf{L}$ be a collection of frames, S, U be sets of formulas, and X be a formula. We write $S \models_{\mathcal{L}} U \to X$ to mean: For every frame $\langle \mathcal{G}, \mathbf{R}\rangle \in \mathbf{L}$, for every model $\mathcal{M}$ based on that frame in which all members of S are valid, for every $s \in \mathcal{G}$ for which all member of U are true, $\mathcal{M}, s \Vdash X$.*

In the last definition, S, U are considered sets of temporary assumptions. S is the set of *global* assumption, all members of which must be true in all the worlds in the model (and thus are necessary truths (given the rule of necessitation)), and U is the set of *local* assumptions.

Theorem 2.17 (Monotonicity) *If $S_1 \subset S_2$, $U_1 \subset U_2$, and $\mathbf{L}_2 \subset \mathbf{L}_1$ (we say that $\mathbf{L}_2$ is* stronger *than $\mathbf{L}_1$), then $S_1 \models_{\mathbf{L}_1} U_1 \to X$ implies $S_2 \models_{\mathbf{L}_2} U_2 \to X$*

Since we have two kinds of premises, we will have two kinds of deduction theorems:

Theorem 2.18 (Local Deduction Theorem) *$S \models_{\mathbf{L}} U \cup \{X\} \to Y$ iff $S \models_{\mathbf{L}} U \to (X \to Y)$*

We will write $\Box^0 X = X$ and recursively $\Box^{n+1} X = \Box(\Box^n X)$.

Theorem 2.19 (Global Deduction Theorem) *1. If for some n $S \models_{\mathbf{L}} U \cup \{\Box^0 X, ..., \Box^n X\} \to Y$ then $S \cup \{X\} \models_{\mathbf{L}} U \to Y$.*
2. If $\mathbf{L}$ is a set of frames closed under generated subframes and ultraproducts, and $S \cup \{X\} \models_{\mathbf{L}} U \to Y$ then there is some n such that $S \models_{\mathbf{L}} U \cup \{\Box^0 X, ..., \Box^n X\} \to Y$.

It was noticed that **L** is closed under ultraproducts, if it can be characterized by a set of first-order formulas (by the Los theorem) (e.g., $\forall x R(x,x)$). Also, closure under generated subframes is immediate for classes of frames characterizable by universal formulas of classical logic. Thus most of the cases we will encounter satisfy our precondition above (being closed under generated subframes and ultraproducts).

Theorem 2.20 (Compactness) *Suppose* **L** *is a collection of frames that is closed under ultraproducts. If* $S \models_{\mathbf{L}} U \to X$ *then there are finite sets* $S_0 \subset S$ *and* $U_0 \subset I$ *such that* $S_0 \models_{\mathbf{L}} U_0 \to X$.

2.3.3 Frame Logics

The following formulas are well-behaved, in the sense that they satisfy the preconditions (and conclusions) of the above theorems.

K The class of all frames. This is the weakest normal modal logic.

$$K: \qquad \Box(P \to Q) \to (\Box P \to \Box Q)$$

T The class of all *reflexive* frames. The motivation is of $\Box$ as *knowable* or *necessary*. The logic is due to Feys [Fey37] who called it **T**, and independently to von Wright [vW51] who called it **M**.

$$T: \qquad \Box P \to P$$

Intuitively this axiom says that whatever is known is true (see section 2.4 about the differences between knowledge and belief). Together with K, T characterizes **T**: $S \models_{\mathbf{T}} U \to X \iff S \cup [\![T]\!] \models_{\mathbf{K}} U \to X$.

K4 The class of all *transitive* frames.

$$4: \qquad \Box P \to \Box\Box P$$

Intuitively, $\Box$ refers to *known*, and it is also called *positive introspection* (if I know something, I know that I know it).

S4/KT4 The class of all *reflexive transitive* frames (Lewis and Langford 1959). $S \models_{\mathbf{S4}} U \to X \iff S \cup [\![T]\!] \cup [\![4]\!] \models_{\mathbf{K}} U \to X$.

KB The class of all *symmetric* frames.

$$B: \qquad P \to \Box\Diamond P$$

$S \models_{\mathbf{KB}} U \to X \iff S \cup [\![B]\!] \models_{\mathbf{K}} U \to X$.

B/KBT The class of all *symmetric reflexive* frames. $S \models_{\mathbf{B}} U \to X \iff S \cup [\![B]\!] \cup [\![T]\!] \models_{\mathbf{K}} U \to X$.

S5/KT45/KT4B The class of all *symmetric, reflexive, transitive* frames (Lewis and Langford [LL59]).

$$5: \qquad \neg\Box P \to \Box\neg\Box P$$

This is *negative introspection* (if I do not know something, I know that I do not know it). $S \models_{\mathbf{S5}} U \to X \iff S \cup [\![T]\!] \cup [\![4]\!] \cup [\![B]\!] \models_{\mathbf{K}} U \to X \iff S \cup [\![T]\!] \cup [\![4]\!] \cup [\![5]\!] \models_{\mathbf{K}} U \to X$.

Read alone, 5 is the condition of *euclidity*: $\forall s,t,u\ R(s,t) \wedge R(s,u) \implies R(t,u)$.

K45 The class of all *transitive and euclidean* frames. This logic is also known as *Weak* **S5**. In this logic it is possible that $\Box\alpha \Rightarrow \Diamond\alpha$ is not true, and thus beliefs are not required to be consistent with each other. It was shown that for the logic of this set of frames one does not have to talk about accessibility *relations*. The same fixed set of worlds can be thought of as being accessible from any world.

D The class of all frames having *idealization* (i.e., every possible world has a possible world accessible from it)[4]. Intuitively $\Box$ is *obligatory*.

$$D: \qquad \Box P \to \Diamond P$$

It is true iff $\Diamond TRUE$.

D4 The class of *transitive idealized* frames.

DB The class of *symmetric idealized* frames.

An example of a class of frames that does not satisfy the above theorems is **GL** (Gödel and Löb). It is the class of all *transitive, irreflexive, finite* frames. Intuitively $\Box$ stands for "provable in first-order Peano arithmetic".

$$GL: \qquad \Box(\Box P \to P) \to \Box P$$

The local deduction theorem holds for it, but the global deduction theorem, and the compactness theorem do not. $S \models_{\mathbf{GL}} U \to X \iff S \cup [\![GL]\!] \models_{\mathbf{K}} U \to X$ provided S, U are finite sets.

2.3.4 The Frege-Hilbert Proof System for Normal Modal Logics

Let us examine **K**. It has two rules, and two axioms:

Classical Axioms all tautologies.

Modal Axioms all formulas of the form $\Box(X \to Y) \to (\Box X \to \Box Y)$.

Modus Ponens Rule $\frac{X, X \to Y}{Y}$.

[4] Also called *serialization*.

Necessitation Rule $\frac{\Box X}{X}$.

A derivation is done in the usual way. There are a few derived rules that are generally useful in shortening proofs:

Regularity Rule for $\Box$ $\frac{X \to Y}{\Box X \to \Box Y}$.

Generalized Regularity Rule for $\Box$ $\frac{A_1 \wedge \ldots \wedge A_n \to B}{\Box A_1 \wedge \ldots \wedge A_n \to \Box B}$.

Regularity Rule for $\Diamond$ $\frac{X \to Y}{\Diamond X \to \Diamond Y}$.

The following definition and theorem rely heavily on the deduction theorems proved earlier for **K**.

Definition 2.21 (Syntactic Entailment for K) $S \vdash_{\mathbf{K}} U \to X$ *says that there is a sequence of formulas, ending with* X*, consisting of a* global part*, coming first, and a* local part *coming last. In the global part each term is an axiom of* **K***, a member of* S*, or follows from earlier formulas using one of the rules* MP *or* Necessitation*. In the local part each term is an axiom of* **K***, a member of* U*, or follows from earlier terms by the rule* MP *(but not* Necessitation*).*

Theorem 2.22 (Completeness and soundness for K) $S \models_{\mathbf{K}} U \to X \iff S \vdash_{\mathbf{K}} U \to X$.

For the logic of **GL** there is a completeness theorem for validity but not for consequence.

2.3.5 Transforming Propositional Modal Logics to First Order Logic

This is quite simple for normal logics for which the translation of axioms into First Order Logic (in terms of restrictions on the accessibility relation **R** in the possible-worlds semantics) is possible. We transform every proposition p to a predicate $P(x)$ (x stands for the world in which p is checked), and follow recursively with the standard connectives. The only significant part in this translation is the interpretation of modalities: $\Box\varphi$ goes to $(\forall y)(R(x,y) \to \varphi(y))$ (again, intuitively x comes for the world we examine right now).

Theorem 2.23 φ *is* **K***-valid iff* $\forall x\ \tau(\varphi, x)$ *is valid (where* τ *is the translation mentioned above).*

The translation to First Order Logic from other modal logics, can be carried by stating the restrictions on R as preconditions for the goal formula. For example, for $\vdash_{\mathbf{T}} X$ we write $(\forall x\ R(x,x)) \to \forall x\ \tau(X,x)$.

This translation leaves us with at most 2 different variables at the same scope, which has some significance in limiting our proofs (see Van Benthem [Ben89]).

2.3.6 First Order Modal Logic

The case for modal logic with First-Order formulas is conceptually straightforward if we limit ourselves to closed formulas in the scope of modal operators. If we wish to consider free variables in the scope of the modal operator, we enter a whole new world of philosophical debates.

In the treatment of First-Order modal formulas, and especially formulas that are have variables in the scope of the modal operator but are quantified outside that scope, there are several tradeoffs. Some concepts of interest are *rigid designators* and *non-rigid designators*. The first refers to holding the constants constant for all the possible worlds. The second says that the constants may change (or may not have any designation at all at some worlds). In the latter case, there is a difference between the *de dicto* and *de re* meanings for $\Box\Phi(C)$, because C may mean the constant in the current world, or the constant in the accessible world. One way to deal with that is using λ-notation: $(\lambda x \Box\Phi(x))(C)$ for the first and $(\Box(\lambda x \Phi(x)))(C)$ for the second.

In the rigid designators case, we can require monotonicity (the universes are only getting larger (considered as sets of elements) as we proceed in the accessibility chain), or fixed universes, which will require all the worlds to have the exact same objects. The latter case is suitable for first-order **S5** and **B**.

The following two formulas summarize the two major axioms used in the literature.

Barcan formula $(\forall x \Box\Phi(x)) \Rightarrow (\Box\forall x \Phi(x))$

Inverse Barcan formula $(\Box\forall x \Phi(x)) \Rightarrow (\forall x \Box\Phi(x))$

The first holds for fixed-worlds models, and the second is equivalent to the monotonicity condition. For further investigation of First-Order Modal Logic, the reader is referred to [Fit93].

Of particular interest to our work are the extensions to first-order logic used in the nonmonotonic community. Some of them can be found in section 5.4 on page 84, but at least one example can be given here. In [Lev84a] and [Lev84b], Levesque considers a language that includes an infinite stock of predicate symbols of every arity, an infinite collection of (individual) variables, an existential quantifier and a special two-place equality symbol. In addition, his language includes a countably infinite set of *standard names* which are considered (like the equality symbol) to be logical symbols.

A sentence is called *objective* if it contains no modal operators. A sentence is called *subjective* if all the nonlogical symbols occur within the scope of a modal operator.

Proposition 2.24 ([Lev84a]) *A proof theory for quantified* **K45** *is formed by the following axioms and inference rules:*

- $(n_i = n_i) \wedge (n_i \neq n_j)$, *where* n_i *and* n_j *are any distinct names.*
- *All instances of theorems of first order logic.*

- $\Box\varphi$, *where* φ *is an instance of a theorem of first-order logic.*
- $\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi)$.
- $(\forall x \Box\varphi) \Rightarrow (\Box \forall x \varphi)$ *(Barcan Formula).*
- $\psi \Rightarrow \Box\psi)$, *where* ψ *is subjective*[5].
- *From* φ *and* $\varphi \Rightarrow \psi$ *infer* ψ.
- *From* $\varphi_{[x/n_1]}, ..., \varphi_{[x/n_k]}$, *infer* $\forall x \varphi$, *if the* n_i *range over all names in* φ *and one not in* φ *(the last one comes as an unconstrained constant).*

2.4 Logics of Knowledge and Belief

This section draws much material from Halpern [Hal95] and Fagin, Halpern, Moses and Vardi [FHMV95], and the reader is referred to these sources for further details.

In general, the most simple modal logic **S5** will do for knowledge. **S5** includes the following axioms:

K	Logical Consequence	$K(p \Rightarrow q) \wedge Kp \Rightarrow Kq$
T	Knowledge	$Kp \Rightarrow p$
4	Symmetry (Positive introspection)	$Kp \Rightarrow KKp$
5	Euclidity (Negative introspection)	$\neg Kp \Rightarrow K\neg Kp$
R1	Propositional tautologies	
R2	Necessitation	$\frac{p}{\Box p}$

The notion of *Belief* mostly drops **T**, which is sometimes replaced by **D**. In both cases, theories of knowledge and belief are not confined to these modal logics and many times prefer different ones, but we shall describe the issues involved in Knowledge and Belief using these simple formalisms, as there is no one agreed upon logic that is the *most suitable*[6].

2.4.1 Multiple Agents

To represent multiple agents we use many accessibility relations $R_1, ..., R_n$ on top of the modal logic that we pick. Each of these accessibility relations comes for a different modal operator in $K_1, ..., K_n$. Given a group of agents we then reason about the beliefs of the group.

To say that *everyone in group* G *knows* φ we use the modal operator E_G. To say that φ *is common knowledge in* G (i.e., that everyone in G knows that

[5] This axiom subsumes the usual positive and negative introspection axioms, and also guarantees correct belief about equalities.

[6] The same thing is argued by Halpern in [Hal95].

everyone in G knows that ...) we use the modal operator C_G. The following semantics define these operators (given with respect to the possible worlds semantics)

$$\begin{array}{lll}(M,s) \models K_i\varphi & \iff & \forall s' \in S\ R_i(s,s') \implies (M,s) \models \varphi \\ (M,s) \models E_G\varphi & \iff & \forall i \in G\ (M,s) \models K_i\varphi \\ (M,s) \models C_G\varphi & \iff & \forall k \in \mathbb{N}(M,s) \models E_G^k\varphi\end{array}$$

where $E_G^1\varphi := E_G\varphi$ and $E_G^{k+1}\varphi := E_G E_G^k\varphi$.

Proof theoretically we add the following axioms on top of those of the basic modal logic.

C1.	$E_G\varphi \iff \bigwedge_{i\in G} K_i\varphi$	Everyone knows
C2.	$C_G\varphi \iff E_G(\varphi \wedge C_G\varphi)$	Fixed point for common knowledge
RC1.	$\frac{\varphi \implies E_G(\varphi\wedge\psi)}{\varphi \implies C_G\varphi}$	induction rule

Once we add common knowledge to the language, the complexity rises to exponential time (from co-NP-complete in the single agent **S5**, or the PSPACE-hard in the multiple agent **S5**).

By adding the element of time (by means of situations, and referring to the world in the possible-worlds semantics as a run and a point in the run) we can formalize the coordinated attack problem:

> Two divisions of an army are positioned on the opposite sides of an enemy. If the two generals attack at the same time, they will win, and otherwise lose. They wish to coordinate attacks by means of sending messengers back and forth. The messengers may pass freely, but may also be captured. In the given night everything goes smoothly, and no messenger is caught. How many passes must they perform before the attack is coordinated?

The answer to that is that they will never achieve common knowledge, and thus will not attack.

Theorem 2.25 *In any run of a system that display unbounded message delays (i.e., messages may be delayed without bound on time), it can never be common knowledge that a message has been delivered. As a result, in such a system, the generals never attack.*

There are two major approaches that try to make things more feasible. The first uses notions of common knowledge that require finite rather than infinite sequences of knowledge, thus achieving an approximation of common knowledge rather than common knowledge. The second approach tries to create common knowledge by increasing the granularity of time (i.e., we use a more coarse time model). If in a certain instant both people know something (e.g., if one of them sneezes) then they have common knowledge. If there is uncertainty about the time needed for one of them to get that knowledge, then again there is no common knowledge.

2.4.2 The Logical Omniscience Problem

The problem of logical omniscience is that all modal logics that rely on the possible-worlds semantics give the agent an infinite inference power, or in other words assume that the agent knows anything that is logically inferred from its knowledge.

One approach to deal with the problem is the *syntactic approach* of Konolige [Kon84], where what the agent knows is represented by a set of formulas that need not be closed under logical consequence or follow any logical system. This approach can be taken to include some inference rules, but does not require them to be complete, and thus has the drawback of losing some of the structure and principles for analysis.

Montague-Scott structures give semantics analogous to the syntactic approach, where the agents knowledge is represented by a set of *sets of possible worlds*, i.e., every sentence is the set of possible worlds that satisfy it. Formally, a function $C_i(s)$ stands for the set of sets of worlds that agent i thinks are possible in world s, and the Montague-Scott structure is then $M = (S, \pi, C_1, ..., C_n)$. We then get nice correlations between properties of the set of sets of world and properties of the consequence relation followed by the agent's knowledge. This semantics does not have the major problem of Logical Omniscience, but since it is a semantic approach, it still does not distinguish two logically equivalent formulas, i.e., the following is a consequence rule in the Montague-Scott semantics:

$$\frac{\varphi \iff \psi}{K\varphi \iff K\psi}$$

The approaches above give us ways to *represent* knowledge rather than to *model* it. The following approaches try to keep the flavor of the kripke semantics and still solve the logical omniscience problem.

One approach is to use *conditional logics* (such as intuitionistic logic, relevance logic and four-valued logic) instead of propositional logic as an underlying logic. This approach was pursued by Halpern et al with their NPL (Nonstandard Propositional logic) where a model of NPL does not necessarily satisfy all tautologies (e.g., no $(p \vee \neg p) \Rightarrow q$), and then every world in the kripke-semantics is an NPL model. In this case not all the tautologies are in the logical omniscience of the agent, but some are.

A similar approach is the *impossible-worlds approach*, where there are worlds in the kripke structure where invalid formulas may be true (such as $p \wedge \neg p$). The K modal operator is interpreted in the same way as before (and same as NPL), but we do not give interpretation to the impossible worlds, and they are used only to infer the semantics of K. Variants of that approach are the NPL above, and also Levesque's logic that uses a 4-valued logic (true, false, both, neither) for his impossible worlds approach.

Another approach is the *logic of general awareness* in which the set of formulas the agent is aware of is a subset of the formulas the agent "knows" (i.e., a subset of the formulas that are true in all the accessible worlds). This set of

formulas, that the agent is aware of, is state dependent (i.e., attached to the state the agent is in) and may be calculated by using the agent's algorithm for computing his knowledge. This last approach is called *algorithmic knowledge*.

Lastly, there is also a probabilistic approach that is not (in my opinion) yet mature, but is mentioned in [Hal95].

3 Nonmonotonic Reasoning Systems

In this section we survey the major nonmonotonic systems and some of their variants. Each one of the systems, without exception, allows some parameters to change, and the methodology for setting these parameters is highly dependent on the application at hand and is still under constant research.

The different nonmonotonic reasoning systems have different restrictions to what the actual players are. In Reiter's Default Logic one has a set of facts W (in either propositional or FOL) and a set of defaults D (in a corresponding way). In Circumscription one adds a formula to the original theory (a formula that depends on the original theory and may be in second-order logic). In Nonmonotonic Modal Logics one uses the same underlying modal logic sentences we saw in section 2.3 to specify defaults, but extends the set of results entailed by the chosen logic by a fixed-point definition relying on what is possibly true.

To have some uniformity, before introducing each family of nonmonotonic systems, we will present one possible view of the the system under consideration using the notation $\Gamma \mid\sim_? \varphi$ ("?" comes for the specific nonmonotonic system in use at that point) saying that φ is nonmonotonically concluded from Γ using this system. In section 4.2 we will see that a given system can define more than one entailment relation. Therefore, the description we choose to display before each system should be taken as a demonstration of the use of the system, but not as the only possible use of the system.

3.1 The Circumscription Family (Montague)

The major idea behind the family of circumscription formulas is that nonmonotonicity of reasoning can be achieved by making the reasoning in two steps: first, add some new facts to your original theory (facts that will depend on your original theory) and only then make deductions using these new facts and your original theory. These *new facts* are frequently of a second-order logic nature (for the first-order original theory case).

Nonmonotonicity is achieved here by the differences between the added sentences to two different theories. This family of nonmonotonic operators is of particular interest in that it does not introduce a new logic but rather use classical second-order logic (or first-order in some cases) to account for the nonmonotonic entailment.

One way to define an entailment relation using this family is, $\Gamma \mid\sim_{C(P,Z)} \varphi$ iff $\Gamma \wedge \Psi(\Gamma, P, Z) \models \varphi$, for $\Psi(\Gamma, P, Z)$ a second-order formula (given a first-order

theory Γ) depending on the form of circumscription chosen, and P, Z vectors of predicate or function constant symbols.

3.1.1 McCarthy's Circumscription

McCarthy's Circumscription [McC80] is one of the first and major nonmonotonic reasoning tools.

McCarthy's Circumscription formula [McC80]

$$Circ[A(P,Z);P;Z] \;=\; A(P,Z) \wedge \forall p,z\,[A(p,z) \Rightarrow \neg(p < P)] \qquad (1)$$

says that in the theory A, with parameter relations and function sequences P, Z, P is a minimal element such that $A(P,Z)$ is still consistent, when we are allowed to vary Z in order to allow P to become smaller.

Notice that $p < P$ stands for the conjunction of $\forall\hat{x}.(p(\hat{x}) \Rightarrow P(\hat{x}))$ (p is no larger than P) and $\exists\hat{x}.(\neg p(\hat{x}) \wedge P(\hat{x}))$ (p is strictly smaller than P).

Take for example the following simple theory:

$$T \;\equiv\; block(B_1) \wedge block(B_2)$$

Then, the circumscription of *block* in T, varying nothing, is $Circ[T;block;] = T \wedge \forall p\,[T_{[block/p]} \Rightarrow \neg(p < block)]$ which is equivalent to

$$Circ[T;block;] \equiv \forall x\;(block(x) \Leftrightarrow (x = B_1 \vee x = B_2))$$

By minimizing *block* we concluded that there are no other blocks in the world other than those mentioned in the original theory T.

Historical notes The formula presented above is in fact the result of a progression of papers.

In [McC77] a formula called *domain circumscription* was presented (there called *minimal entailment*) with the intuition that we try to make the domains of our models as small as possible:

$$Axiom(\Phi) \wedge A^{\Phi} \wedge \forall x.\Phi(x)$$

$Axiom(\Phi)$ comes for the conjunction of $\Phi(a)$'s, for every object constant symbol a, and $\forall x.(\Phi(x) \Rightarrow \Phi(f(x)))$, for every function constant symbol f (and similarly for predicate symbols). A^{Φ} is the relativization of A to the domain defined by Φ, written as the axiom that is achieved from A by rewriting "$\forall x$" into "$\forall x.\Phi(x) \Rightarrow$" and "$\exists x$" into "$\exists x.\Phi(x) \wedge$".

As a second step, [McC80] introduced a more intuitive formula, similar to the one we mentioned at the beginning of this section, in which there were no varied constant symbols mentioned:

$$A(\Phi) \wedge [(\Phi \leq P) \Rightarrow (\Phi = P)]$$

Notice that this is in fact a schema (on the formula Φ) while (1) is a second-order formula. Properties of the first-order schema were studied in [PM86]. In

[McC80] it was also shown that domain circumscription is reducible (under a simple translation) to predicate circumscription.

In response to some criticism given by [EMR85] and others, a slightly different formula was given by McCarthy in [McC86], a formula which he calls *formula circumscription*, which he distinguishes from the less-general previous formula which he calls *predicate circumscription* in two ways: the new one is a second order formula, and the minimized formula is an arbitrary wff whereas the original version minimizes a single predicate. For P being a tuple of some predicate symbols, the *formula circumscription* of the wff $E(P,x)$ in $A(P)$ is

$$A(P) \wedge \forall P'.[A(P') \wedge [\forall x.E(P',x) \Rightarrow E(P,x)] \Rightarrow [\forall x.E(P',x) \equiv E(P,x)]]. \quad (2)$$

This formula, although not mentioning the varied predicates, is in fact equivalent to predicate circumscription with varied constant symbols, which is the formula (1) in the opening of this section. Because of this equivalence, we will continue considering only (1) in the rest of this paper.

3.1.2 Parallel and Prioritized Circumscription

Prioritized Circumscription was introduced by Lifschitz and McCarthy in [Lif85b] and [McC86] to be a simple generalization over the Circumscription idea. Instead of minimizing only one specific predicate, we may want to minimize several predicates, allowing for some priorities among them.

The example below describes an inheritance hierarchy in which every bird is a thing and every penguin is a bird. Things normally do not fly, birds normally do and penguins normally do not.

$$\left.\begin{array}{l} thing \wedge \neg Ab_1 \Rightarrow \neg fly \\ bird \wedge \neg Ab_2 \Rightarrow fly \\ penguin \wedge \neg Ab_3 \Rightarrow \neg fly \end{array}\right\} \text{Deliberated properties}$$
$$\left.\begin{array}{l} bird \Rightarrow thing \\ penguin \Rightarrow bird \end{array}\right\} \text{Hierarchy definition} \quad (3)$$

In this theory we would like to say that all the abnormalities Ab_1, Ab_2, Ab_3 are normally FALSE, but that Ab_3 takes priority over Ab_2 and Ab_2 takes priority over Ab_1. This is done by lexicographically saying that, for $ab = \langle ab_1, ab_2, ab_3 \rangle$ and $Ab = \langle Ab_1, Ab_2, Ab_3 \rangle$,

$$\begin{array}{lll} ab \leq Ab \equiv [& (ab_3 \leq Ab_3) \wedge & \\ & (ab_3 = Ab_3 \Rightarrow [& (ab_2 \leq Ab_2) \wedge \\ & & (ab_2 = Ab_2 \Rightarrow [(ab_1 \leq Ab_1)])])] \end{array}$$

The use of this order $ab \leq Ab \wedge ab \neq Ab$ instead of $p < P$ in our circumscription formula (1) then yields the intended solution (i.e., that a penguin does not fly).

A generalization of this idea is the following (here we follow the account given in [Lif93]). Let $\mathcal{P}$ be a tuple of predicate constants that is broken into disjoint sets $\mathcal{P}_1, ..., \mathcal{P}_n$. The predicates in each $\mathcal{P}_i$ have equal priority and are

therefore minimized concurrently. They also all have more priority than those in $\mathcal{P}_j$ for $j > i$. This is formulated by the following order relations:

$$\langle P_1, ..., P_k \rangle \leq \langle Q_1, ..., Q_k \rangle \equiv P_1 \leq Q_1 \wedge ... \wedge P_k \leq Q_k$$

$$\mathcal{P} \preccurlyeq \mathcal{Q} \equiv$$
$$\bigwedge_{i=1}^{n} \left(\bigwedge_{j=1}^{i-1} (\mathcal{P}_j = \mathcal{Q}_j) \Rightarrow (\mathcal{P}_i \leq \mathcal{Q}_i) \right)$$
$$\mathcal{P} \prec \mathcal{Q} \equiv (\mathcal{P} \preccurlyeq \mathcal{Q}) \wedge \neg(\mathcal{P} = \mathcal{Q}) \quad (4)$$

The *prioritized circumscription* $Circ[A; \mathcal{P}_1, ..., \mathcal{P}_n; Z]$ is then

$$A(\mathcal{P}, Z) \wedge \forall pz[A(p, z) \Rightarrow \neg(p \prec \mathcal{P})] \quad (5)$$

Fortunately, there is a simple theorem that relates prioritized circumscription with predicate circumscription.

Theorem 3.1 *The circumscription* $Circ[A; \mathcal{P}_1, ..., \mathcal{P}_n; Z]$ *is equivalent to*

$$\bigwedge_{i=1}^{n} Circ[A; \mathcal{P}_i; \mathcal{P}_{i+1}, ..., \mathcal{P}_n, Z].$$

Finally, the reader should notice that prioritized circumscription does not treat the cases in which we would like to put priorities that are not in the order format mentioned above. For example, arbitrary partial orders are not treated (see also [Ami98a]). A generalization of prioritization was suggested by Grosof [Gro91] to the case of any well-founded strict partial order, yielding a revised version of Prioritized Circumscription.

3.1.3 Pointwise Circumscription (Benvolio)

In various papers (such as [MP84b], [MP84a]) it was argued that the original predicate circumscription was not able to specify what is allowed to change and vary as we minimize some predicates. For example, we may want, in general, to minimize the predicate P, but we would like to leave some things undetermined (e.g., $P(c)$ must be determined by First-Order deduction and not non-monotonicity). Although not directly in response to this criticism, Pointwise Circumscription treats this case and others.

Pointwise Circumscription was first proposed in [Lif86b] and then expanded in [Lif87b] and [Lif89c][7]. In [DL94] it was demonstrated that Pointwise Circumscription (and Circumscription) can be used to formalize various logics of action[8] in the framework of Features and Fluents [San94]. These logics include OCM

[7] This last one also includes an interesting representation of the circumscription policy in the language itself.

[8] Some of the formalizations needed *filtering*.

(Original Chronological Minimization), PCM (Prototypical Chronological Minimization), PCMF (PCM with Filtering), CMON (Chronological Minimization of Occlusion with Nochange Premises) and CMOC (Chronological Minimization of Occlusion and Change).

The general idea behind Pointwise Circumscription is to capture the minimization idea of *Circumscription* while making the circumscription on different subsets of the domain for different stages of the minimization process.

The **basic case for Pointwise Circumscription** is the formula

$$A(P) \wedge \forall x \neg [Px \wedge A(\lambda y(Py \wedge x \neq y))]$$

(where we minimize the subscript P again, varying nothing), intuitively saying that we remove one element at a time. One of the benefits of such an approach is the first-orderedness of the circumscription formula. This property disappears, though, in the most general form, which we use below.

For Pointwise Circumscription, let $A(S_1, ..., S_n)$ be a sentence in which each S_i is a predicate symbol or a function symbol (in particular, it can be a 0-ary function symbol, i.e., an object constant). We want to minimize one of the predicate symbols from this list, S_{i_0} (Thus, S_{i_0} corresponds to P, and the other members of the list correspond to Z in the notation used above). Let us write $EQ_V(P, Q)$ ("P and Q are equal outside V") for

$$EQ_V(P, Q) \stackrel{def}{=} \forall x(\neg Vx \Rightarrow (Px \equiv Qx)) \tag{6}$$

for P, Q predicates or functions, and V a predicate, all with same arity. The *Pointwise circumscription* of S_{i_0} in A with S_i allowed to vary on V_i is, by definition,

$$A(S) \wedge \forall x s \neg [S_{i_0} x \wedge \neg s_{i_0} x \wedge \bigwedge_{i=1}^{n} EQ_{V_i x}(s_i, S_i) \wedge A(s)] \tag{7}$$

Here, S stands for $S_1, ..., S_n$, s is a list $s_1, ..., s_n$ of predicate and function variables corresponding to the predicate and function constants S, and $\lambda x u V_i(x, u)$ $(i = 1, ..., n)$ is a predicate without parameters that does not contain $S_1, ..., S_n$ and whose arity is the arity of S_{i_0} plus the arity of S_i. We denote (7) by $C_{PW}[A; S_{i_0}; S_1/V_1, ..., S_n/V_n]$. Lifschitz names this form of Pointwise Circumscription **Generalized Pointwise Circumscription**.

The Pointwise Circumscription formula (7) has the property that we remove x from S_{i_0} while varying S (recall that S is the list of predicates/functions $S_1, ..., S_n$) in each of the $V_i x$ separately, fixing all the other parts of the domain, at that time.

In other words, the bracketed part of the formula says that, in the process of removing x from S_{i_0}, we allow only some parts of S to change. These parts are determined by $V_i x$ by saying that only where $V_i x$ is TRUE s_i may be different from S_i (recall that $EQ_{V_i x}(s_i, S_i)$ means that s_i is equal to S_i outside $\lambda u V_i(x, u)$). In section 7.2.2 on page 96 we detail an example of the use of Pointwise Circumscription for common-sense reasoning.

Theorem 3.2 ([Lif87b]) *If $Circ[T;P]$, $PW_B[T;P]$ are the circumscription and the basic pointwise circumscription of P in T, respectively, then*

$$\models Circ[T;P] \Rightarrow PW_B[T;P]$$

[Moi88] and [IO96] report on various results in the opposite direction, showing that there are many interesting cases in which Basic Pointwise Circumscription is equivalent to simple circumscription (thus giving a first-order-equivalent form for the circumscription formula in those cases).

Lately, different generalizations of Pointwise Circumscription were devised. [IO96] present **Extended Pointwise Circumscription** which extends the basic case of Pointwise Circumscription to minimizing multiple predicates in finitely many points concurrently, while preserving a first-order formulation. [Ami98b] presented **Point-Sensitive Circumscription** which generalizes both Circumscription and Pointwise Circumscription (the general case) forming a continuum of formulas between them. Amir also shows that there are some cases in which one would rather use a Point-Sensitive Circumscription form that does not correspond to either Circumscription or Pointwise Circumscription.

3.1.4 Other Circumscription Formulae

Different other forms of circumscription formulas were proposed over the years. Some of particular interest are Autocircumscription [Per88] and Introspective Circumscription [Lif89b], both detailed in section 3.4.1.

One of the major drawbacks of the circumscription paradigm is the need to reduce a second-order circumscription formula (in each special case) to a first-order form in order to reason about it (more about that in section 7.1.

Perlis talks about one way to avoid a second order formula by using set theory [Per87]. The difficulty with using Set Theory comes from the fact that the usual Zermelo-Frankel-Choice (ZFC) theory is too difficult to prove theorems with in an automatic way (and even in proof-checkers), due to two axiom schemas that are included in the theory and the theory's incompleteness [Coh63]. This incompleteness of the theory in fact substitutes the incompleteness of the logic (if we use second-order logic instead), leaving us at the same point we started. To get over this obstacle, Perlis proposed an *elementary Set Theory* [Per87] (a good survey of the topic of common-sense set theory is [PA95]) and showed how with the simple axioms he can circumscribe in first-order logic.

Another problem with circumscriptive theories is that they do not preserve consistency (see [Mak93] for a summary), i.e., that there are theories that are consistent but their circumscription is not. One of the most notoriously known examples is the theory of natural numbers without the constant symbol 0. If we try to minimize the domain (or minimize the predicate $Nat(x)$ of natural numbers), we end up with no models at all, since no model is minimal. Different circumscription formulas were proposed to handle that problem:

Closed Nonrecursive Circumscription Mott [Mot87] *adds* to a theory $T(P,Z)$

the axiom schema

$$T(\Phi, \Psi) \wedge (\Phi \leq P \Rightarrow P \leq \Phi)$$

where Φ and Ψ are tuples of first-order formulae with the restriction that each Φ_i has no more free variables than P_i has (*closeness*) and all symbols in (P,Z) do not occur in any of the formulae Φ_i (*nonrecursiveness*). This formula was proved to preserve consistency, but Besnard et al. [BMM89] showed that it results in a significant reduction in the "conclusion power" compared with Circumscription.

Abstract Circumscription Liau and Lin's [LL92] formula intuitively does not distinguish between isomorphic models, and so the problem of no minimal models for a consistent theory is decreased. Essentially, they prove that *each ω-categorical*[9] *theory has abstractly-minimal models.* Their Abstract Parallel Circumscription (they have a version for Domain Circumscription as well) is the formula

$$T(P, Z) \wedge \forall pz.[T(p, z) \wedge p < P \Rightarrow \\ \exists h.(Function(h) \wedge \text{1-1}(h) \wedge Onto(h) \wedge Pred(h, p, z))]$$

($Pred(h, p, z)$ intuitively says that using h does not change the domains of p and z (automorphism)). Abstract Circumscription coincides with Circumscription for well-founded theories.

Last, another criticism of predicate circumscription is that it can not influence the equality relation (one of the ramifications of this is that one has to write Unique Names Axioms by hand). For that purpose and also due to the consistency preservation problem, Rathmann et al. [RWM90] proposed **Equality Circumscription** which uses the notion of a homomorphism (similar to Abstract Circumscription) to create a preference among models instead of a preference among predicates (see section 5.1 for more details on the different semantics).

[Hin88] proposed an alternative notion he calls **Model minimization** which is based on an extension of Tableaux methods. He argues that there are some cases where one will want $\varphi = Isblock(a) \vee Isblock(b)$ to nonmonotonically entail $(\forall x)(x = a \vee x = b)$ rather than $(\forall x)(x = a) \vee (\forall x)(x = b)$ (which is the result of the domain circumscription of φ). Hintikka also shows that Circumscription can not prove that "all humans are descendents of Adam and Eve" from "Adam and Eve are the only humans without parents" (and a suitable theory of parenthood), while his notion can.

3.1.5 Summary

Diagram 1 on the next page describes the development of Circumscription variants we discussed above and the reasons behind them.

There are still problems with a good methodology for using Circumscription. The most prominent problems are

[9] a theory T is κ-categorical if all models of T with cardinality κ are isomorphic.

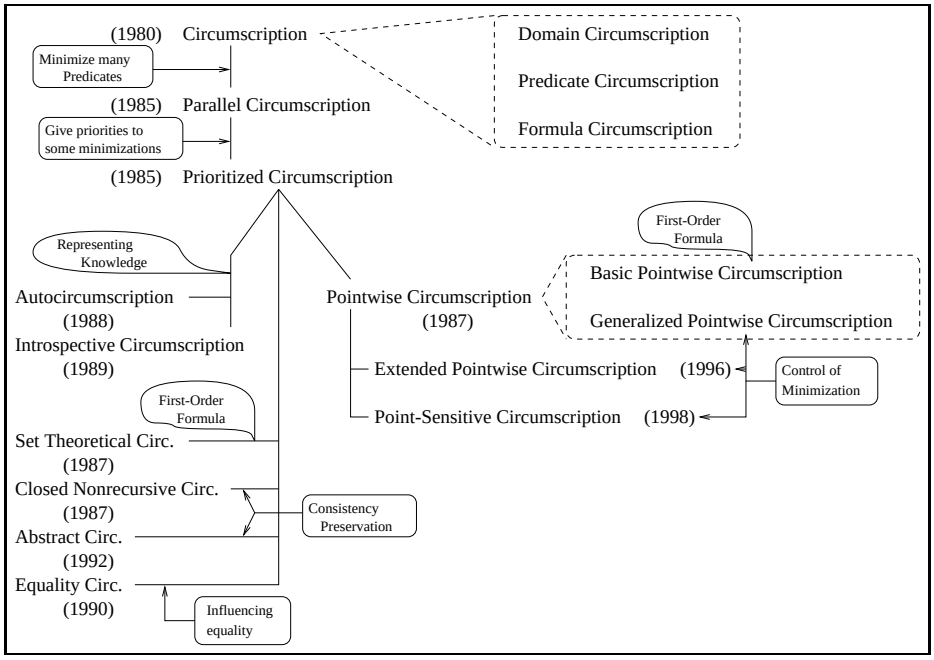

Figure 1: Circumscription Systems

- There is no clear methodology for choosing what to vary and what to minimize, with what priorities.
- There are many circumscription systems, but no clear description of the cases to which a given system is better than others.

3.2 The Default Logic Family (Capulet)

The major idea behind the Default Logic family of nonmonotonic systems is to have an extended set of inference rules on top of the classical ones. This set of "default rules" has a special *syntactic* semantics and is considered part of the set of parameters to be set for the nonmonotonic system.

The family of Default Logic systems has the theory consist of two parts: the base-language theory W and a set of default rules D. Intuitively, then, we can define $W \mid\sim_D \varphi$ according to the appropriate semantics, depending on the specific default logic under consideration.

Our exposition below follows the lines of the definition originally given by Reiter. Antoniou [Ant97] and [Ant96] display some of the logics presented below using *processes*, which are operational characterizations of default extensions and the reader is referred to [Ant97] for an extensive segment on Default Logic and its variants using this presentation.

3.2.1 Reiter's Default Logic

In Reiter's Default Logic [Rei80] one has a set of facts W (in either propositional or FOL) and a set of defaults D (in a corresponding language). Defaults in D are of the form $\frac{\alpha:\beta_1,...,\beta_n}{\gamma}$ with the intuitive idea that if α is proved, and $\beta_1, ..., \beta_n$ are consistent (throughout the proof), then γ is proved. α is usually called the *prerequisite*, $\beta_1, ..., \beta_n$ are the *justifications* and γ is the consequent. Notice that the justifications are checked for consistency one at a time (and not conjoined).

Take, for example, the following default theory $T = \langle W, D \rangle$:

$$D = \{\frac{bird(x) : fly(x)}{fly(x)}\} \quad W = \{bird(Tweety)\} \tag{8}$$

Intuitively, this theory says that birds normally fly and that $Tweety$ is a bird.

Some restricted forms of defaults may be more intuitive than others, such as defaults of the form $\frac{\alpha:\beta}{\beta}$ which basically say that if α is proved, then β is true *by default*. These defaults are called *normal defaults*. Other special kinds of defaults are *semi-normal defaults*, of the form $\frac{\alpha:\beta\wedge\gamma}{\gamma}$, and *open defaults* of the form $\frac{:\beta_1,...,\beta_n}{\gamma}$.

Returning to our general nonmonotonic framework, proving φ from $\langle W, D \rangle$ corresponds to the statement $W \mid\sim_{DL} \varphi$ where φ is a sentence in our language (propositional or FOL), and $W \mid\sim_{DL} \varphi$ corresponds to our *intended meaning*. In Default Logic there are different entailment operators one can take. For *cautious* entailment we define, $W \mid\sim^{c}_{DL} \varphi$ iff every extension (defined below) of $\langle W, D \rangle$ satisfies φ (unless mentioned otherwise, our entailment of choice will be cautious entailment). The *brave* entailment is defined by $W \mid\sim^{b}_{DL} \varphi$ iff at least one extension of $\langle W, D \rangle$ satisfies φ. Thus, entailment in Default Logic is determined by the set of extensions of the default theory.

An *extension* of $\langle W, D \rangle$ is a set of sentences E that satisfies W, follows the defaults in D, and is minimal. More formally, E is an extension if it is minimal (as a set) such that $\Gamma(E) = E$, where we define $\Gamma(S_0)$ to be a minimal set S such that

1. $W \subset S$.
2. For all $\frac{\alpha:\beta_1,...,\beta_n}{\gamma} \in D$ if $\alpha \in S$ and $\forall i \ \neg\beta_i \notin S_0$, then $\gamma \in S$.

An equivalent way to write that rule is the following ([Rei80]): Let $E \subset L$ a set of sentences. Define $E_0 = W$ and

$$E_{i+1} = Th(E_i) \cup \{w \mid \frac{\alpha : \beta_1, ..., \beta_n}{\gamma} \in D \text{ and } \alpha \in E_i \text{ and } \beta_1, ..., \beta_n \in E\}$$

Then E is an extension of $\langle W, D \rangle$ iff $E = \bigcup_{i=0}^{n} E_i$. Take, for example, the default theory

$$D = \{\frac{bird(x) : fly(x)}{fly(x)}\} \quad W = \{bird(Tweety)\}.$$

This theory has exactly one extension, $E = \{bird(Tweety), fly(Tweety)\}$. If we have the more ambitious theory

$$D = \{\frac{bird(x):fly(x)}{fly(x)}, \frac{penguin(x):\neg fly(x)}{\neg fly(x)}\}$$
$$W = \{penguin(Tweety), (\forall x)(penguin(x) \Rightarrow bird(x))\}$$

we get two extensions.

$$E_1 = \{penguin(Tweety), bird(Tweety), fly(Tweety)\}$$
$$E_2 = \{penguin(Tweety), bird(Tweety), \neg fly(Tweety)\}$$

E_1 and E_2 result from the application of the first and second default respectively.

An informal description of a simple way to develop extensions is as follows: We try using the first definition to *construct* extensions. In each *cycle* in the extension buildup one has the choice of which default to apply next. After applying the default, we must make sure that we did not add something to S that violated a previously used default. If that happened, we must backtrack to the last default-choice point. When we have no more defaults to apply (i.e., there is no default that is justified, whose prerequisite was not negatively proved and whose consequent is not in our S), we got an extension. If for every choice of default, applying it causes a later default to contradict the earlier default's justification, then there is no extension.

[Ant97] contains a simple prototype of an implementation of an "extension finder" for a slightly different (but equivalent) definition of an extension.

Many interesting theorems exist in the Default Logic literature, and we present only a few of them here.

Theorem 3.3 *If, for two extensions E and E' of a default theory T, $E \subset E'$ then $E = E'$.*

Theorem 3.4 *A default theory $T = \langle W, D\rangle$ has an inconsistent extension iff W is inconsistent.*

Theorem 3.5 (Unique extension) *Let $T = \langle W, D\rangle$ be a default theory such that the set of formulae $E = W \cup \{\psi_1 \wedge ... \wedge \psi_n \wedge \xi | \frac{\varphi:\psi_1 ... \psi_n}{\xi} \in D\}$ is consistent. Then, T has exactly one extension E.*

One "problem" with Default Logic is that a default theory is not guaranteed to have extensions. We partially deal with this problem by introducing *normal* defaults below. Other attitudes are expressed in the various default logics presented in sections to follow.

Finally, an important thing to notice about the First-Order-Logic case of Default Logic is that a default with *free variable* symbols in fact stands for that set of defaults instantiated with ground terms. This fact has some important consequences with respect to the treatment of equality and the expressivity of the system. First, unlike the case for circumscription, the Unique Names Axioms can be simply stated by the default $\frac{:x \neq y}{x \neq y}$. Second, with respect to representation, one can not prove $\forall x(bird(x) \Rightarrow fly(x))$ from the default $\frac{bird(x):fly(x)}{fly(x)}$ or even from $\frac{:bird(x) \Rightarrow fly(x)}{bird(x) \Rightarrow fly(x)}$.

Normal Defaults Normal defaults are defaults of the form $\frac{\alpha:\beta}{\beta}$. These defaults are interesting both because they are fairly intuitive in nature ("if we proved α then β is proved unless previously proved inconsistent"). We say that a default theory is *normal*, if all of its defaults are normal.

Some of the most important properties of normal default theories are the following:

Theorem 3.6 (Existence of extensions) *Every normal default theory has an extension.*

Theorem 3.7 (Semi-monotonicity) *Let $T = \langle W, D\rangle$ and $T' = \langle W', D'\rangle$ be normal default theories such that $D \subset D'$, and E an extension of T. Then, there is an extension E' such that $E \subset E'$ (remember that an extension is a* set *of sentences).*

Other words, adding defaults always keeps the original extensions, possibly enlarged (by the addition of further conclusions).

Theorem 3.8 *Let E and F be different extension of a normal default theory T. Then $E \cup F$ is inconsistent.*

As a result, the number of extensions never decreases as we add defaults to a normal default theory.

One of the most important results about normal default theories is the existence of a proof theory. A *default proof* of φ in a normal default theory $T = \langle W, D\rangle$ is a finite sequence $\langle D_0, ..., D_n\rangle$ of subsets of D such that

- $W \cup cons(D_0) \models \varphi$ ($cons(D_0)$ is the set of consequents the defaults in D_0).
- $\forall i < n$, the prerequisites of defaults in D_i follow from $W \cup cons(D_{i+1})$.
- $D_n = \emptyset$.
- $W \cup cons(\bigcup_i D_i)$ is consistent.

Theorem 3.9 (Proof Theory) *A formula φ has a default proof in a normal default theory T iff there is an extension E of T such that $\varphi \in E$.*

Notice that this proof theory corresponds to the *brave* entailment operator $\mid\sim^b_{DL}$. The appropriate proof theory for the *cautious* entailment operator $\mid\sim^c_{DL}$ can be defined to be the nonexistence of a proof for $\neg\varphi$. More on that when we talk about computational complexity in section 7.1.3.

Returning to our example (8), we can now prove (cautiously) that $Tweety$ indeed flies. First notice that from theorem 3.5 we can conclude that there is exactly one extension for T. Using the proof theory above, take $D_0 = \{\frac{bird(x):fly(x)}{fly(x)}\}$, $D_1 = \emptyset$ and $\varphi = \{fly(Tweety)\}$. Using the proof theory mentioned above, one can easily check that $\langle D_0, D_1\rangle$ in fact is a default proof of φ.

Semi-normal Defaults Sometimes normal default are not enough. Take for example (adapted from [Poo93]) the following default rule:

$$\frac{crime(c) : committed_crime(x,c) \wedge suspect(x,c)}{suspect(x,c)}$$

This default is not normal and we would not like to write $\frac{crime(c):suspect(x,c)}{suspect(x,c)}$ which implies everyone is a suspect.

Semi-normal defaults are defaults of the form $\frac{\alpha:\xi\wedge\beta}{\beta}$. Since they are generalizations of normal defaults (every normal default is a semi-normal one), we have to impose further conditions on our theory to achieve nice results.

We do not define the notion of an *ordered* default theory and refer the reader to [Ant97] p.60. A default theory is *plain* if it is propositional and the set of defaults is finite. For this case we can prove the following

Theorem 3.10 (Existence of extensions) *A semi-normal, plain, ordered default theory has at least one extension.*

A nice property of default theories with one justification is that they can be translated to a semi-normal default theory in a way that preserves the extensions.

For a default $\delta = \frac{\alpha:\beta}{\gamma}$, let $\kappa(\delta) = \frac{\alpha:\beta\wedge\gamma}{\gamma}$. For a theory $T = \langle W, D\rangle$ let $\kappa(T) = \langle W, \{\kappa(\delta)|\delta \in \mathbb{D}\}\rangle$.

Theorem 3.11 *Let T be a default theory in which all the defaults have one justification. Then, every extension of T is also an extension of $\kappa(T)$.*

Notice that in fact the default $\kappa(\delta)$ is more cautious than δ and so there may be extensions of $\kappa(T)$ that are not extensions of T.

Translating Semi-normal default rules to prerequisite-free defaults, can yield an interesting result:

Theorem 3.12 *Let $\langle W, D\rangle$ be a semi-normal default theory and let $\langle W, D'\rangle$ be the theory where $\frac{\alpha:(\beta\wedge\gamma)}{\beta} \in D$ iff $\frac{:(\alpha\Rightarrow\beta)\wedge\gamma}{\alpha\Rightarrow\beta} \in D'$. If E is an extension of $\langle W, D\rangle$ then there is an extension of $\langle W, D'\rangle$, E' such that $E \subset E'$.*

3.2.2 Justified Defaults

Lukaszewicz presented in [Luk88] and [Luk90] a different fixed-point definition of extensions, allowing for the separation of the extension from its justifications. In this new formalization, we stop the "construction" of the Extension before we add a consequent that contradict earlier justifications. This way, we ensure that there is always an extension.

We define $\Gamma_1(S_0, J_0) = S$ (defining the extension) and $\Gamma_2(S_0, J_0) = J$ (defining the justification) to be the smallest S, J (as sets) such that

1. $S = Th(S)$.
2. $W \subset S$.

3. If $\frac{\alpha:\beta_1,...,\beta_n}{\gamma} \in D$, $\alpha \in S$ and $\forall X \in J \cup \{\beta_1, ..., \beta_n\}$ $S_0 \cup \{\gamma\} \not\vdash \neg X$, then

   (a) $\gamma \in S$.

   (b) $\beta_1, ..., \beta_n \in J$.

E is a *modified extension* of T if there is a set of justifications $F \subset D$ such that $E = \Gamma_1(E, F)$ and $F = \Gamma_2(E, F)$.

Take for example the default theory $T = \langle W, D \rangle$

$$D = \{\frac{:fly(x), \neg fly(x)}{\neg fly(x)}\}$$
$$W = \{\}$$

which says that if we can not decide whether something flies or not, we should assume it does not. We assume that our language includes the object constant $Bill$. In Reiter's default logic we get no extensions because this default contradicts itself. After trying to add $\neg fly(Bill)$ using the default, we discover that the default is no longer applicable, and so there is no fixed point. On the other hand, there is a modified extension, which is the empty one. To see this, notice that the precondition for adding facts and justifications to S and J is never satisfied (because $\neg fly(Bill) \vdash \neg fly(Bill)$).

Theorem 3.13 *Every extension of T is a modified extension of T.*

Theorem 3.14 (Existence of modified extensions) *Every default theory has at least one modified extension.*

For normal default theories the definitions converge.

Theorem 3.15 *Let T be a normal default theory. Then, E is an extension of T iff it is a modified extension T.*

Finally, we relate some default theories to to their semi-normal form (see section 3.2.1 above).

Theorem 3.16 *Let $T = \langle W, D \rangle$ be a default theory such that all default in D have one justification. Then T and $\kappa(T)$ have the same modified extensions.*

[Ant97] gives another (equivalent) characterization of modified extensions that is somewhat easier to apply to concrete examples.

3.2.3 Cumulative Default Logic

Developed by Brewka in [Bre91a], Cumulative Default Logic was built after Brewka noticed that contradicting justifications may be used in constructing extensions (in Reiter's defaults) and modified extensions (in Justified defaults). Brewka argues that this situation is not favorable. For example, the default theory $T = \langle W, D \rangle$ with $T = \emptyset$ and $D = \{\frac{:p}{q}, \frac{:\neg p}{r}\}$ has the extension $Th(\{q, r\})$.

To formalize this approach, Brewka introduced *Assertional Default Theories*. An *assertion* ξ is a tuple $\langle \alpha, \{\gamma_1, ..., \gamma_n\} \rangle$ where $\alpha, \gamma_1, ..., \gamma_n$ are formulas, $\alpha =$

$Form(\xi)$ is the asserted formula and $\{\gamma_1, ..., \gamma_n\} = Supp(\xi)$ is the support of α. For a set of assertions $\mathcal{S}$, he then defines $\hat{Th}(\mathcal{S})$ to be the assertional closure of $\mathcal{S}$, being the minimal set that includes $\mathcal{S}$ and its consequences:

$$\text{If } \xi_1, ..., \xi_n \in \hat{Th}(\mathcal{S}) \text{ and } Form(\xi_1), ..., Form(\xi_n) \vdash \gamma, \text{ then } \langle \gamma, \cup_{i=1}^{n} Supp(\xi_i) \rangle \in \hat{Th}(\mathcal{S})$$

An *assertional default theory* is a pair $\langle D, \mathcal{W} \rangle$, where D is a set of defaults and $\mathcal{W}$ a set of assertions.

Definition 3.17 ([Bre91a]) *Let $\langle D, \mathcal{W} \rangle$ be an assertional default theory. For any set of assertions $\mathcal{S}$, let $\Gamma(\mathcal{S})$ be the smallest set of assertions $\mathcal{S}'$ such that*

1. $\mathcal{W} \subset \mathcal{S}'$,
2. $\hat{Th}(\mathcal{S}') = \mathcal{S}'$,
3. *for any $\frac{\alpha:\beta_1,...,\beta_n}{\gamma} \in D$, if $\langle \alpha, \Gamma \rangle \in \mathcal{S}'$ and $Form(\mathcal{S}) \cup Supp(\mathcal{S}) \cup \{\beta_1, ..., \beta_n, \gamma\} \nvdash \bot$, then $\langle \gamma, \Gamma \cup \{\beta_1, ..., \beta_n, \gamma\} \rangle \in \mathcal{S}'$.*

A set of assertions $\mathbb{E}$ is an assertional extension for $\langle D, \mathcal{W} \rangle$ iff $\Gamma(\mathbb{E}) = \mathbb{E}$.

Example 3.18 (a modified version of [Eth86]) *The assertional default theory*

$$D = \{\frac{: day}{work}, \frac{: night}{sleep}\} \quad \mathcal{W} = \{\langle day \iff \neg night, \emptyset \rangle\}$$

entails (in Reiter's default logic) that we both work and sleep. If we use Brewka's Cumulative DL, we conclude that we either work or sleep (two different assertional extensions).

Another example is given in section 4.2.2. [Sch92] and [DSJ94] give a focused models semantics for Cumulative DL.

Another default logic that accounts for similar considerations is **Rational Default Logic**, which was introduced in [MT93b]. This proposal is somewhat different in application from Cumulative DL, in that it uses default theories instead of assertional default theories and in easing up the requirement that the consequent of a default be consistent with the current justification. Thus, it does not guarantee the existence of extensions (and is not semi-monotonic) but it does commit to default justifications. Another difference between the two logics is that Rational Default Logic is not cumulative (cumulativity is introduced in section 4.1). Two other versions of Default Logic that do satisfy Cumulativity (one of them also satisfies commitment to assumptions) were proposed by Giordano and Martelli [GM94b].

3.2.4 Constrained Default Logic (Tybalt)

Schaub's work on Constrained Default Logic ([Sch92] and [DSJ94]) is in fact a buildup on the Justified Default Logic (section 3.2.2 on page 31) and on *Cumulative Default Logic* (section 3.2.3 on page 32).

In Schaub's Constrained Default Logic, both ideas (maintaining joint consistency of justifications and halting the extension construction short of causing an inconsistency) are used. The following is the definition of a *constrained extension* (a semi-inductive definition can be found in [DSJ94]). .

Definition 3.19 (Constrained Extensions) *Let $\langle W, D\rangle$ be a default theory. For any set of formulas T let $\Gamma(T)$ be the pair of smallest sets of formulas $\langle S', T'\rangle$ such that*

1. *$W \subset S' \subset T'$,*
2. *$S' = Th(S')$ and $T' = Th(T')$,*
3. *for any $\frac{\alpha:\beta}{\gamma} \in D$, if $\alpha \in S'$ and $T \cup \{\beta, \gamma\}$ is consistent, then $\gamma \in S'$ and $\beta, \gamma \in T'$.*

A pair of sets of formulas $\langle E, C\rangle$ is a constrained extension *of $\langle W, D\rangle$ iff $\Gamma(C) = \langle E, C\rangle$.*

Example 3.20 (from [DSJ94]) *The default theory $(\{A\}, \{\frac{:B}{A}, \frac{:D}{\neg B}\})$ has two constrained extensions:*

1. *$(Th(\{A\}), Th(\{A, B\}))$: $Th(\{A\})$ justified by $Th(\{A, B\})$ and*
2. *$(Th(\{A, \neg B\}), Th(\{A, \neg B, D\}))$: $Th(\{A, \neg B\})$ justified by $Th(\{A, \neg B, D\})$.*

Theorem 3.21 (Existence of Extensions) *Every default theory has a constrained extension.*

Theorem 3.22 (Semi-monotonicity) *Let $T = \langle W, D\rangle$ be a normal default theory and D' a set of defaults such that $D \subset D'$, and (E, C) a constrained extension of T. Then, there is a constrained extension (E', C') of $\langle W, D'\rangle$ such that $E \subset E'$ and $C \subset C'$.*

In normal default theories, every extension is a constrained extension (and a restriction on the other direction).

Theorem 3.23 *Let T be a normal default theory and E a set of sentences. Then, E is a standard extension of T iff (E, E) is a constrained extension of T.*

The notion of the set of *generating defaults* is important and useful in relating Default Logic (DL) and Constrained DL (ConDL). For a standard extension E,

$$C_E = \left\{ \beta \;\middle|\; \frac{\alpha : \beta}{\gamma} \in D,\ \alpha \in E,\ \neg\beta \notin E \right\}$$

The next two theorems are for general default theories.

Theorem 3.24 *Let T be a default theory. If (E, C) is a constrained extension of T and E is a standard extension of T then $C \subset Th(E \cup C_E)$.*

Theorem 3.25 *Let T be a default theory and let E be a standard extension of T. If $E \cup C_E$ is consistent, then $(E, Th(E \cup C_E))$ is a constrained extension of T.*

Last, we relate Constrained DL with Cumulative DL.

Theorem 3.26 *Let $< W, D\rangle$ be a default theory and $\langle \mathcal{W}, D\rangle$ the assertional default theory, where $\mathcal{W} = \{\langle \alpha, \emptyset\rangle | \alpha \in \mathcal{W}\}$. Then, if $\langle E, C\rangle$ is a constrained extension of $\langle D, W\rangle$, then there is an assertional extension $\mathbb{E}$ of $\langle D, \mathcal{W}\rangle$ such that $E = Form(\mathbb{E})$ and $C = Th(Form(\mathbb{E}) \cup Supp(\mathbb{E}))$; and, conversely, if $\mathbb{E}$ is an assertional extension of $\langle D, \mathcal{W}\rangle$, then $\langle Form(\mathbb{E}), Th(Form(\mathbb{E})cupSupp(\mathbb{E}))\rangle$ is a constrained extension of $\langle D, W\rangle$.*

[DSJ94] gives semantics to ConDL.

3.2.5 Other Default Logics

One issue, that the Default Logics described above did not address, is the priorities between defaults. For Circumscription, this topic turned out to have many useful outcomes (see section 3.1.2 on page 22), and in fact, throughout the years there have been some interest in enabling priorities in Default Logic as well. One easy way of representing defaults (see [RC81] and [ER83]) is to replace normal defaults by some appropriate semi-normal defaults. One takes the prerequisite of the preferred default and makes it a weak exception of the other default. For example, in the default theory

$$D = \{\frac{Student : \neg Married}{\neg Married}, \frac{Adult : Married}{Married}\} \quad W = \{Student, Adult\}$$

we would like to prefer the first default over the second. We can then achieve that by replacing the second default with

$$\frac{Adult : Married \wedge \neg Student}{Married}$$

Other approaches include [Bre94] PDL (Prioritized Default Logic) and PRDL (Prioritized Reasoning in DL)) and [HB93]. For example, in PRDL one can write the following default theory:

$$D = \{d_1 = \frac{bird : flies}{flies}, d_2 = \frac{penguin : \neg flies}{\neg flies}\} \quad W = \{d_2 \prec d_1, bird, penguin\}$$

where $d_2 \prec d_1$ means that the default d_2 takes priority over d_1 and so we should conclude $\neg flies$.

[DS94], [DS97b], and [DS97a] describe **ordered default theories** which consist of default rules, world knowledge and **preference** on the default rules.

[DS97b] shows that these theories are translatable into default theories via naming of defaults, so that reference may be made to a default rule within a theory.

One of the problems with Default Logic is its inability to handle disjunctions properly: if $a \vee b$ is proved (but neither a nor b is proved) and we have the default rules $\frac{a:\beta}{\gamma}, \frac{b:\beta}{\gamma}$, then we can not prove γ (see also section 5 on page 70 on the *OR* rule). [GLPT91] proposed **Disjunctive Defaults** of the form

$$\frac{\alpha : \beta_1, ..., \beta_n}{\gamma_1 | ... | \gamma_m}$$

such that, when a Disjunctive Default is applied, some of the consequents $\gamma_1, ..., \gamma_m$ are inferred. This, together with avoiding *Facts* (facts are represented by prerequisite- and justification-free defaults), solves the disjunction problem in some cases[10]. Some other methods that enable reasoning by cases were devised by Besnard et al [BQQ83], Delgrande and Jackson [DJ91], Konolige [Kon88b], Moinard [Moi94] and Roos [Roo98]. Notice that the problem with disjunctions arises in all other default theories mentioned above.

Other default logics that we do not consider here are [DL92] (**NML3**), [GM94b] and [Ryc91].

Relating Default Logics With respect to default systems that generalize over those that were presented above, we would only mention Teng [Ten97] who presented Possible World Partition Sequences that allowed him to represent Reiter's DL, Justified DL, Disjunctive DL and Cumulative DL.

3.2.6 Summary

The major properties of the Default Logics presented above are summarized in table 2.

	DL	Con. DL	Jus. DL	Cum. DL
Existence	N	G	G	G
Semi-monotonicity	N	G	G	G
Orthogonality	N	N	N	N
Weak-Orthogonality	-	G	N	G
Commitment	N	G	N	G
Reasoning about defaults	PfN	Pf	?	?

Figure 2: Properties of the different Default Logic systems.

[10] One can still have disjunctions, and so the problem is not entirely solved, but a theory can be modeled so that the problem does not arise.

The Legend for this table is as follows:

G	General default theories.
N	Normal default theories.
-	not relevant.
Pf	Prerequisite free defaults.
PfN	Prerequisite free normal defaults.
Existence	Existence of extensions is guaranteed for consistent W.
Semi-monotonicity	Adding defaults only adds extensions (that are possibly larger).
Orthogonality	Consistent extensions are identical.
Weak-Orthogonality	Distinct extensions have inconsistent justifications.

Diagram 3 describes the development of the Default Logic variants we discussed above and the reasons behind them.

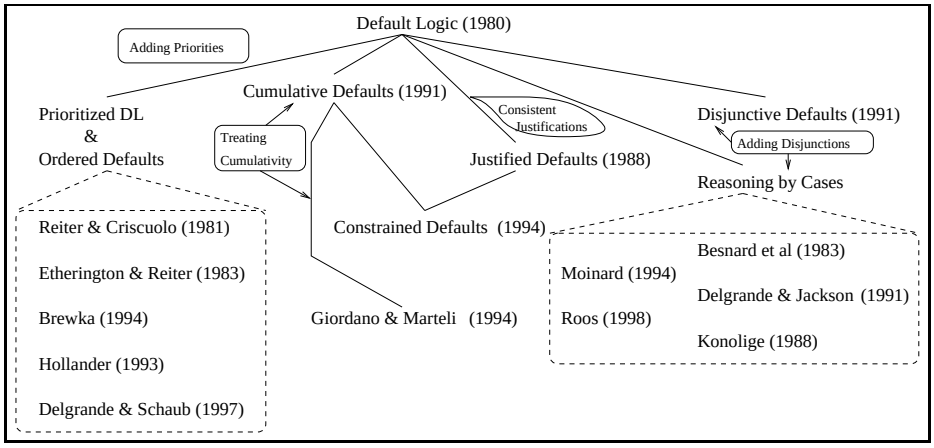

Figure 3: Default Logic Systems

There are still problems with a good methodology for using Default Logic. The most prominent problems are

- One can not reason about defaults with prerequisites.
- There is no methodology for preferring between the various seemingly-equivalent representations for defaults.
- The nature of the Default Logic family implies that every default theory is a brand new logic, with different inference rules, thus making studying Default Logic difficult.
- Inference in the First-Order case requires us to reason using ground terms, making it incompatible with regular First-Order inference and non-intuitive.

3.3 The Modal Approaches

The approaches we discuss in this section use modal operators to model the elusive "jumping to conclusions". The different nonmonotonic systems that are based on modal logics, use modalities with different intuitions. Some refer to the modality as characterizing consistency, others use it to characterize knowledge and yet some others use more than one modality operator. For the basics of modal logics, the reader is referred to section 2.3.

The well known modal logics (e.g., all normal modal logics described in section 2.3) are monotonic. That is, adding an axiom does not retract any of the previously valid conclusions. Nevertheless, modal operators allow us to distinguish different kinds of *truths* and that feature proves very useful in defining extended versions of modal logics that do not obey monotonicity.

In this section we will put an emphasis on three main systems: Nonmonotonic Logic [MD80], Autoepistemic Logic [Moo87] and the logic of the operator "All I Know" [Lev90]. Since the result of Schwarz [Shv90], we know that Autoepistemic Logic is in fact Nonmonotonic Logic with the underlying entailment of modal logic **K45**. Nevertheless, Autoepistemic Logic takes a central place in the literature, and we dedicate a special section for it.

In the following, we will use two modal operators L and M, where $M \stackrel{def}{=} \neg L \neg$. L and M will change meanings, depending on the logic we present in each section, but typically they will mean "Believe" and "Possible", respectively.

3.3.1 Nonmonotonic Logic

McDermott & Doyle's *nonmonotonic logic* (NML) [MD80], [McD82] was the first modal nonmonotonic system. The intuitive interpretation of "assuming A by default" is that "if A can be consistently assumed, assume it."

McDermott and Doyle [MD80] use a modal operator M with an underlying first-order language. $M\varphi$ is intended to mean "φ is consistent". To say that all birds fly by default, we write

$$\forall x.\, Bird(x) \wedge MFly(x) \Rightarrow Fly(x)$$

Since the "consistency" we ascribe to $Fly(x)$ depends on the rest of the theory, there was some difficulty in making this notion precise and intuitive. [MD80] defines the nonmonotonic extension T of a modal theory A by a fixed point.

Definition 3.27 (Expansions) *T is a fixed point of A just in case T is the set of tautological consequences of $A \cup \{\neg LP | P \notin T\}$. Every such a fixed point is called an* expansion.

Given a theory $\Sigma \subset \mathcal{L}$ and a formula $\varphi \in \mathcal{L}$, $\Sigma \models_{NM} \varphi$ iff φ belongs to all the expansions of Σ.

A more modern description of the fixed point is

$$T = Th_{FOL}(A \cup \{Mw | \neg w \notin T\}) \tag{9}$$

At the time, Nonmonotonic Logic did not get clear semantics and it was shown inadequate in capturing the intended meaning of the modal operator M (e.g., $\{M\varphi, \neg\varphi\}$ is consistent). McDermott's attempt [McD82] at correcting NML by replacing FOL (in the semantics for the closure operator $Th(...)$) by the modal logics **T**,**S4** and **S5** turned out to be either too weak to adequately characterize the M operator as intended[11] (in the case of T and $S4$), or to reduce the logic to the monotonic $S5$ (in the case of $S5$).

Due to these problems and the emergence of Autoepistemic Logic (next section), research on NML was mostly abandoned in the mid 80's. Non-the-less, at the turn of the decade, Schwarz [Shv90] showed that if one uses the modal logic **K45** as the underlying logic, one gets exactly Autoepistemic Logic (which turned out to be adequate for the job, if not realizing the intended meaning for the modal operator M).

Marek et al. [MST91],[MT90] and Truszczyński [Tru91b], [Tru91a] subsequently made an extensive investigation of propositional modal nonmonotonic logics using different modal logics composed from **K**,**D**,**T**,**4**,**5** (see section 2.3) and the three axiom schemata

$$\begin{array}{ll} (WK) & L(\varphi \Rightarrow \psi) \wedge L\varphi \Rightarrow ML\psi \\ (W5) & ML\varphi \Rightarrow (\varphi \Rightarrow L\varphi) \\ (F) & \varphi \wedge ML\psi \Rightarrow L(M\varphi \vee \psi) \end{array}$$

(WK) is a weakened form of **K** and (W5) and (F) are weakened forms of **5**.

One modal system S is properly included in another system S' if all of the theorems of S are theorems of S', but not vice versa. Diagram 4 on the next page gives the inclusion relations of some relevant (monotonic) modal systems. This figure is a slightly modified version of the one appearing in [Kon93].

The system **N** contains no schemata (only the rule of necessitation and base-language tautologies) and the system T^- contains only the axiom schema **T** (but not the rule of necessitation). Recall that any system not containing **K** is nonnormal and does not have a possible worlds semantics.

The corresponding inclusion for modal nonmonotonic logics uses the expansions defined by the logic.

Definition 3.28 ([McD82]) *Let $S(A)$ and $S'(A)$ be two nonmonotonic logics generating extensions of a premise set A. S is included in S' if for every set A, every extension of $S(A)$ is also an extension of $S'(A)$. S is properly included in S' if it is included in S' but S' is not included in S.*

McDermott [McD82] showed that the inclusion relation is preserved when moving from modal logics to the corresponding nonmonotonic modal logics, for modal logics included in **S5** (recall that nonmonotonic **S5** is equivalent to **S5**).

[11] Interestingly enough, these exact systems (NML with T and $S4$ as the background entailment relation, to be precise) were later proved adequate to represent Default Logic (see [Tru91b] and section 6.2.2 on page 87), which proves that the abandonment of NML was a little premature.

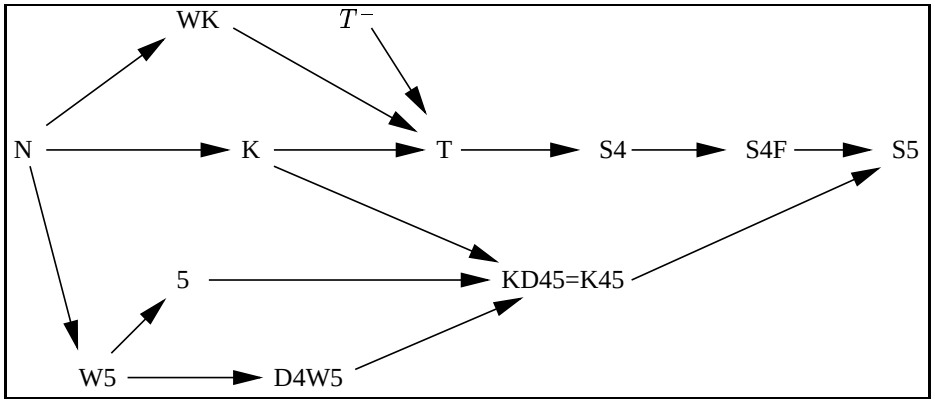

Figure 4: The proper inclusion relation between modal systems

Proposition 3.29 ([**MST91**]) *Diagram 5 on the following page gives the proper inclusion relations among the modal nonmonotonic logics based on the systems of figure 4 on the next page.*

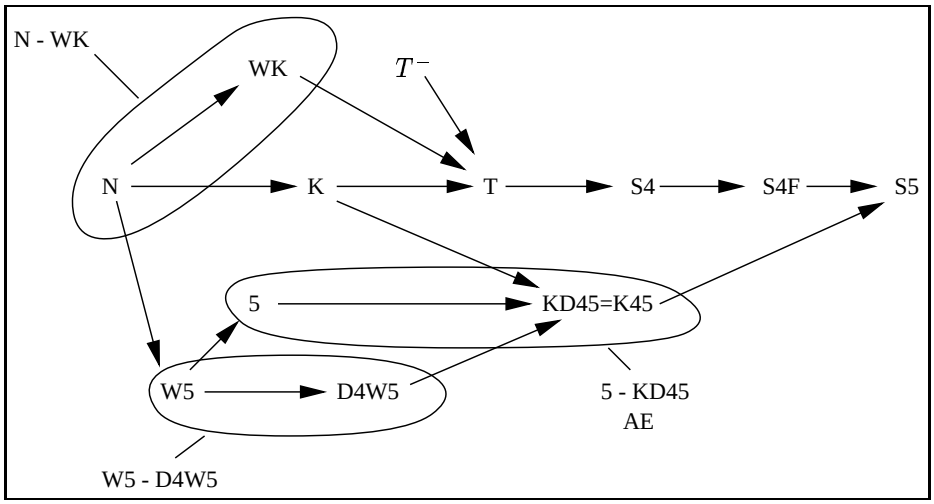

Figure 5: The proper inclusion relation between nonmonotonic modal systems

Each consistent nonmodal theory has exactly one extension for any system between N and KD45 or S4F.

A recent book [Sch97] and a somewhat more dated book [MT93a] give a more detailed treatment to Nonmonotonic Logic.

3.3.2 Autoepistemic Logic

Moore's Autoepistemic Logic (AEL) [Moo87] grew out of Nonmonotonic Logic as an attempt to present a coherent nonmonotonic framework based on Modal Logic. The result, later known as Nonmonotonic K45 (a result due to [Shv90]), is defined using a fixed-point, similar to Nonmonotonic Logic (and also similar to the family of Default Logic), but with a small difference.

Unlike NML which regarded $M\varphi$ as saying "φ is consistent", Autoepistemic logic tries to present a nonmonotonic logic of one's beliefs about his own beliefs (thus "autoepistemic"). The language we use is the propositional logic with the modal operator L for *belief*. In this context, M means *Possible*.

Definition 3.30 (Stable Expansions) *T is a* stable expansion *of A just in case T is the set of tautological consequences of $A \cup \{LP | P \in T\} \cup \{\neg LP | P \notin T\}$.*

Given a theory $\Sigma \subset \mathcal{L}$ and a formula $\varphi \in \mathcal{L}$, $\Sigma \models_{AE} \varphi$ iff φ belongs to all the stable expansions of Σ.

Comparing this definition to definition 3.27 on page 38 of Nonmonotonic Logic, we see that Moore added the part $\{LP | P \in T\}$ to the fixed point.

Theorem 3.31 ([Shv90]) *For a consistent set T, T is an extension of A iff*

$$T = Cn_{K45}(\Sigma \cup \{\neg L\varphi | \varphi \notin T\})$$

where Cn_{KD45} is the logical consequence operator[12] *of modal logic KD45.*

Comparing this equivalent definition for an expansion with that of (9) we can see that the difference between AEL and NML can be seen to be precisely the use of KD45 in AEL and classical entailment in the original NML.

Each stable expansion of T is a *stable set*, i.e.,

- T is closed under propositional consequence;
- If $\varphi \in T$, then $L\varphi \in T$;
- If $\varphi \notin T$ then $\neg L\varphi \in T$.

Example 3.32 *From the theory $T = \{bird \wedge Mflies \Rightarrow flies, bird\}$ we want to prove $flies$. Since $\neg flies \notin T$ we get $\neg L \neg flies \in T$ and so $Mflies \in T$. Thus $flies \in T$.*

Theorem 3.33 ([Moo85]) *If two stable sets agree on nonmodal formulae, they are equal.*

For a given modal language $\mathcal{L}$, we write $\mathcal{L}_0$ for the base language with no modal operators. For a theory Γ, we write Γ_0 for its restriction to $\mathcal{L}_0$.

Theorem 3.34 ([Kon88b]) *Every stable set Γ is an AE stable expansion of Γ_0.*

[12] C is the consequence operator of an entailment relation $\vdash$ iff $C(A) = \{\varphi | A \vdash \varphi\}$.

Theorem 3.35 ([Kon88b]) *Let W be a set of nonmodal formulae closed under logical consequence. There is a unique stable set Γ such that $\Gamma_0 = W$. W is called the* kernel *of the stable set, and the stable set is called $ST(W)$.*

We discuss semantics for AE logic in section 5.4 on page 84.

Historical Notes Instead of having axioms like in Modal logics of belief (e.g. KD45), Moore presented his autoepistemic logic by defining the notions of an *autoepistemic theory*, an *autoepistemic interpretation*, an *autoepistemic model*, a *stable interpretation* and a *stable expansion*. Here we use the modality L for belief. We say that T is a stable expansion of a set of axioms A, if

- $A \subset T$
- T is *grounded in* A, i.e, is the tautological consequence of $A \cup \{LP | P \in T\} \cup \{\neg LP | P \notin T\}$, and
- T is *stable*, i.e., $P \in T \implies LP \in T$, and $P \notin T \implies \neg LP \in T$.

Defining the consequence relation is then simple: $A \mid\sim_A \varphi$ iff every stable expansion of A contains φ.

Moderately Grounded Expansions Stated simply, groundedness is the principle that the result of reasoning should in some way be based on the premises. In Autoepistemic logic, fixed points use the premise set but also the final belief set. The notion of groundedness is important because it distinguishes AEL from nonmonotonic formalisms that use minimization techniques (e.g., Circumscription) as well as correlates AEL with other derivation-sensitive noncumulative formalisms (e.g., Default Logic, JTMSs). Since the fewer assumptions that are made in the fixed-point equation, the stronger the groundedness of the resulting system, there was an attempt to strengthen the notion of stable expansions.

There are three main classes of groundedness: weak (these are all the regular stable expansions), moderate and strong.

Definition 3.36 (Moderate Groundedness) *A set of sentences T is moderately grounded in A if*

$$T = Cn_{K45}(A \cup \{\neg L\varphi | \varphi \in \mathcal{L}_0 \wedge \varphi \notin T\})$$

where $\mathcal{L}_0$ is the base language with no modal operators.

Every moderately grounded expansion is an expansion (see the following theorem), but not every expansion is moderately grounded. The premise set $A = \{LP \Rightarrow P\}$ has two expansions, $ST(P)$ and $ST()$, but only one moderately grounded expansion, $ST()$.

We say that an extension T of A is *minimal for* A if there is no stable set S containing A such that $S_0 \subset T_0$ (S_0, T_0 are S, T restricted to the base language).

Theorem 3.37 ([Kon88b]) *A set of sentences is moderately grounded in A iff it is a minimal AE expansion of A.*

Strongly Grounded Expansions The notion of Strongly Grounded Expansion is a more syntactical one (a property that will prove useful in translating Default Logic Extensions to Strongly Grounded Autoepistemic Expansions (see section 6.2 on page 86)).

Definition 3.38 (Normal Form) *The AE theory A is in* normal form *if its sentences have the form*

$$\neg L\alpha \vee L\beta_1 \vee ... \vee L\beta_n \vee \omega$$

with α, β_i, ω all being nonmodal sentences. Any of the disjuncts may be absent and if $\neg L\alpha$ is absent, the normal form is called simple.

A is said to be in *extended normal form* if it is in normal form and all of the disjuncts are present.

Definition 3.39 (Strongly Grounded Expansion) *Let A be a set of AE sentences in normal form, and let T be an extension of A. Let A' be the set of sentences $\neg L\alpha \vee L\beta_1 \vee ... \vee L\beta_n \vee \omega$ of A such that none of $\beta_1, ..., \beta_n$ is contained in T. Then T is* strongly grounded in A *iff*

$$T = Cn_{K45}(A' \cup \{\neg L\varphi | \varphi \in \mathcal{L}_0 \wedge \varphi \notin T\})$$

Strongly Grounded AEL differs significantly from AEL. In particular, in SGAEL it may be the case that two logically equivalent but syntactically different sets of premisses have different semantics, that is, different strongly grounded expansions.

Theorem 3.40 ([Kon88b]) *If T is a strongly grounded expansion of A, it is moderately grounded in A.*

Not every moderately grounded expansion is strongly grounded. The set of facts $A = \{\neg LP \vee P, LP \vee \bot\}$ has the moderately grounded expansion $ST(P)$. This is not a strongly grounded expansion in A, as ignoring $LP \vee \bot$ in selecting A' (and the strongly grounded expansions) leaves us no way to derive P. Notice that the equivalent theory $B = \{P, LP \vee \bot\}$ has $ST(P)$ as a strongly grounded expansion.

Theorem 3.41 ([Gel87]) *If a theory A has no negative modal atoms (it is then said to be in* simple normal form*) then every extension of A is strongly grounded.*

3.3.3 "All I Know"

Autoepistemic logic, as originally conceived, is a logic of self-belief used internally by an agent. Levesque [Lev90] decided to learn autoepistemic logic on a meta-level. The advantage of doing so is a more expressive formalism and the

possibility of studying self-belief in a more declarative setting. The major drawback of using meta-theory is that complicated objects must often be introduced and soon enough the picture becomes too cluttered to decipher.

In his study of AE logic, Levesque [Lev90] has made the interesting choice of adding just enough additional expressivity to form a meta-theory, while retaining most of the formal apparatus of AE logic. Instead of axiomatizing sets of AE sentences and their inferential relations, he introduces a modal construct $O\alpha$ whose intended meaning is "α is all I know".

The language $\mathcal{OL}$ of "objective" AEL has two modalities: L and O. $L\varphi$ now refers to the beliefs of the agent according to an outside observer, rather than self-belief. $L\varphi$ means that the agent believes φ, while $LL\varphi$ means that the agent believes that he believes φ.

The semantics[13] of O is a possible-worlds semantics that is built on top of that of **K45** (see section 2.3.3 on page 13):

$$w \models_W L\varphi \iff \forall w'.(w' \in W \implies w' \models_W \varphi)$$
$$w \models_W O\varphi \iff \forall w'.(w' \in W \iff w' \models_W \varphi)$$

As was mentioned in section 2.3.3, for the logic **K45** one can assume that the set of accessible worlds is fixed, and so there is no need to explicitly use an accessibility relation. W here is the set of accessible worlds (independent of w) in the structure.

Theorem 3.42 ([Lev90]) *For any basic sentence α (no modal operators) and a possible-worlds set W, $S(W)$ (the set of basic sentences true in all worlds of W) is an extension of α iff $\models_W O\alpha$.*

Theorem 3.43 ([Lev90]) *A basic sentence α has exactly as many stable expansions as there are maximal sets of worlds where $O\alpha$ is true.*

By maximal we mean that any new world w' added to W would change the truth value that W assigns to some basic sentence.

For an axiomatic account of the first-order case, the reader should look at section 2.3.6.

3.3.4 Other Modal Systems

There are many variants of the modal systems discussed above. [Kon93] discusses Autoepistemic Logic and its variants. [MST91] discusses variants of nonmonotonic logic using different modal logics.

There are cases in which a consistent theory has no stable expansions (e.g., the theory $A = \{\neg LP \Rightarrow P\}$). This problem is similar to the corresponding problem with Reiter's Default Logic, as the construction of an expansion encounters a self contradicting loop. The solutions for Autoepistemic Logic come in two main flavors. The first is adding some additional assumptions, as in

[13] We provide the semantic account here because this is the definition for the operator O as given by Levesque. The semantics for AEL are given in section 5.4 on page 84.

Stable Closures [Mor89]. This approach advocates adding a random minimal (in the sense of set-inclusion) set of sentences that will allow the production of an expansion.

The other approach, which had several followers is the **Three-valued AE Logic** approach. Here, we would rather stay uncommitted to either LP or $\neg LP$ in the case that the assumption P is judged as overly strong (i.e., give rise to inconsistent AEL). This approach was pursued by several authors such as Przymusinski [Prz89a] and Bonatti [Bon91],[Bon95]. Recently, Denecker et al. [DMT98] gave a 3-valued semantics for AEL.

A **Hierarchic AE Logic** [Kon88a] is a modified version of the logic that addresses the problem of priorities among extensions. It is derived from AE logic by splitting a uniform theory into components, called *subtheories*. The subtheories are meant to serve as bodies of evidence: those lower in the hierarchy are considered to be stronger evidence, and conclusions derived from them take precedence over defaults in subtheories higher in the hierarchy. An interesting application of HAEL is to the formalization of speech acts using defaults about what is commonly inferred to be believed among the agents [AK88].

3.4 Hybrid Nonmonotonic Systems

For different reasons, mostly related to the superiority of different paradigms in different areas (application areas in common-sense reasoning, computational complexity, expressivity etc.), some hybrid systems were proposed over the years to bridge over the distances between the systems mentioned above. These are described in this section.

The first two systems we discuss try to use some form of circumscription yet caring to deal with modal logics and some notions of knowledge and belief. These systems come to bridge the gap between Autoepistemic Logic and Circumscription or to allow Circumscription to easily capture modalities[14]. Other systems that we do not count here that go along the same lines are due to Lin [Lin88] and Wainer [Wai92].

3.4.1 Autocircumscription

Perlis' Autocircumscription [Per88] was the first to try using an idea from AEL and an idea from Circumscription to create a nonmonotonic system that combines both. Perlis' idea was to view the Circumscription formula as indeed saying two things. First, one notices that something is not concluded. Then, one conjectures that this thing is not true.

In fact, Perlis introduced two different formulas: GenK (General Knowledge) and MinK (Minimal Knowledge). Both formulas are in fact *First-Order formulas (not modal) but with reified wffs* (he uses Feferman's scheme [Fef84] and mentioned his own work [Per85] as an alternative). To better compare his

[14]It was argued that one of the reasons Circumscription can not simulate AEL is that it can not make the difference between the truth of P and the knowledge of P (LP).

proposals, let us write them all down next to the original Circumscription formula (the formula circumscription form). W is a wff and $A(P,Z)$ is a set of axioms.

$$
\begin{array}{l}
Circumscription: \\
A(P,Z) \wedge \forall p,z,y.[A(p,z) \wedge (\forall x)(W(p,z,x) \Rightarrow W(P,Z,x)) \wedge \\
\qquad\qquad \neg W(p,z,y) \Rightarrow \neg W(p,z,y)] \\
MinKAutocirc.: \\
A(P,Z) \wedge \forall p,z,y.[A(p,z) \wedge (\forall x)(W(p,z,x) \Rightarrow W(P,Z,x)) \wedge \\
\qquad\qquad \neg W(p,z,y) \Rightarrow \neg KW(p,z,y)] \\
GenKAutocirc.: \\
A(P,Z) \wedge \forall p,z,y.[A(p,z) \wedge \neg W(p,z,y) \Rightarrow \neg KW(p,z,y)]
\end{array}
$$

If one is interested in simulating Circumscription, one should use MinK together with the axiom

$$\neg K(Py) \Rightarrow \neg Py \tag{10}$$

On the other hand, if one is interested in the extent of ignorance, $GenK$ is more suitable. Notice, that adding formula (10) can lead to unwanted conclusions (e.g., if $A = Pa \vee Pb$, we infer from GenK that both $\neg KPa$ and $\neg KPb$ are true, thus concluding $\neg Pa \wedge \neg Pb$ implying contradiction (with A)). It is $GenK$ that Perlis prefers, calling it *Autocircumscription*.

The predicate K gets it meaning using the demand that the theory T is *autoconsistent*.

Definition 3.44 *A theory T is* autoconsistent *(for the predicate letter K) if the language of T is reified and has the predicate letter K and for every wff α:*

1. *$T \vdash K\alpha$ implies $T \vdash \alpha$ (similar to the modal rule* **4***), and*
2. *$T \vdash \neg K\alpha$ implies $T \nvdash \alpha$ (similar to the modal rule* **5***).*

Note that, so far, there is no reason to conclude $K\varphi$ unless we have some other axioms about K in our theory T. If we add the axioms of an arbitrary modal logic about K, this situation may change.

Theorem 3.45 *If the only proper axioms of T that involve K are literals, and if T is autoconsistent, then $AUTO[T]$ is consistent.*

As a special case, this implies consistency for Autocircumscription of autoconsistent theories without the symbol K. Notice that this theorem is also true for $MinK$ as that formula is a restricted form of $GenK$.

Example 3.46 *How can an agent decide that, on the basis of all it knows, the question as to whether Bill Clinton is (currently) standing or seated is indeterminate? Let an agent have the following axiom $A(Seated, Standing)$:*

$$\left\{ \begin{array}{l} Seated(Ron), \neg Seated(Sue), Seated(x) \iff \neg Standing(x), \\ Bill\ Clinton \neq Ron, Bill\ Clinton \neq Sue \end{array} \right\}$$

Then, for $W(P, Z, y) = Seated(y)$, $p = \lambda x.(x = Ron), z = \lambda x.(x \neq Ron)$, *we get that* $A(p, z)$ *and* $\neg Seated(BillClinton)$. *Thus, we conclude*

$$AUTO[A] \vdash \neg K(\text{“}Seated(BillClinton)''\text{)}$$

Perlis did not provide relationships with Autoepistemic Logic or the other modal approaches[15].

3.4.2 Introspective Circumscription

Lifschitz's Introspective Circumscription [Lif89b] is a subsequent effort of combining AEL and Circumscription. The language is second-order logic, and instead of reifying wffs, Lifschitz restricts himself to expressing "believing P" for a predicate constant symbol P only.

For a predicate symbol P_i, a new predicate constant symbol LP_i is added to the language, of the same arity. For any sentence $A(P, LP)$, its *Introspective Circumscription* $A^I(P, LP)$ is the formula

$$A(P, LP) \wedge \bigwedge_{i=1}^{n} \forall x_i[LP_i(x^i) \equiv \forall p(A(p, LP) \Rightarrow p_i(x_i))]$$

The conjunctive term (the second argument of the top-level $\wedge$) is called the *introspection condition for* P_i and it expresses the "rationality" of the agent: he believes that $P_i(x_i)$ holds iff this follows from his beliefs A.

There are three interesting special cases of the Introspective Circumscription formula. The first is when the set of axioms A does not contain LP, in which case we get a specific definition for LP: $A(P) \wedge \forall x[LP(x) \equiv \forall p(A(p) \Rightarrow p(x))]$ (for example $A = P(a)$ implies that $A^I \equiv P(a) \wedge \forall x(LP(x) \equiv (x = a))$). A second interesting case is when A does not contain P, in which $A^I \equiv A(LP) \wedge \forall x \neg LP(x)$. This case is a counterpart to the fact that, in AEL, the axiom LP has no stable expansion.

The third interesting case is for an arbitrary set of axioms $A_0(P)$ that does not contain LP, to have $A(P, LP) = A_0(P) \wedge \forall x(P(x) \Rightarrow LP(x))$. Then we get an equivalent formula to Circumscription (for the language without LP).

Theorem 3.47 *Let* $A(P, LP) = A_0(P) \wedge \forall x(P(x) \Rightarrow LP(x))$ *for* $A_0(P)$ *a formula with no occurrences of* LP. *Then,* A^I *is equivalent to*

$$Circ[A_0(P); P] \wedge \forall x(LP(x) \equiv P(x))$$

Lifschitz [Lif89b] shows how different formulas (on top of A_0) can allow us to get the same theorem for more extended forms of circumscription, such as Parallel Circumscription and Prioritized Circumscription.

[Lif89b] shows how to give semantics to Logic Programs using Introspective Circumscription.

[15] My main feeling about it is that it is half-baked (compared to the modal approaches). If there was a clear relationship with one of the modal nonmonotonic systems, it would have made things much clearer.

A simple example of using Introspective Circumscription for commonsense reasoning is the following: Given that a block A is on the table, conclude that for any block other than A it is not known whether it is on the table ([Lif89a] Problem E3). The only axiom is $ontable A$, and introspective circumscription gives $L_ontable x \equiv x = A$.

3.4.3 Model Preference Default Theories (Juliet)

Kautz and Selman [KS89], [KS91], [SK90] defined *Model Preference Default Theories* with the intention of producing an efficient (computational complexity-wise) mode of nonmonotonic reasoning. They defined their system for propositional language only and focused their attempts at horn theories, although their treatment is for general propositional logic clauses.

Semantically-defined (as is Levesque's All-I-know), the model theory for Model Preference Default Theories is similar to that of circumscription, in that it involves considering models which are preferred, based on some order relation. Where they differ is that the conclusions of a circumscriptive proof must hold in *all* preferred models (cautious entailment), and in the fact that the order relation in a circumscriptive theory is defined solely in terms of minimizing predicates. The syntax is non-the-less similar to that of Default Logic, arguing that the representation of defaults in Default Logic is compellingly intuitive.

The original system of defaults that Selman and Kautz propose is called $\mathcal{D}$. Later they define an extension to $\mathcal{D}$ by adding priorities to defaults ($\mathcal{D}^+$) and examine some restricted cases of defaults such as Horn Defaults $\mathcal{DH}$ (and $\mathcal{DH}^+$) and acyclic sets of horn defaults $\mathcal{DH}_a^+$.

Definition 3.48 (Default Rule) *A* default rule *d is an expression of the form $\alpha \Rightarrow x$ where α is an arbitrary propositional sentence and x a single literal (i.e., a positive or negative propositional constant symbol).*

The rule d is called a horn default rule *iff α contains only positive literals. A rule d is* positive *if it has a positive literal on the right-hand side. Otherwise, d is* negative.

Interpretations/Models are usual truth assignments. A default rule $\alpha \Rightarrow x$ is said to be *applicable at the model M* iff $M \models \alpha$.

Definition 3.49 (Model-Preference Relation) *Let d be a default rule, of the form $\alpha \Rightarrow x$, that is applicable at M. The* application of rule d at M *is any model (there may be several) M' such that $M' \models x$ and for all propositional constant p that does not show in x, $M' \models p \iff M \models p$. For each such M' we write $M \Rightarrow_d M'$ and if $d \in D$, we also write $M \Rightarrow_D M'$.*

For a set of default rules D, the model-preference relation $\leq_D$ is the reflexive, transitive closure of $\Rightarrow_D$.

Note that the entailment relation depends on the syntactic form of the formula x.

Definition 3.50 (Model-Preference Default Entailment) *Let D be a set of defaults and A a set of propositional formulae. M is a maximal model of D and A iff $M \models A$ and $\forall M'(M' \leq_D M \wedge M' \models A \implies M \leq_D M)$ (i.e., M is minimal (modulo $\leq_D$-equivalence) satisfying A).*

One way to write an entailment relation for MPD is $M \models_D \varphi$ iff M is a maximal model of D and A. Then, $T \models_D \varphi$ iff $\forall M.(M \models_D T \implies M \models_D \varphi$. This nonmonotonic system is called Model-Preference Default Theory (MPD), $\mathcal{D}$.

Example 3.51 *Let us observe the following set of default rules D:*

$$\begin{array}{l} student \Rightarrow adult \\ adult \Rightarrow employed \\ student \Rightarrow \neg employed \end{array}$$

If $A = \{student\}$ is the set of facts, then figure 6 displays the preference ordering $\leq_D$, crossing off the models that do not satisfy A. Thus, we can conclude adult but cannot conclude employed, as both $sa\bar{e}$ and $sa\bar{e}$ are maximal models.

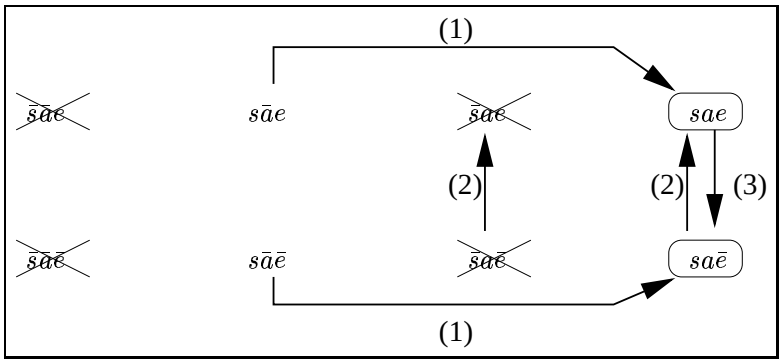

Figure 6: The preference ordering $\leq_D$. s, a, e stand for $student, adult, employed$ respectively.

From this system (which they call $\mathcal{D}$), Selman and Kautz go to define a system of defaults with specificity built-in.

Definition 3.52 (Specificity-Ordered Defaults: $\mathcal{D}^+$) *Given a set of defaults D, a default $d = \alpha \Rightarrow x$ is* blocked *at a model M iff D contains a default of the form $(\beta \cup \alpha) \Rightarrow \neg x$ and $M \models (\beta \cup \alpha)$. It is* applicable *at M iff $M \models \alpha$ and d is not blocked at M.*

Thus, they define a default to be more *specific* than another if the other's precondition includes the former's. Redefining applicability for that system $\mathcal{D}^+$ then defines a new entailment relation following the minimal models in $\leq_D^{\mathcal{D}^+}$ that satisfy a set of facts A.

It is interesting to notice that $\leq_D^{\mathcal{D}^+} \subset \leq_D$ as a result of the fact that $\Rightarrow_D^{\mathcal{D}^+} \subset \Rightarrow_D$.

Example 3.53 *In our previous example (3.51 on the preceding page), we can now specify a new default*

$$student \wedge adult \Rightarrow \neg employed$$

which blocks the default adult ⇒ employed, allowing us to conclude that the only preferred model is saê, concluding ¬employed.

This is the only priority information that Selman and Kautz describe and there is no general *Prioritized* version of their method.

A subclass of default rules is the class of *Horn default rules* (only rules of the form $p_1, ..., p_n \Rightarrow q$ are allowed, for $p_1, ..., p_n, q$ positive or negative propositional constant symbols), written $\mathcal{DH}, \mathcal{DH}^+$ for the horn versions of the regular $\mathcal{D}$ and specificity-ordered $\mathcal{D}^+$ default rule systems.

Another subclass of default rules is the class of *acyclic defaults*. A set of defaults D is called *acyclic* if there is no chain of default rules $d_1, ..., d_n$ such that the consequent of each rule d_i has some literals that show (either positive or negative) in the precondition of d_{i+1}. The system of sets of defaults that are Horn, Specificity-ordered and Acyclic is written $\mathcal{DH}_a^+$. Examples 3.51 and 3.53 are in $\mathcal{DH}_a^+$.

It was shown by Selman and Kautz that some subclass of $\mathcal{DH}^+$ is representable as a prerequisite-free semi-normal default theory in Reiter's Default Logic. More on that in section 6.5.1 on page 91.

3.5 Other Nonmonotonic Systems

In this section we briefly overview some other nonmonotonic systems that came from different backgrounds and different intentions. We do not give a comprehensive survey of these systems, but try to give the reader a glimpse at the important behaviors these systems offer. At the end we list yet other systems that came over the years and we did not have the space to comprehensively describe.

3.5.1 Truth Maintenance Systems

Truth Maintenance Systems (TMS) were devised as book-keepers for the knowledge of companion problem solvers. Typically, a problem solver updates the TMS with the reasoning that it did, and the TMS updates the dependencies and tells the problem solver what to believe in, or what are possible causes for a contradiction that arose. This architecture is appealing because one separates two functionalities of the agent, creating a better manageable and augmentable system.

In this section we will informally describe the different major TMS architectures, touching at their major characteristics, up-sides and down-sides. At the end of this section, we give some references to some formal work done on giving TMS semantics and interpretation in other nonmonotonic systems.

JTMS Justification based TMSs [Doy79] keep a network of nodes, each related to some other nodes with connections that describe their inter-dependencies. Each node represents an outside world proposition (i.e., a proposition of the problem solver), which is kept in that node as a label with no semantics, a classification of the node as *IN* or *OUT* (i.e., believed in, or not believed), and a set of justifications (each with a form of a list of nodes that should be IN, and a list of nodes that should be OUT). If no justification is fulfilled, the node is assumed (nonmonotonically) to be OUT.

The problem solver can define a node to be a contradiction node, in which case, if it is IN, it is a sign that some of the assumptions (those drawn nonmonotonically) should be changed, or that there is an inconsistency in some of the assertions of the problem solver. In the first case - the JTMS chooses a *culprit* to be the one to blame for the inconsistency, and then a justification is added to that node that includes the other values of the nodes that caused the contradiction. In the second case the JTMS notifies the problem solver of a contradiction, and of the possible reasons for that contradiction.

JTMSs are used for dependency directed backtracking (in Search), Reasoning tasks (as described above, in the problem-solver—TMS relationship), CSPs, and realizing nonmonotonic systems and belief revision.

Positive Gives a first set of hypotheses fast, disconnected from the reasoner and the meaning of its nodes, is elaboration tolerant, and established a nice protocol for problem-solver—TMS interaction.

Negative Difficult to decide on the culprit, does not give the set of all possible hypotheses sets, does not represent the contexts explicitly, expensive to switch contexts (i.e., to find a new solution once an assumption was withdrawn), detection of loops requires special treatment.

Complexity issues SAT is co-NP-complete, and thus the JTMS algorithms to date are worst case exponential. Also, the detection of loops is complicated.

LTMS Logical TMSs [McA80] add semantics to the nodes of JTMSs, letting each node be a proposition, so that the LTMS can detect contradictions even when the problem solver did not select a node to be a contradiction node.

ATMS Assumption based TMSs ([dK86a], [dK86b])

can be seen as running a set of JTMSs in parallel, each one choosing a different set of assumptions, so that all the possible sets of assumptions that can be taken by the JTMS at a given time are accounted for.

In practice, an ATMS has a set of assumptions for each node such that the node holds under this set of assumptions. The set of assumptions is then updated as the problem solver adds more assertions or nodes to the ATMS.

If a set of assumptions is contradictory, then every set of assumptions that contains it is contradictory. As a result, there is no need to hold all the possible

contradictory assumptions, but only a small set that spans all the contradictory assumptions (the prime implicants). This set is called the *no good* set. This property is used to keep the updating process efficient and simple.

Among the uses of ATMSs, is realizing nonmonotonic systems. [Gin89] had a circumscriptive theorem prover that relies on an ATMS. [RDB89] and [JK90] use ATMSs to compute extensions of Autoepistemic Logic and Default Logic. Other applications include diagnosis ([de 86], [dW87], [RdK87]), vision [Pro88] and qualitative reasoning [For86].

Positive Gives a complete set of sets of hypotheses for every given node, disconnected from the reasoner and the meaning of its nodes, no need to decide on culprit, because we account for all the possible sets of assumptions (in a sense), the contexts are expressed explicitly, and loops are dealt with naturally as the labels of a node (its list of contexts) being updated does not update any more.

Negative It is more complicated to come with only one set of hypotheses if that is all we need.

Complexity issues SAT is co-NP-complete, and thus the ATMS algorithms to date are also worst case exponential.

Summary Every type of TMS has its own special behavior and associated problems, but at least one recurring problem with this architecture, that is worth noting, is the *unouting* problem. By the *unouting* problem we refer to the phenomenon of a TMS changing its mind and, as a result, some computations must be redone in the process (in the problem solver).

Several formal treatments were given to the various TMSs. Reiter and de Kleer [RdK87] and Caryol, Palmade and Schiex [CPS93] give formal semantics to ATMSs. Elkan [Elk90] and Reinfrank, Dressler and Brewka [RDB89] present TMSs as formal nonmonotonic systems. Finally, McAllester [McA90] compared the different TMSs with respect to the respective algorithms and applications.

It is noteworthy that although JTMSs are nonmonotonic, ATMSs are actually monotonic. This is interesting especially given the other observation that ATMSs are in fact generalizations of JTMSs.

3.5.2 Negation As Failure and Closed World Assumptions

The notion of *Negation as Failure to Prove* (NAF) refers to any of a set of variants on the idea that if something is not proved, it can be assumed false. This set of variants includes *Naive NAF* (what was originally called Negation as Failure) and the various forms of the Closed World Assumption (described below). Referred to by itself, NAF refers to what we call here Naive NAF.

Originally, the notion of NAF was studied by Clark [Cla78], trying to formalize the primal form of non-monotonic reasoning that was put into extensive use in the form of logic programming languages such as PLANNER [Hew72]

and PROLOG [Coh88], [SS94b] (in fact, these programming languages used an approximation called "negation as finite failure").

Now days, the notion of *Naive Negation as Failure* is mostly used to denote the practical method that treats failure to prove a statement as proving the negation of the statement. I.e., using Negation As Failure for φ in T, we can assume $\neg\varphi$ if $T \not\vdash \varphi$. If not used with caution, NAF can be harmful.

Example 3.54 ([Rei78]) *Suppose that* $T = Teach(pascal) \vee Teach(prolog)$. *The naive NAF rule applied to* T *allows us to assume* $\neg Teach(pascal)$ *and* $\neg Teach(prolog)$ *and therefore leads to contradiction.*

Reiter [Rei78] studied the notion of (Closed World Assumption (see below)), motivated mostly by the theory of closure in Databases. Whereas Negation as Failure is defined for all sentences φ, Reiter [Rei78] defined a refined version called the *Closed World Assumption* (CWA) that works similarly, but restricts the NAF rule to apply only for ground atoms.

Definition 3.55 (Closed World Assumption) *The* Closed World Assumption with respect to P *is the assumption that the positive information existing in our theory* T *about predicates from* P *is complete, i.e., no other positive information about* P *is available.*

More formally,

$$CWA(T) = T + \{\neg p \mid p \text{ is a ground atom such that } T \not\vdash p\}$$

Theorem 3.56 ([Rei78]) *For a theory* T *consisting of Horn clauses*[16] *the following holds:*

1. *If* T *is consistent, then* $CWA(T)$ *is consistent.*
2. *If* T *is consistent, then for any ground formula* f *we have* $T \not\vdash f \iff CWA(T) \vdash \neg f$.

In [VK76] a semantical characterization of $CWA(T)$ was given in terms of minimal Herbrand models of T. Shepherdson [She84] generalizes Reiter's consistency result to definite theories (a *definite theory* T is a theory for which for any ground disjunction $D = p_1 \vee ... \vee p_n$, $T \vdash D$ iff $T \vdash p_i$ for some i) and shows that for Horn theories the resolution used in PROLOG is sound with respect to $CWA(T)$.

Shepherdson [She84] compares the two notions of NAF and CWA.

Predicate Completion *Predicate Completion* [Cla78],[Kow78] is a device for "closing off" a first order representation. This concept stems from the observation that frequently a world description provides sufficient, but not necessary, conditions on one or more of its predicates and hence is an incomplete description of that world.

[16] Horn clauses are disjunctive clauses with at most one positive literal.

After aggregating all the sufficient information for $P(x)$ in T into a formula $\forall x.\varphi(x) \Rightarrow P(x)$, Predicate Completion for P adds the *completion axiom* $\forall x.P(x) \Rightarrow \varphi(x)$ to T.

For Horn Theories, CWA defined above is enough to produce the needed Predicate Completion. For more complicated cases, Predicate Completion may not be that intuitive any more (for example, what should happen if there is no finite one prime-implicant as a sufficient condition for P?). Nevertheless, if our theory is equivalent to $\forall x.\mathcal{F}(x) \Rightarrow P(x)$, where the formula F does not contain an instance of P, Circumscription and some of the Negation as Failure variants that we describe below are capable of producing the right completion axiom.

Theorem 3.57 ([Rei82],[Lif85b]) *If $F(x)$ does not contain P, then the circumscription*

$$Circ[\forall x(F(x) \Rightarrow P(x)); P]$$

is equivalent to $\forall x(F(x) \equiv P(x))$.

Generalized Closed World Assumption Minker [Min82] tried to generalize the Closed World Assumption to be "safe" for more general theories (than Horn).

Definition 3.58 (Generalized Closed World Assumption) *The Generalized Closed World Assumption applied to a theory T results in the closure $GCWA(T)$ of T defined as follows:*

$$GCWA(T) = T + \left\{ \neg p \;\middle|\; \begin{array}{l} p \text{ is a ground atom and there is no ground} \\ \text{disjunction } B \text{ of atoms such that } T \vdash p \vee B \\ \text{and } T \nvdash B \end{array} \right\}$$

GCWA allows us to construct query answering systems capable of handling common-sense situations as the one discussed in example 3.54 on the preceding page. Such a system will give us a negative answer to a query $Teach(KR)$ but it will reply "maybe" to a query $Teach(pascal)$. Unlike the NAF, Minker's rule works only for literals. Yahya and and Henschen [YH85] extended Minker's system and results to allow application of Minker's GCWA to arbitrary formulae.

Extended Closed World Assumption Both CWA and GCWA do not allow us to prove any new positive facts.

Theorem 3.59 ([Min82]) *For any positive ground formula f*

$$GCWA(T) \vdash f \iff T \vdash f$$

For example, if we wish to conclude $flies(Tweety)$ from $\{bird(Tweety),\ bird(x) \wedge \neg ab(x) \Rightarrow flies(x)\}$, neither of them is helpful. An extension of Minker's GCWA is *Careful Closed World Assumption* (CCWA) proposed by Gelfond and Przymusinska [GP86] to counter this problem. Later, Gelfond, Przymusinska and

Przymusinska and Przymusinski [GPP86],[GPP89] devised a generalization of CCWA that allows formulas other than ground literals to be "closed" (CCWA allowed application of the closure to ground literals only). This version is called Extended Closed World Assumption (ECWA) and we bring it here.

Given a list R of predicates, R^+ [R^-] will denote the set of all positive [negative] ground literals with predicate symbols in R. In the following, T is a theory and P, Z are sets of constant symbols (P is a set of *predicate* constant symbols) as in Circumscription (section 3.1). Let Q be the list of all remaining predicate symbols from T.

Definition 3.60 *Let K be an arbitrary sentence* not *involving predicate symbols from Z. We will say that K is* free for negation in (T, P, Z) *if there exists* no *disjunction φ of ground literals from $P^+ \cup Q^+ \cup Q^-$, such that*

1. $T \vdash K \vee \varphi$
2. $T \nvdash \varphi$

Roughly speaking, a sentence is free for negation if it will not be forced to be positive after we make some closed world assumption about P.

Definition 3.61 (Extended Closed World Assumption) *The* Extended Closed World Assumption *with respect to P and Z applied to the theory T results in the closure*

$$ECWA(T; P; Z) = T + \{\neg K \mid K \text{ is free for negation in } T\}$$

Theorem 3.62 *Assume that T is a first-order theory over a language $\mathcal{L}$ with a finite number of individual constant symbols and no function symbols. The theory T is also assumed to include equality axioms and the Domain Closure Assumption (DCA) and the Unique Name Assumption (UNA) axioms (see section 7.4 on page 98 for their definition).*

Then, for any sentence $\varphi \in \mathcal{L}$,

$$Circ[T; P; Z] \models \varphi \iff ECWA(T; P; Z) \vdash \varphi$$

Other Negation-As-Failure Notions A further extension of ECWA that we do not consider here for lack of space is Iterated Closed World Assumption [GPP89] (ICWA). Roughly speaking, it applies ECWA to the less prioritized predicates first ("varying" those predicates with even lower priority), and then applies ECWA in a recursive fashion on its way up to the predicate with the highest priority. Thus, it covers the case of priorities between predicates and a correspondence to Prioritized Circumscription, similar to the one mentioned in the last theorem, is shown in [GPP89].

Figure 7 on the following page gives a summary of the development of the systems we described above.

Other works on Negation as Failure that we do not consider here are works on delivering a variant of Negation As Failure that is "safe" (as demonstrated

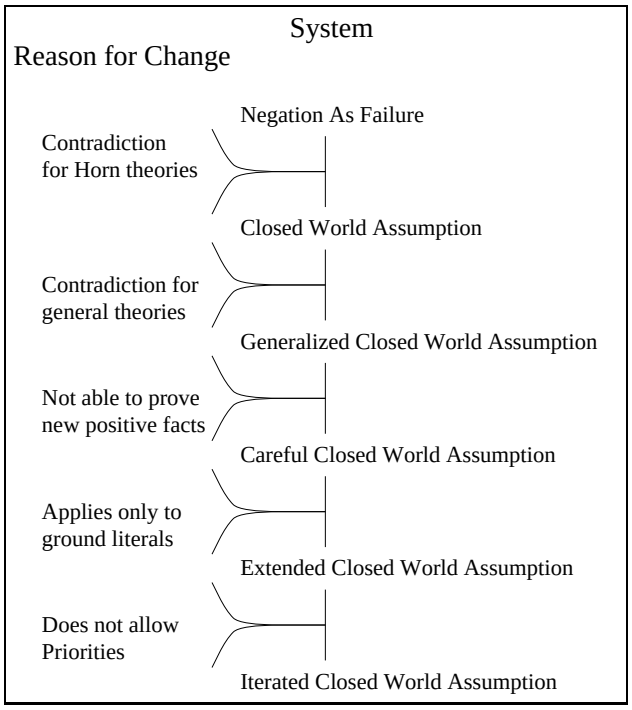

Figure 7: Progress through the different variants of Negation as Failure

above) but also has nice computational properties (the versions described above (especially ECWA and its variants) are not easily computable). The work of Geffner [Gef91] tries to define such a notion, and discusses different variants.

Others that studied variants of NAF include Naqvi [Naq86] and Rajasekar et al [RLM89].

3.5.3 Poole's Maxiconsistent Sets

Poole's Maxiconsistent sets [Poo88b], [Fre98] bear the same intuition as that of Reiter's DL, i.e., that certain items are assumed true unless special circumstances prevent their use or application.

A *Poole system* is a pair $\langle D, K \rangle$ where D, K are both sets of closed propositions in the language $\mathcal{L}$ of ordinary first-order or propositional logic. D is called the set of *default propositions* of the system, and K is called the set of *constraints* of the system.

Intuitively, the idea is that we would like to assume true as many of the propositions in D as possible. We would like to take the maximal (with respect to set-inclusion) subsets of D that are consistent with each other and with the constraints of K.

Given any Poole system $\langle D, K \rangle$, we define its associated extension family

function by the rule that for every set $A \subseteq \mathcal{L}$ of propositions,

$$e(A) = \left\{ Cn(A \cup D) \;\middle|\; \begin{array}{l} D \text{ is maximal among the subsets } D' \text{ of } D \\ \text{such that } A \cup D' \cup K \text{ is consistent} \end{array} \right\}$$

The *cautious* inference relation $\mid\sim_{D,K}$ is then defined as for Reiter's DL: $\alpha \mid\sim_{D,K} \beta$ iff $\beta \in \cap e(\alpha)$.

3.5.4 Inheritance Networks with Exceptions

Inheritance Networks with Exceptions were gradually introduced in [ER83], [Tou86], [THT86], and [DG93b]. Later works include [Nar96], [Mor96], [Mor98]. Here we include only an informal description of this mechanism, and the reader is invited to look at [DG93b] for precise details.

Inheritance systems were originally developed within artificial intelligence in response to the practical need for an efficient way of representing and accessing taxonomic information. Of particular interest were the works on semantics nets and frame-based systems [Woo75].

An *inheritance net* is taken to be a set of links, where a *link* is a proposition of one of the forms "an individual i is (is not) a q" and "p-s typically are (typically are not) q-s". We write links as iq, $i\neg q$, pq and $p\neg q$ (the links are intuitively pointing *up* the inheritance network). Each link thus has a sign and a direction.

Typically, nonmonotonic inheritance networks are represented using graphs in which links are represented with arrows, as in $p \rightarrow q$ for positive links and $p \not\rightarrow q$, $i \not\rightarrow p$ for negative links. However, this should not be confused with the classical negation of a positive link: "p-s are typically not q-s" says more than "p-s are not typically q-s".

A *path* over a net is a finite sequence $p_1, ..., p_{n+1}$, where p_1 is either an individual or a predicate, all the other p_j's are predicates, and for each $k \leq n$ the net contains a link from p_k to p_{k+1}, positive in all cases, except possibly the last. A path is called positive or negative according to the sign of its last link.

Two paths are said to *contradict* each other, if they have the same initial node, the same last node, and are of opposite sign. We want to be able to compose links into paths as much as possible whilst avoiding, in a principled way, the acceptance of mutually contradictory paths. Finding good principles for this is the main task of the theory of defeasible nets.

Figure 8 displays the an example in which there are two contradicting paths.

Most sophisticated procedures for managing conflict, beginning with those of Touretzky [Tou86], make use of the notion of *preclusion* (also known as pre-emption). A positive path $p_1, ..., p_{n+1}$ (negative path $p_1, ..., p_n, \neg p_{n+1}$) over a net A is said to be precluded by a pair $q_1, ..., q_m$ and $q_j \neg p_{n+1}$ (respectively, $q_j p_{n+1}$) iff $q_1, ..., q_m$ is a positive path with $q_1 = p_1$ and $q_m = p_k$ for some $k \leq n$, and for some $j < m, q_j \neg p_{n+1}$ (respectively $q_j p_{n+1}$) is a link in the net A.

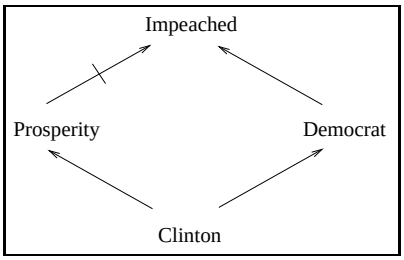

Figure 8: The modern Nixon Diamond

In the example we see in figure 8 the path $Clinton, likable, \neg Impeached$ does not preclude the path $Clinton, Democrat, Impeached$, as the j that we are seeking is not smaller than our m.

Given the notion of preclusion we can define an entailment relation $A \mid\sim_I (p, q)$ (respectively, $A \mid\sim_I (p, \neg q)$) iff there is a positive (respectively, negative) path from p to q that is acceptable (i.e., not precluded) from A. Notice, that this is a *brave* entailment relation, as we can get contradictory information (instead of taking the intersection of the "extensions").

Many other treatments and principles exist and the reader can find this material as well as an extensive account of inheritance hierarchies in [DG93b].

3.6 Summary

There are simple differences between the different systems that are worth mentioning.

- Default logic and AEL can be brave or cautious, while Circumscription is naturally cautious.
- Default logic and AEL can make nonmonotonic interferences about equality and Circumscription cannot.
- Circumscriptive statements can apply to all individuals, whereas default rules are restricted to those individuals with names in the languages. This situation is also depicted in AEL in the form of the problem of quantifying in.
- In default logic, there seems to be no way to capture the distinction between *fixed* and *variable* predicates in Circumscription.
- Proving defaults is more natural in the modal approaches and Circumscription-like than in Default Logic-like systems, as defaults are part of the language about which we can prove theorems. In Default Logic we can prove such theorems by finding all the extensions and checking for validity of the default.

- Default Logic and Autoepistemic Logic are naturally propositional, while Circumscription is naturally First-Order. Researchers went to great length to make plausible Nonmonotonic Modal Logics First-Order (Default Logic was defined for a First-Order base language) and it is still difficult and unnatural in both Default Logic and AEL/NML.

Diagram 9 describes the development of Nonmonotonic systems.

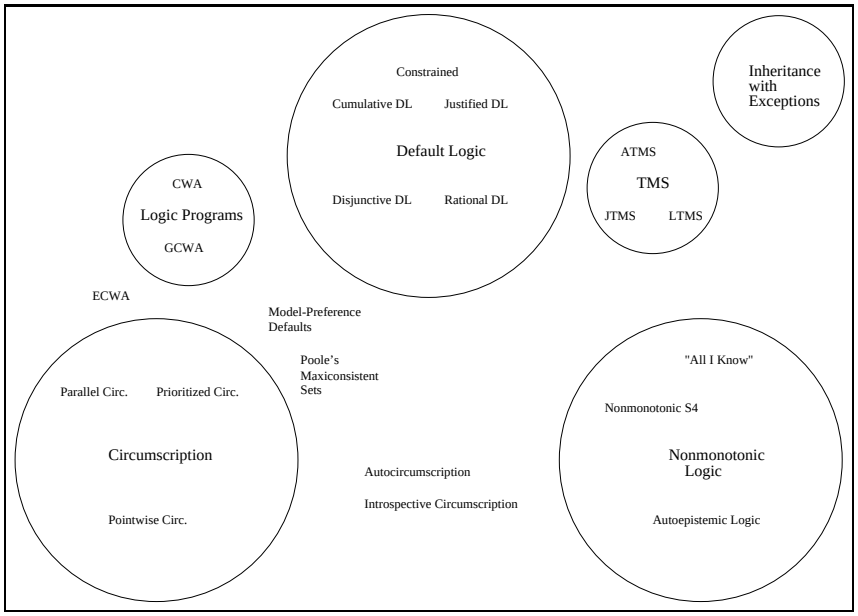

Figure 9: Nonmonotonic Reasoning Systems. Circumscription, Nonmonotonic Logic and Default Logic are shown as *attractors* in a space of nonmonotonic systems. This figure is only an illustration and has no scientific background.

4 Relating Principles: Axiomatic Properties

A central concept in classical predicate logic is that of entailment or inference. The syntactic relation $\vdash$ and the semantic relation $\models$ provide us with a possibility to determine whether a formula φ is entailed by a set of premises T. On the other hand, the nonmonotonic systems presented in section 3 are built around a somewhat different machinery. No specific entailment relations were specified (although in Default Logic one may construct entailment relations using intersections of default extensions) for each of the systems. In order to relate and compare the various systems, get a more abstract, high-level picture of nonmonotonic reasoning, and correspond more closely to the ideas of inference and entailment, we would like to introduce inference relations that we shall call

nonmonotonic. This section discusses the general concepts and results obtained in that direction.

Gabbay [Gab85] was the first to propose the study the consequence relations of the various nonmonotonic systems. He proposed three axioms that he saw to be fundamental to the nonmonotonic reasoning systems: *reflexivity*, *cut* and *cautious monotonicity*. Kraus, Lehmann and Magidor [KLM90], Makinson [Mak89], Dubois and Prade [DP88] and later Lehmann and Magidor [LM92] augmented this set of properties, now called *KLM properties*, to allow the study and characterization of different nonmonotonic systems.

In the study of different rules of inference and axioms the inference relation satisfies, one considers an inference relation $\mid\sim$ as a relation between a set of sentences and a sentence in a language $\mathcal{L}$. Typically, $\mathcal{L}$ is either propositional or first order. [KLM90] touched only the propositional case and so did most of the literature (some recent treatments of the first-order case include [FHK96]). Also, compared with the language of conditional logic (see section 5.2 on page 74), this language of conditionals ($a \mid\sim b$) is restricted to simple conditional statements (e.g., one can not write $a \mid\sim (b \mid\sim c)$).

There is a different (yet equivalent) approach to the investigation we present below, in terms of consequence operators $C(\varphi)$, and we will touch it only lightly. The reader is invited to read [Mak93] or a simpler description in [Ant97] for an overview of this methodology.

4.1 Meta-Rules and Restrictions on Inference Relations

The field of Nonmonotonic Reasoning has now got to some maturity and there are many systems and several guiding principles in preferring one system to another. The first unifying principle, by which we can start evaluating the different nonmonotonic systems, is the set of meta-rules that these systems satisfy.

In this section we explore some of the meta-rules that were cited in the literature for their significance, either in distinguishing between nonmonotonic systems or in following some common-sense examples. These rules were elicited mostly from [KLM90], [LM92] and [Mak93] but some of them date to earlier studies such as [Gab85][17].

The following are the original rule systems that were cited by [KLM90] and [LM92] to be of importance in a logical system (monotonic or not).

Cumulative Entailment Relations A consequence relation $\mid\sim$ is said to be *cumulative* iff it contains all instances of the Reflexivity axiom and is closed under the inference rules of Left Logical Equivalence, Right Weakening, Cut and Cautious Monotonicity. The set of cumulative inference relations is called *the system* **C**.

- *Reflexivity:* $\alpha \mid\sim \alpha$.

[17] We omit the conditional axioms [Lew73] from consideration here, as will be explained in what follows.

- *Left Logical Equivalence (LLE):* $\frac{\models \alpha \Longleftrightarrow \beta,\ \alpha \mid\sim \gamma}{\beta \mid\sim \gamma}$.
- *Right Weakening:* $\frac{\models \alpha \Rightarrow \beta,\ \gamma \mid\sim \alpha}{\gamma \mid\sim \beta}$.
- *Cut:* $\frac{\alpha \wedge \beta \mid\sim \gamma,\ \alpha \mid\sim \beta}{\alpha \mid\sim \gamma}$. If we prove something using an assumption and later prove our assumption, we can prove the original goal without the assumption.
- *Cautious Monotonicity:* $\frac{\alpha \mid\sim \beta,\ \alpha \mid\sim \gamma}{\alpha \wedge \beta \mid\sim \gamma}$. If we proved something, we can add it to our background knowledge without fear of retracting current results.

Loop-Cumulativity The set of cumulative inference relations that satisfy *Loop* is called *the system* **CL**. Every relation is said to be Loop-Cumulative.

$$\textit{Loop:} \qquad \frac{\alpha_0 \mid\sim \alpha_1, ..., \alpha_{k-1} \mid\sim \alpha_k,\ \alpha_k \mid\sim \alpha_0}{\alpha_0 \mid\sim \alpha_k}$$

Preferential Reasoning The system **P** consists of all the rules of **C** and the rule *OR*:

$$\textit{OR:} \qquad \frac{\alpha \mid\sim \gamma,\ \beta \mid\sim \gamma}{\alpha \vee \beta \mid\sim \gamma}$$

This rule intuitively means that one can reason by cases, then combine the cases to one theorem.

Cumulative Monotonic The system **CM** includes the rules of **C** and *Monotonicity*:

$$\textit{Monotonicity:} \qquad \frac{\models \alpha \Rightarrow \beta,\ \beta \mid\sim \gamma}{\alpha \mid\sim \gamma}$$

In other words, strengthening the premises causes no retraction of previous conclusions.

The system M The system **M** consists of all the rules of **C** and the rule of *Contraposition.*

$$\textit{Contraposition:} \qquad \frac{\alpha \mid\sim \beta}{\neg\beta \mid\sim \neg\alpha}$$

A consequence relation that satisfies all the rules of **M** is said to be *monotonic.*

Rational Reasoning The system **R** consists of all the rules of **P** and the rules of *Rational Monotonicity.*

$$\textit{Rational Monotonicity:} \qquad \frac{\alpha \wedge \beta \not\mid\sim \gamma,\ \alpha \not\mid\sim \neg\beta}{\alpha \not\mid\sim \gamma}$$

A consequence relation that satisfies all the rules of **R** is said to be *Rational.*

Theorem 4.1 ([KLM90],[LM92]) *The following are relationships between the systems mentioned above:*

$$\mathbf{C} \subsetneq \mathbf{CL} \subsetneq \mathbf{P} \subsetneq \mathbf{R} \subsetneq \mathbf{M} \tag{11}$$

$$\mathbf{CL} \subsetneq \mathbf{CM} \subsetneq \mathbf{M} \tag{12}$$

For example, that means that there are nonmonotonic entailment relations that are Preferential (in **P**) but are not Rational (in **R**). For an example of proving that a certain rule holds in a system, consider proving that Rational Monotonicity holds in **M**. Assume, by contradiction, that $\alpha \mathrel{|\!\sim} \gamma$. Then $\neg\gamma \mathrel{|\!\sim} \neg\alpha$, by Contraposition. By Right Weakening, $\neg\gamma \mathrel{|\!\sim} \neg\alpha \vee \neg\beta$. But then, using Contraposition again, $\alpha \vee \beta \mathrel{|\!\sim} \gamma$, contradicting the assumptions of Rational Monotonicity.

Some other rules were proposed in the literature ([KLM90], [LM92], [Mak93], [Mak89], [FL96], [FLM91]), and are of some interest. Here we divide them according to the different systems that satisfy them (we quote a separate theorem below).

Cumulative Properties

- *Equivalence:* $\frac{\alpha \mathrel{|\!\sim} \beta,\ \beta \mathrel{|\!\sim} \alpha,\ \alpha \mathrel{|\!\sim} \gamma}{\beta \mathrel{|\!\sim} \gamma}$.
- *And:* $\frac{\alpha \mathrel{|\!\sim} \beta,\ \alpha \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \beta \wedge \gamma}$.
- *MPC:* or Modus Ponens in the Consequent: $\frac{\alpha \mathrel{|\!\sim} \beta \Rightarrow \gamma,\ \alpha \mathrel{|\!\sim} \beta}{\alpha \mathrel{|\!\sim} \gamma}$.

Monotonic Properties

- *EHD:* or Easy Half of the Deduction theorem is $\frac{\alpha \mathrel{|\!\sim} \beta \Rightarrow \gamma}{\alpha \wedge \beta \mathrel{|\!\sim} \gamma}$.
- *Transitivity:* $\frac{\alpha \mathrel{|\!\sim} \beta,\ \beta \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \gamma}$.

Properties of P

- *Rule S:* or the *hard half of the deduction theorem* $\frac{\alpha \wedge \beta \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \beta \Rightarrow \gamma}$.
- *Rule D:* or Proof by Cases: $\frac{\alpha \wedge \neg\beta \mathrel{|\!\sim} \gamma,\ \alpha \wedge \beta \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \gamma}$.

Properties of R

- *Negation Rationality:* $\frac{\alpha \wedge \beta \not\mathrel{|\!\sim} \gamma,\ \alpha \wedge \neg\beta \not\mathrel{|\!\sim} \gamma}{\alpha \not\mathrel{|\!\sim} \gamma}$. Intuitively, that means that anything we can prove with α can be proved with either the addition of β or the addition of $\neg\beta$.
- *Disjunctive Rationality:* $\frac{\alpha \not\mathrel{|\!\sim} \gamma,\ \beta \not\mathrel{|\!\sim} \gamma}{\alpha \vee \beta \not\mathrel{|\!\sim} \gamma}$.
- *Weak Contraposition:* $\frac{\gamma \wedge \alpha \mathrel{|\!\sim} \beta,\ \gamma \not\mathrel{|\!\sim} \beta}{\gamma \wedge \neg\beta \mathrel{|\!\sim} \neg\alpha}$.

Some inter-properties of the rules defined above are

Theorem 4.2 ([KLM90],[LM92]) *1.* **C** *implies Equivalence, And and MPC.*

2. *Monotonicity, EHD and Transitivity are equivalent, in the presence of the rules of* **C**.

3. *In the presence of Left Logical Equivalence and Right Weakening, Contraposition implies Monotonicity.*

4. *In* **CL**, *the following rule holds for any* $i, j = 0, ..., k$: $\frac{\alpha_0 \mid\sim \alpha_1, ..., \alpha_{k-1} \mid\sim \alpha_k,\ \alpha_k \mid\sim \alpha_0}{\alpha_i \mid\sim \alpha_j}$.

5. **P***implies the rules S and D.*

6. *Any preferential relation that satisfies Rational Monotonicity satisfies Disjunctive Rationality (but not vice versa). Any preferential relation that satisfies Disjunctive Rationality satisfies Negation Rationality and Weak Contraposition (but not vice versa).*

7. *Negation Rationality and Weak Contraposition are incomparable. Also, even together, they do not imply Disjunctive Rationality.*

There are some other properties that we did not account for here, that are specified for the corresponding properties in Conditional Logic. Some can be found in section 5.2 on page 74 where we describe some aspects of Conditional Logics.

Some of the properties that one can ascribe to entailment relations are better understood if written in the format of consequence operators. The consequence operator C associated with an entailment relation $\mid\sim$ is defined by

$$C(A) = \{\varphi \mid \varphi \in \mathcal{L},\ A \mid\sim \varphi\}$$

Below we shall count a few properties in both their entailment relation form and their consequence operator form, where $Cn(A)$ is the consequence operator associated with the classical logical entailment.

- *Supraclassicality:* $\frac{\alpha \models \beta}{\alpha \mid\sim \beta}$ or equivalently, $C(A) \supseteq Cn(A)$. Our inference includes the classical entailment.

- *Idempotence:* there is γ such that $\alpha \mid\sim_1 \gamma$ and $C(C(A)) = C(A)$. Reapplying our inference relation to the *set of consequences*, leads to the same set of consequences[18].

- *Consistency Preservation:* $\alpha \not\vdash FALSE \implies \alpha \not\mid\sim FALSE$. If A is consistent, then so is $C(A)$.

- *Left Absorption:* $Cn(C(A)) = C(A)$. Equivalently, $\mid\sim\ = \mid\sim \circ \vdash$ (which is actually Right Weakening).

- *Right Absorption:*[19] $C(Cn(A)) = C(A)$.

[18] Notice that trying to write this property as $\mid\sim \circ \mid\sim = \mid\sim$ is not equivalent to Idempotence, but rather is equivalent to *transitivity*!

[19] Trying to write this property using entailment relations, $\vdash \circ \mid\sim\ = \mid\sim$, is not successful, and is in fact equivalent to Monotonicity.

- *Full Absorption:* $\vdash \circ \mid\sim = \mid\sim = \mid\sim \circ \vdash$. Equivalently, $C(Cn(A)) = C(A) = Cn(C(A))$.
- *Distribution (of Conjunction):* $C(A) \cap C(B) \subset C(Cn(A) \cap Cn(B))$. Equivalently, if $\alpha \mid\sim \gamma,\ \beta \mid\sim \gamma$, then there is δ such that $\alpha \vdash \delta,\ \beta \vdash \delta$ and $\delta \mid\sim \gamma$.
- *Distribution of Disjunction:* For a finite language $\mathcal{L}$, for every pair of formulas α, β, $C(\alpha \vee \beta) \subset Cn(C(\alpha) \cup C(\beta))$. Equivalently, if $\alpha \vee \beta \mid\sim \gamma$ then there are $\delta_\alpha, \delta_\beta$ such that $\alpha \mid\sim \delta_\alpha$, $\beta \mid\sim \delta_\beta$ and $\delta_\alpha \wedge \delta_\beta \vdash \gamma$.

Some inter-properties of the rules defined above are

Theorem 4.3 ([KLM90],[Mak89],[LM92],[FL96],[Fre93]) *1. Supraclassicality implies Inclusion.*

2. *OR implies Distribution.*
3. *In the presence of Supraclassicality, Idempotence implies Left Absorption, and Cumulativity implies Full Absorption.*
4. *Left Absorption implies And and Right Weakening.*
5. *Right Absorption implies LLE.*
6. *If $\mid\sim$ satisfies Distribution, Supraclassicality and Full Absorption, then it satisfies OR, D and S.*
7. *If $\mid\sim$ satisfies Distribution, Supraclassicality and Cumulativity, then it satisfies Loop and OR.*
8. *Negation rationality for a consequence operation C is equivalent to each of the following:*
   (a) *For all formulas α, β, $C(\alpha) \subseteq C(\alpha \wedge \beta) \cup C(\alpha \wedge \neg\beta)$.*
   (b) *For all formulas α, β, $C(\alpha) \subseteq C(\alpha \wedge \beta)$ or $C(\alpha) \subseteq C(\alpha \wedge \neg\beta)$.*
   (c) *For all formulas α, β, $C(\alpha \wedge \beta) = Cn(\{\beta\} \cup C(\alpha)$ or $C(\alpha \wedge \neg\beta) = Cn(\{\neg\beta\} \cup C(\alpha)$.*
9. *Disjunctive Rationality is equivalent to requiring that $C(\alpha \vee \beta) \subseteq C(\alpha) \cup C(\beta)$.*
10. *Disjunctive Rationality implies Distribution of Disjunctions but not vice versa.*

Figure 10 on the following page illustrates the relationships between the major different systems.

It is interesting to notice that originally some of the rules mentioned above were expressed as *conditional* assertions (see e.g., [KLM90]). It is a bit deceiving, but, as a conditional, $\alpha \mid\sim \beta$ does *not* say that "from α nonmonotonically entail β." It does say "Given some *conditional* background knowledge base $\mathbf{K}$, one can

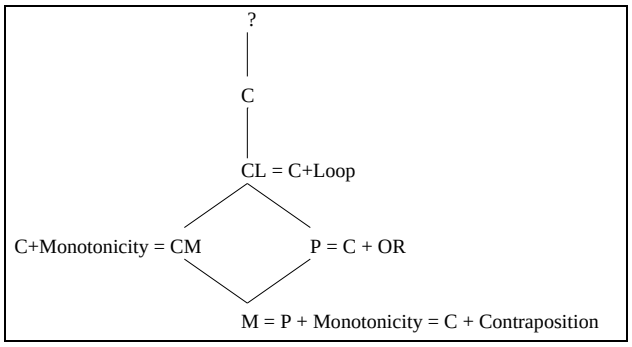

Figure 10: Relationships between rule systems

prove the conditional $\alpha \Rightarrow \beta$." The view of nonmonotonic systems using this framework is explored in section 4.2, in which the difference between conditionals and entailment relations will be further explained.

In later works, [Fla98] showed how some meta-rules systems can be reduced to some other meta-rules systems, preserving the property "*is* $\mid\sim$ *an entailment of system* **X***?*" Recently, [BMP97] offered some more non-horn rules (we define and discuss some non-horn rules in section 5.2 on page 74). In section 5, we will see how these properties are matched with semantics.

4.2 Nonmonotonic Systems as Inference Relations

In this section we describe some of the methodologies used to represent the different nonmonotonic systems presented in section 3 as inference relations. Once represented as an inference relation, these systems (with the appropriate methodology) can be examined with respect to the rules that they follow. The tricky part is choosing the inference relation that we produce using the nonmonotonic system.

4.2.1 Circumscription

In some uses of Circumscription (e.g., Satoh [Sat90], [Sat92] for the case without variables), the nonmonotonic entailment of β from α, written $\alpha \mid\sim \beta$, corresponds to

$$Circ[\alpha; P; Z] \models \beta.$$

This methodology says that, given a KB α, one first circumscribes ab in α and then draws the conclusion β. Let us call this entailment relation $\mid\sim_{C[P,Z]}$, or for short $\mid\sim_C$.

Theorem 4.4 *The entailment relation* $\mid\sim_C$ *satisfies Cumulativity and OR. It does not satisfy Monotonicity.*

PROOF Reflexivity, LLE and RW are obvious. For Cut, notice that $Circ[\alpha \wedge \beta; ab; \mathcal{L}] \models \gamma$, $Circ[\alpha; ab; \mathcal{L}] \models \beta$ means that in the minimal models of α, β is true, thus having the minimal models of $\alpha \wedge \beta$ being the same as those of α. Therefore $Circ[\alpha; ab; \mathcal{L}] \models \gamma$. For Cautious Monotonicity the same idea works. Since $\alpha \mid\sim_C \beta$, the minimal models of α are the same as those of the minimal models of $\alpha \wedge \beta$. Thus $\alpha \mid\sim \gamma \implies \alpha \wedge \beta \mid\sim \gamma$ and $\mid\sim_C$ satisfies Cumulativity.

For OR, notice that the minimal models of $\alpha \vee \beta$ are minimal models of either α or β. If a minimal model exists that is not a minimal model of either α or β, then either α or β is satisfied, and the corresponding formula has a more preferred model. Since that model also satisfies $\alpha \vee \beta$, and we have the same preference over models, we get that the given model is not minimal for $\alpha \vee \beta$. Contradiction. Thus OR holds.

Monotonicity does not hold simply because $Circ[p; ab; \mathcal{L}] \models \neg ab$ but $Circ[p \wedge ab; ab; \mathcal{L}] \models ab$ and $p \wedge ab \Rightarrow p$. ■

Thus, Circumscription defines an entailment relation $\mid\sim_C$ in the meta-rule system **P**. As Satoh [Sat90] showed, there are some cases for which $\mid\sim_C$ is rational and there are some others for which $\mid\sim_C$ is not rational. Notice, though, that the theorem is stated for *propositional* circumscription *without varied constants.*

Theorem 4.5 ([Sat90]) *Let P be a tuple of propositions in the finite language $\mathcal{L}$ (in particular, there are only finitely-many models for any (monotonic) theory).*

1. *If P does not contain all of the propositions of $\mathcal{L}$ and there is a set of axioms A such that $Circ[A; P] \not\equiv A$, then $\mid\sim_C$ is not Rational.*
2. *If P contains all the propositions of $\mathcal{L}$, which count no more than two, then $\mid\sim_C$ is Rational.*
3. *If P contains all the propositions of $\mathcal{L}$, which count more than two, then $\mid\sim_C$ is not Rational.*

One result of this theorem is that the CWA for propositional languages (section 3.5.2) is not rational, as we always minimize all the propositions, suggesting that sometimes rationality is not commonsensical. In [KLM90] it was further shown that there are cases of Circumscription that do not satisfy negation rationality. The same counter examples (in the theorem above and in [KLM90]) work for Pointwise Circumscription as well[20].

4.2.2 Default Logic

The point of view taken by meta-rule over Default Logic, is investigating the properties of each set of defaults by itself. Notice how this degree of change in default logics corresponds to the degree of change for Circumscription (choosing

[20]The author's observation.

the minimized and the varied constant symbols (and related other predicates in the different variants of Circumscription)).

For a set of defaults $\mathcal{D}$, one can describe the different properties of the induced entailment relation $\mid\sim^1_{DL[\mathcal{D}]}$, which we will call for short $\mid\sim_{\mathcal{D}}$. $\alpha \mid\sim^1_{\mathcal{D}} \beta$ iff all of the extensions[21] of $\langle\{\alpha\}, \mathcal{D}\rangle$ include β.

Theorem 4.6 ([KLM90],[Mak93]) *For a set of default rules $\mathcal{D}$, the entailment relation $\mid\sim^1_{\mathcal{D}}$ satisfies Reflexivity, LLE, Right Weakening and Cut. It also satisfies Full Absorption. It does not necessarily satisfy Monotonicity, Cautious Monotonicity (and thus Cumulativity), OR, Loop and Distributivity.*

The following example, due to Makinson, demonstrates how Default Logic does not satisfy cautious monotonicity: $T = \langle\emptyset, D\rangle$ with $D = \{\frac{:a}{a}, \frac{a\vee b:\neg a}{\neg a}$. From T we can prove both a and $a \vee b$, but after adding $a \vee b$ to T we can not prove a any more (remember that we use cautious entailment). Intuitively, the difference between having a sentence being included in the theory and being derived by the theory is that, in the former, the set of *assumptions* (the *justifications* that we use) that we make is empty, while in the latter it already includes this sentence that we just derived. This same argument makes Cautious Monotonicity fail for Constrained DL.

The concern with the failure of Cautious Monotonicity is that one cannot prove *lemmas* and then use them. Antoniou [Ant97] argues that the right way to represent lemmas is by introducing a new default rather than as a fact, since a lemma that is derived using default consequents is *defeasible* in nature, and should be represented as such.

In [Ant97] p.75 it is shown that if instead of adding the lemma to W we add it to D then the cumulativity is satisfied.

Theorem 4.7 ([Ant97] pp.82,174) *Let $T = \langle W, D\rangle$ be a default theory, and ξ a lemma of T. Also, let $\delta_\xi = \frac{:\varphi_1\wedge\ldots\wedge\varphi_n}{\xi}$ where $\varphi_1, ..., \varphi_n$ are the consequent or justification in defaults used in a minimal proof ("construction") of ξ. Then, every extension of T is an extension of $\langle W, D\cup\{\delta_\xi\}\rangle$ and vice versa. Also, every constrained extension of T is a constrained extension of $\langle W, D\cup\{\delta_\xi\}\rangle$ and vice versa.*

A consequence from theorem 3.6 on page 30 is the following:

Conclusion 4.8 *If T is a normal default theory, it satisfies consistency preservation.*

Theorem 4.9 (Dix, [Mak93]) *If $\mathcal{D}$ is a set of normal default rules with no prerequisites, then the entailment relation $\mid\sim^1_{\mathcal{D}}$ satisfies Cautious Monotonicity, Supraclassicality and Distribution.*

We define the brave entailment similarly. For a set of defaults $\mathcal{D}$, $\mid\sim^2_{DL[\mathcal{D}]}$ is the brave entailment relation. We will call for it for short $\mid\sim^2_{\mathcal{D}}$. $\alpha \mid\sim^2_{\mathcal{D}} \beta$ iff one of the extensions[22] of $\langle\{\alpha\}, \mathcal{D}\rangle$ includes β.

[21] We take the conservative entailment here.
[22] We take the brave entailment here.

Theorem 4.10 ([Mak93]) *For a set of default rules $\mathcal{D}$, the entailment relation $\mid\sim^2_{\mathcal{D}}$ does not necessarily satisfy Cut. It does satisfy Reflexivity, LLE and Right Weakening.*

As for rationality, the following observation of Makinson is mentioned in [FL96].

Proposition 4.11 *There exists a set D of prerequisite-free, normal defaults such that the associated sceptical inference relation $\mid\sim_D$ does not satisfy Negation Rationality.*

Cumulative Default Logic As hinted by its name, Cumulative Default Logic is indeed cumulative.

Theorem 4.12 ([Bre91a]) *Cumulative Default Logic is Cumulative.*

The following example demonstrates how Cumulativity holds in Brewka's theory.

Example 4.13 ([DSJ94]) *The assertional default theory $\langle \emptyset, \{\frac{:A}{A}, \frac{A\vee B:\neg A}{\neg A}\}\rangle$ has one assertional extension: $\hat{Th}(\{\langle A, \{A\}\rangle\})$. Consequently, the assertional extension contains $\langle A \vee B, \{A\}\rangle$.*

The assertional default theory $\langle \{\langle A \vee B, \{A\}\rangle\}, \{\frac{:A}{A}, \frac{A\vee B:\neg A}{\neg A}\}\rangle$ has one assertional extension: $\hat{Th}(\{\langle A, \{A\}\rangle\})$.

Notice that the reason cumulativity holds here is that we are not allowed to add the plain result $A \vee B$ to the original theory, but rather have to add the result *with its justification.* This ensures cumulativity in Brewka's system.

4.2.3 Autoepistemic Logic

For Autoepistemic Logic, we define $\alpha \vdash_{AE} \beta$ to hold iff every stable expansion A of α, satisfies $\beta \in A$.

Theorem 4.14 *$\mid\sim_{AE}$ satisfies Reflexivity, LLE, Right weakening and Cut. It does not satisfy Cautious Monotonicity.*

PROOF [KLM90] proves that $\mid\sim_{AE}$ satisfies Reflexivity, LLE, Right Weakening and Cut. To see that it does not satisfy Cautious Monotonicity, we have to find an example where $\alpha \mid\sim_{AE} \beta$, $\alpha \mid\sim_{AE} \gamma$ but $\alpha \wedge \beta \not\mid\sim_{AE} \gamma$.

Take $A = \{MP \Rightarrow P, (P \vee Q) \wedge M\neg P \Rightarrow \neg P\}$. Since $\neg P \notin A$, $\neg L\neg P \in T$ and so $MP \in T$. Thus $P \in T$. There is nothing we can conclude about Q, because we can not prove $M\neg P$ and thus $ST(P)$ is the only stable expansion for A. Thus, we can also prove $P \vee Q$. If we have $B = A \cup \{P \vee Q\}$, then B has the two stable expansions $E_1 = ST(P)$ and $E_2 = ST(Q \wedge \neg P)$. The second one is a stable expansion, because it is stable and

$$E_2 = Cn_{KD45}(B \cup \{\neg L\varphi | \varphi \notin E_2\})$$

■

A similar counter-example was given for Default Logic in [Mak93].

4.2.4 Summary

The major properties of the nonmonotonic systems presented in section 3 on page 20 are summarized in table 11.

The material in the table below and its proofs can be found in [DSJ94] (the Default Logic variants), [KLM90] (AEL, DL, Circumscription), [Mak93] (DL, Circumscription), [Kon93] (AEL translations). Some of the properties (such as for Pointwise Circumscription, or different properties of the different systems) that were not explicitly stated in the literature but were easily provable (as consequences of other theorems (in this section and sections 5, 6)) were put in by the author himself without a proof[23].

		Default Logic	DL (brave)	Constrained DL	Justified DL	Cumulative DL	Circumscription	Pointwise Circ.	Autoepistemic Logic	Moderately Grounded AEL	Strongly Grounded AEL
Cumulativity	Reflexivity	+	+	+	+	+	+	+	+	=	+
	LLE	+	+	+	+	+	+	+	=	=	+
	Right Weakening	+	+	+	+	+	+	+	+	=	+
	Cut	+	—	+	+	+	+	+	+	=	+
	Cautious Monotonic.	PfN	=	Pf	?	+	S	S	—	=	—
Pref.	Loop	PfN	=	PfN*	PfN*	PfN*	+	+	=	=	—
	OR	PfN	=	PfN*	PfN*	PfN*	+	+	=	=	—
	D	Pf	=	Pf	Pf	Pf	+	+	sim	sim	sim
	Consistency Preserv.	N	=	+	+	+	S	S	—	—	—
Rat.	Negation Rationality	—	=	—	—	—	—	—	=	=	—
	Disjunctive Rat.	—	=	—	—	—	—	—	=	=	—
	Rat. Monotonicity	—	=	—	—	—	—	—	=	=	—
	Full Absorption	+	=	+	+	+	S	S	=	=	=
	Distribution	PfN	=	PfN*	PfN*	PfN*	+	+	=	=	=
	Supraclassicality	+	+	+	+	+	+	+	=	=	=

Figure 11: Meta-rules obeyed by the different nonmonotonic systems.

The Legend for this table is as follows:

+	Always holds.	N	Normal default theories.
—	Does not hold.	Pf	Prerequisite free defaults.
?	Open Problem.	PfN	Prerequisite free normal defaults.
=	Not known	S	Smooth preference relation[24]
*	at least that class	sim	Simple normal form theories

By "Not known" we mean that the author could not find any theorem stating the property and no proof was immediate and trivial. Open Problems are problems that were declared as such by other authors.

[23]The author made an effort not to put anything, that is not an "obvious" conclusion of current theorems, without a proof.

[24]Smoothness is defined in section 5.1, but for now we can use the similar requirement of

All of the properties of Default Logic systems in table 11 refer to the cautious entailment relation ($\mid\sim^1_{\mathcal{D}}$), besides the one explicitly mentioned as brave entailment.

5 Relating Principles: Nonmonotonic Reasoning Semantics

In this section we consider the different semantics proposed for nonmonotonic systems. Some of these semantics come to propose a new nonmonotonic entailment relation, defined using the semantics (we already saw some examples with "All-I-Know" (section 3.3.3) and Model Preference Default Theories (section 3.4.3)). Nevertheless, the majority of semantics came as a result of trying to give semantics to existing systems and trying to unify the existing systems under a single semantics.

We will first describe the different general semantics devised in the literature, and then show the specific semantics given to the nonmonotonic systems we considered in section 3. We reserve considering the translations between the different semantics to section 6 on page 85.

5.1 Preferential Semantics

By Preferential Semantics we refer to all semantics that have structures combined of a set of worlds W, an interpretation function i on these worlds (and the underlying language) and a relation defining order between the worlds $\leq_W$, which define their entailment relation by observing those worlds that are the most preferred (according to the relation $\leq_W$).

Intuitively, the idea is that we can model a nonmonotonic entailment relation by fixing a preferential structure for that nonmonotonic relation. Nonmonotonicity is then simulated by changing the preferred worlds according to our axioms (what we know). The preference relation determines for us the most preferred worlds (which may or may not exist, depending, among others, on the specific preferential semantics and its requirements), and we say that for a preference model W, $\alpha \mid\sim_W \beta$ iff for every world in which α is true, and that world is minimal satisfying α with respect to $\leq_W$, β is true (the expanded treatment for conditionals is described in section 5.2.1).

There are works as early as Lewis [Lew73], describing preferential semantics for conditional logics. Shoham [Sho87] introduced the preferential models semantics for nonmonotonic reasoning systems and argued that it is capable of giving semantics for various nonmonotonic systems. Kraus et al [KLM90] expanded on Shoham's work by showing how different conditions put on somewhat generalized preferential-models semantics can correspond to different meta-rules that the entailment relation should satisfy.

consistency preservation.

In what follows, we shall show many *Representation Theorems* for some of the systems of meta-rules we defined in section 4.1 on page 60. A Representation Theorem is a theorem stating that a model in the given representation satisfies the axioms of the system (soundness) and that every entailment relation that satisfies these axioms has a model in the given representation that defines that entailment relation (completeness).

5.1.1 Cumulative Models

Kraus, Lehmann and Magidor [KLM90] defined a structure, which they call a *Cumulative Model*, that allowed them to give semantics to all cumulative entailment systems,

A model essentially consists of a set of states and a binary relation on those states. In addition, every state is labeled with a set of worlds (worlds here are classic base-language structures), which determine the formulas true in that state (thus a state has only a syntactic value by itself). This scheme allows us to have "duplicate" worlds (states labeled with the same set of worlds) in different positions in the partial order $\prec$.

In the following $\mathcal{U}$ stands for the set of classic, base-language structures.

Definition 5.1 *For a state s and a labeling function $l : S \mapsto 2^{\mathcal{U}}$ that labels s with a nonempty set of worlds $l(s)$, we say that s* satisfies a formula α*, writing $s \models \alpha$, if $\forall m \in l(s).(m \models \alpha)$.*

Definition 5.2 (Smoothness) *We say that a set P is* smooth[25] *w.r.t. a binary relation $\prec \subset P \times P$, if for all $t \in P$, t is $\prec$-minimal or there is $s \prec t$ that is minimal in P.*

Notice that the condition of smoothness is weaker than well-foundedness (no infinite descending chains). Not every smooth set is also well-founded[26].

Definition 5.3 (Cumulative Model) *A* Cumulative Model *is a triple $\langle S, l, \prec \rangle$, where S is a set of elements, $l : S \mapsto 2^{\mathcal{U}}$ is a function that labels every state with a nonempty set of worlds and $\prec$ is a binary relation on S. In addition, we require that, for every formula $\alpha \in \mathcal{L}$, the set of states that satisfy α is smooth (this is sometimes called the* smoothness condition*).*

Notice that $\prec$ is not required to be a partial order. The smoothness condition is needed to ensure the validity of Cautious Monotonicity (for theorem 5.4 below).

To define a consequence relation, given a cumulative model $W = \langle S, l, \prec \rangle$, the consequence relation $\mid\sim_W$ defined by W is defined by

$$\alpha \mid\sim_W \beta \iff \forall s(s \text{ is } \prec\text{-minimal such that } s \models \alpha \implies s \models \beta)$$

Theorem 5.4 (Representation Theorem for Cumulative Relations) *A consequence relation is a cumulative consequence relation (i.e., in the system* **C***) iff it is defined by some cumulative model.*

[25] Makinson [Mak93] calls this property *stoppered*.

[26] In some papers, such as [Lif86a] and [EMR85], well-foundedness did refer to this property.

An interesting subset of the cumulative models are *cumulative ordered models.*

Definition 5.5 *A* cumulative ordered model *is a cumulative model in which the relation $\prec$ is a strict partial order.*

Theorem 5.6 (Representation Theorem for Loop-Cumulative Relations) *A consequence relation is a loop-cumulative relation (i.e., in the system* **CL***) iff it is defined by some cumulative ordered model.*

Another interesting set of cumulative models are the *simple cumulative models.*

Definition 5.7 *A cumulative model will be called a* simple cumulative model *iff the binary relation $\prec$ on its states is empty.*

Theorem 5.8 (Representation Theorem for Cumulative Monotonic Relations) *A consequence relation is cumulative monotonic (i.e., in the system* **CM***) iff it is defined by some simple cumulative model.*

5.1.2 Preferential Models

Although *Preferential Models* can be seen as a restricted form of Cumulative Models, variants their-of have been studied much before [KLM90]. The first to introduce them was Lewis [Lew73] for his Conditional Logic. In [Sho87], Shoham introduced a form of preferential models to give semantics to various nonmonotonic systems. The version we present here was devised by Kraus, Lehmann and Magidor [KLM90] and is a slight generalization of Shoham's notion, that allowed them to prove a representation theorem for the system of meta-rules **P**.

Definition 5.9 (Preferential Model) *A preferential model W is a triple $\langle S, l, \prec \rangle$, where S is a set, the elements of which will be called states, $l : S \mapsto \mathcal{U}$ assigns a world to each state and $\prec$ is a strict partial order on S (i.e., an irreflexive, transitive, relation), satisfying the smoothness condition.*

An equivalent definition is that a preferential model is a cumulative ordered model in which states are labeled by singletons (instead of sets of possibly multiple worlds).

Theorem 5.10 (Representation Theorem for Preferential Relations) *A consequence relation is a preferential consequence relation (i.e., in the system* **P***) iff it is defined by some preferential model.*

This representation theorem holds even if we allow $\prec$ to be an arbitrary binary relation, and [KLM90] say that the reason $\prec$ is required to be a strict partial order is only for esthetics of the models and ease of checking smoothness in them.

The semantics proposed by Shoham [Sho87] is different from definition 5.9 in that he essentially required that the mapping l be *injective* (i.e., every state can show only once in the order) requiring only that the relation $\prec$ be a strict partial order. Without smoothness one cannot guarantee cautious monotonicity, and so the representation theorem does not hold for his preferential models. Also, it was showed by Kraus et al that accepting a stronger condition of well-foundedness on the order relation will deny us the representation theorem (its completeness side).

Bossu and Siegel [BS85] displayed a similar approach to that of Shoham, requiring smoothness (but keeping the injectiveness of the model). These models allow us to get another representation theorem.

Theorem 5.11 ([Fre93]) *Let $\mathcal{L}$ be a logically finite language and $\mid\sim$ a preferential inference relation on $\mathcal{L}$. Then $\mid\sim$ is represented by an injective model iff $\mid\sim$ satisfies Distribution of Disjunction*[27].

If we assume that the preference relation is empty, then we get a monotonic system in **M**.

Definition 5.12 *A* simple preferential model *is a preferential model in which the binary relation $\prec$ is empty.*

Theorem 5.13 (Representation Theorem for Monotonic Relations) *A consequence relation is monotonic iff it is defined by some simple preferential model.*

In [LM92], Lehmann and Magidor proposed another set of preferential models, modeling the meta-rules of rationality (see section 4.1).

Definition 5.14 (Modular Partial Orders) *Let $\prec$ be a partial order*[28] *over a set V. We say that $\prec$ is* modular *if for every $x, y, z \in V$, $x \prec y \implies z \prec y \vee x \prec z$.*

[LM92] gives two other equivalent conditions for modularity.

Definition 5.15 (Ranked Models) *A* ranked model *W is a preferential model $\langle S, l, \prec \rangle$ for which the strict partial order $\prec$ is modular.*

Theorem 5.16 (Representation Theorem for Rational Relations) *A binary relation $\mid\sim$ on $\mathcal{L}$ is a rational consequence relation iff it is the consequence relation defined by some ranked model.*

For preferential entailment relations that satisfy only Negation Rationality, Freund [Fre93] gives a representation theorem using models that are slight generalizations of ranked models.

[27] See section 4.1 on page 64 for the definition of Distribution of Disjunction.

[28] Or a strict partial order.

Definition 5.17 (Filtered Models) *A preferential model $W = \langle S, l, \prec \rangle$ is said to be* filtered *if whenever two states $s, t \in S$ satisfy a formula α without being minimal satisfying α, there exists a state $r \in S$, such that $r \prec s$, $r \prec t$ and $r \models \alpha$.*

Ranked models are easily seen to be also filtered.

Theorem 5.18 (Representation theorem for P+NR) *Let $\mathcal{L}$ be a finite language. A preferential relation is disjunctive (i.e., satisfies* **P** *and Negation Rationality) iff it can be defined by an injective filtered model.*

Freund actually gave his representation theorem for the general case of $\mathcal{L}$, but replacing *injective* by *standard* (a notion we do not cover here).

Among the nonmonotonic systems that have preferential semantics are Circumscription [McC80], Pointwise Circumscription [Lif87b], Ground S5 [HM85], Minimal Temporal Epistemic Logic [ET96] and Poole's Maxiconsistents sets [Poo88b].

5.2 Conditional Logic and Plausibility Measures

Conditionals, such as "if I was to hit a match, it would have lit", existed long before the computer age, but Conditional Logics were not introduced until recently. In 1973, Lewis [Lew73] proposed conditionals to capture counterfactuals. The syntax of Conditional Logic is similar to that of propositional logic (or First-order logic, respectively) with the additional connective $\rightarrow$. That is, if φ and ψ are formulas in the language, so is $\phi \rightarrow \psi$. Given some rules of inference ρ, the notation $\Delta \vdash_\rho \phi \rightarrow \psi$ denotes that $\phi \rightarrow \psi$ can be deduced from Δ using these rules.

Many semantics have been proposed in the literature for conditionals. Most of them involve preference ordering [Lew73],[KLM90] and probabilistic models [Ada75], [GMP93], [Sat92]. In this section we describe these semantic accounts (preferential semantics is described in section 5.1 on page 70 but we expand here from a conditional logic point of view) and recount the way *Plausibility Measures* [FH96b],[FH98] generalizes them into one framework.

The literature on Conditional Logic is rich and wide and we do not intend to give a comprehensive survey of results but rather mention those aspects that are more important and relevant to research in Nonmonotonic Reasoning. Some of the results in this section may be read as results for both conditional logics and nonmonotonic entailment relations, while other results are stated specifically for conditional logics. In the latter case, we shall explain the meaning of these results for nonmonotonic entailment relations. The most important thing to notice, at this time, is that a conditional entailment relation corresponds to a *class* of nonmonotonic entailment relations. This point is especially sharpened given the fact that conditional logic is *monotonic*. More on these points in section 5.2.6 on page 79.

In the following descriptions, we use the structure (W, X, π) (together with different other objects) as a model of the world. In this structure, W is a set of

possible worlds, $\pi(\omega)$ is a truth assignment to primitive propositions ($\omega \in W$) and X is some "measure" on W. Given such a structure, $[\![\varphi]\!] \subset W$ is the set of worlds satisfying φ.

5.2.1 Preferential Entailment from a Conditional Logic Perspective

With respect to preferential entailment (section 5.1), in conditional logic there is the extra case of proving $\Delta \models \phi \Rightarrow \psi$, where Δ is a conditional knowledge base (when considering nonmonotonic entailment relations $\mid\sim$, there is only the consideration of $\phi \mid\sim \psi$). A knowledge base Δ *preferentially entails* $\phi \Rightarrow \psi$, denoted $\Delta \models_p \phi \Rightarrow \psi$, if every preferential structure that satisfies (all the defaults in) Δ also satisfies $\phi \Rightarrow \psi$.

Recall the definitions of the rule systems **P** and **CL** from section 4.1.

Theorem 5.19 ([KLM90]) *If a conditional knowledge base Δ contains only horn assertions (i.e., of the form $p_1 \wedge ... \wedge p_n \mid\sim q$, where $p_1, ..., p_n, q$ are propositional symbols), and $\mathcal{A}$ is a horn assertion, then $\Delta \mid\sim_{\mathbf{P}} \mathcal{A} \implies \Delta \mid\sim_{\mathbf{CL}} \mathcal{A}$.*

Theorem 5.20 ([LM92],[Bou94a]) *For a conditional knowledge base Δ, $\Delta \models_p \phi \rightarrow \psi$ (preferential entailment) iff $\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi$ (the system of rules **P**).*

This means that if we take the conjunction of entailments true in all the models of Δ, then this set of formulas is exactly those that are proved using the rules of **P** from Δ.

The following theorem emphasizes the difference between conditional logic interpretation of such theorems and that of nonmonotonic entailment relations. We say that a conditional knowledge base Δ *rationally entails* $\phi \rightarrow \psi$, denoted $\Delta \models_r \phi \rightarrow \psi$, if every rational structure (ranked model (definition 5.15 on page 73)) that satisfies Δ also satisfies $\phi \rightarrow \psi$.

Theorem 5.21 ([LM92]) $\Delta \models_r \phi \rightarrow \psi$ *iff* $\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi$.

This is somewhat surprising, since we know (from theorem 5.16 and theorem 4.2) that every ranked model satisfies *rational monotonicity* while there are some preferential models that do not.

Friedman and Halpern [FH98] propose the explanation that, although Rational Monotonicity is an additional property satisfied by rational structures, it is not one that is expressible in the language of defaults (because negated defaults are not allowed). A different explanation is given by Lehmann and Magidor [LM92] that note that the intersection of rational entailment relations may fail to be rational.

5.2.2 Parameterized Probabilistic Distribution and ϵ-Semantics

Pearl [Pea89] considered a semantics for conditionals grounded in probability, using an approach due to Adams [Ada75].Pearl's ϵ-semantics was naturally induced by a *class* of preferential models [LM92], but it had no direct counterpart

to the notion of an *individual* preferential model (which is what corresponds to a nonmonotonic entailment relation (see sections 5.2.6, 5.1)).

As a reformulation, Goldszmidt Morris and Pearl [GMP90],[GMP93], introduced PPD structures. A *parameterized probabilistic distribution* (PPD) on W is a sequence $\{Pr_i : i \geq 0\}$ of probability measures over W. A *PPD structure* is a tuple $(W, \{Pr_i : i \geq 0\}, \pi)$, where $\{Pr_i\}$ is a PPD over W. It satisfies the conditional $\varphi \rightarrow \psi$ if $lim_{i \rightarrow \infty} Pr_i([\![\psi]\!]|[\![\varphi]\!]) = 1$ (where $Pr_i([\![\psi]\!]|[\![\varphi]\!])$ is taken to be 1 if $Pr_i([\![\varphi]\!]) = 0$).

Δ ϵ-entails $\phi \rightarrow \psi$, denoted $\Delta \models_\epsilon \phi \rightarrow \psi$, if every PPD structure that satisfies all the defaults in Δ also satisfies $\phi \rightarrow \psi$.

Theorem 5.22 ([Gef92b],[GMP93]) *Let Δ be a conditional knowledge base. $\Delta \models_\epsilon \phi \rightarrow \psi$ iff $\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi$*

5.2.3 Possibility Measures

[DP86], [DP88] discuss Possibility Measures and their uses in reasoning about uncertainty and defaults. The following representation was presented in [DP91].

$Poss$ is a function $Poss : 2^W \mapsto [0, 1]$ such that $Poss(W) = 1$, $Poss(\emptyset) = 0$, and $Poss(A) = sup_{\omega \in \mathcal{A}}(Poss(\{\omega\})$. A *possibility structure* is a tuple $(W, Poss, \pi)$, where $Poss$ is a possibility measure on W. It satisfies a conditional $\phi \rightarrow \psi$, written $(W, Poss, \pi) \models_{Poss} \phi \rightarrow \psi$ if either $Poss([\![\varphi]\!]) = 0$ or $Poss([\![\varphi \wedge \psi]\!]) > Poss([\![\varphi \wedge \neg\psi]\!])$. That is, either φ is impossible, in which case the conditional holds vacuously, or $\varphi \wedge \psi$ is more possible that $\varphi \wedge \neg\psi$.

For example, $Bird \rightarrow Fly$ is satisfies when $Bird \wedge Fly$ is more "possible" than $Bird \wedge \neg Fly$. In Conditional Logic terms, $\Delta \models_{Poss} Bird \rightarrow Fly$ iff all the models that agree on the conditionals of Δ agree that the "Possibility" associated with $Bird \wedge Fly$ is larger than that associated with $Bird \wedge \neg Fly$.

Theorem 5.23 ([DP91]) *Let Δ be a conditional knowledge base. $\Delta \models_{Poss} \phi \rightarrow \psi$ iff $\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi$*

5.2.4 κ-Rankings

Defined by [GP92a], and based on ideas that go back to [Spo87], κ-ranking (or *ordinal ranking*) on W is a function $\kappa : 2^W \mapsto \mathcal{N} \cup \{\infty\}$ such that $\kappa(W) = 0$, $\kappa(\emptyset) = \infty$ and $\kappa(A) = min_{\omega \in A}(\kappa(\{\omega\}))$. Intuitively, an ordinal ranking assigns a degree of surprise to each subset of worlds in W, where 0 means unsurprising and higher numbers denote greater surprise. A *κ-structure* is a tuple (W, κ, π) where κ is an ordinal ranking on W. It satisfies a conditional $\varphi \rightarrow \psi$ if either $\kappa([\![\varphi]\!]) = \infty$ or $\kappa([\![\varphi \wedge \psi]\!]) < \kappa([\![\varphi \wedge \neg\psi]\!])$.

For example, $Bird \rightarrow Fly$ if the ordinal associated with $Bird \wedge Fly$ is smaller than that associated with $Bird \wedge \neg Fly$. In Conditional Logic terms, $\Delta \models_\kappa Bird \rightarrow Fly$ iff all the models that agree on the conditionals of Δ agree that the ordinal associated with $Bird \wedge Fly$ is smaller than that associated with $Bird \wedge \neg Fly$.

Theorem 5.24 ([Gef92b]) *Let Δ be a conditional knowledge base. $\Delta \models_\kappa \phi \rightarrow \psi$ iff $\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi$*

5.2.5 Plausibility Measures

A *plausibility measure* [FH96b], [FH98] Pl on W is a function that maps subsets of W to elements in some arbitrary partially ordered set. $Pl(A)$ reads the "plausibility set of A". If $Pl(A) \leq Pl(B)$, then B is at least as plausible as A. Formally a *plausibility space* is a tuple $S = (W, \mathcal{F}, Pl)$, where W is a set of worlds, $\mathcal{F}$ is an algebra on "measurable" (by Pl) subsets of W, and Pl maps elements of $\mathcal{F}$ to some set D, partially ordered by a relation $\leq$ (so that $\leq$ is reflexive, transitive, and anti-symmetric). As usual, the strict ordering $<$ is defined by taking $d_1 < d_2$ if $d_1 \leq d_2$ and $d_1 \neq d_2$. The set D is assumed to be *pointed*, that is, it contains two special elements $\top$ and $\bot$ such that $\bot \leq d \leq \top$ for all $d \in D$. We further assumer that $Pl(W) = \top$ and $Pl(\emptyset) = \bot$.

Finally, the following requirement is added to the description of plausibility spaces. It says that a set is at least as plausible as its subsets.

A1. If $A \subset B$, then $Pl(A) \leq Pl(B)$.

For the semantics for conditionals, a *plausibility structure* is a tuple $PL = (W, \mathcal{F}, Pl, \pi)$, where $(W, \mathcal{F}, Pl)$ is a plausibility space and π a truth assignment for the propositions of $\mathcal{L}$ in each of the worlds in W. For the language $\mathcal{L}$ (propositional!) we take $[\![\varphi]\!]$ to represent the set of worlds in W that satisfy φ, and require that $\forall \varphi \in \mathcal{L}.[\![\varphi]\!] \in \mathcal{F}$. This way, all of the formulas in $\mathcal{L}$ have a plausibility value in our plausibility space. We then define,

$$PL \models \varphi \rightarrow \psi \text{ if either } Pl([\![\varphi]\!]) = \bot \text{ or } Pl([\![\varphi \wedge \psi]\!]) > Pl([\![\varphi \wedge \neg\psi]\!])$$

Intuitively, $\varphi \rightarrow \psi$ holds vacuously if φ is impossible. Otherwise, it holds if $\varphi \wedge \psi$ is more plausible than $\varphi \wedge \neg\psi$.

It was shown in [FH96b], [FH98] that plausibility spaces and measures generalize probability spaces, possibility measures, κ-ranking, preference orderings, parameterized probability distributions and *belief functions* ([Sha76], [Sha90a], [Sha90b]).

The following describe two of the mappings given by [FH98].

PPDs For a PPD $PP = (W, \{Pr_i\}, \pi)$, $PP \models \varphi$ iff $(W, Pl_{PP}, \pi) \models \varphi$, where Pl_{PP} is a plausibility measure on W such that

$$Pl_{PP}(A) \leq Pl_{PP}(B) \iff lim_{i \rightarrow \infty} Pr_i(B|A \cup B) = 1$$

Preferential Structures For a preferential structure $(W, \prec, \pi)$, $(W, \prec, \pi) \models \varphi$ iff $(W, Pl_\prec, \pi) \models \varphi$, where

$$Pl_\prec(A) \leq Pl_\prec(B) \iff \forall w \in A \setminus B \exists w' \in B.(w' \prec W \wedge \neg\exists w'' \in A \setminus B.w'' \prec w')$$

The domain D for plausibility values consists of one element d_w for each element $w \in W$, complemented with least upper bounds for each subset of D.

It was shown that there are Plausibility Structures that do not obey the rules Cautious Monotonicity, Cut and OR. For the other properties of **P** (which are also properties of **C**), namely, Reflexivity, Right Weakening and LLE, we have the following theorem.

Theorem 5.25 ([FH98]) *Let* $\mathbf{C}'$ *be the system consisting of the rules REF, RW and LLE, and* $\mathcal{S}^{PL}$ *denote the* class *of Plausibility Structures. If* $\Delta \vdash_{\mathbf{C}'} \phi \rightarrow \psi$, *then* $\Delta \models_{\mathcal{S}^{PL}} \phi \rightarrow \psi$.

To satisfy the KLM properties [KLM90], we must restrict our view to plausibility structures that satisfy the following axioms:

A2. If A, B, C are pairwise disjoint sets, $Pl(A \cup B) > Pl(C)$ and $Pl(A \cup C) > Pl(B)$, then $Pl(A) > Pl(B \cup C)$.

A3. If $Pl(A) = Pl(B) = \bot$, then $Pl(A \cup B) = \bot$.

In this case, we call the plausibility space $(W, \mathcal{F}, Pl)$ *qualitative*. In [FH96b] it was shown that the KLM properties are sound with respect to a class of plausibility structures iff the class consists of qualitative plausibility structures.

Theorem 5.26 ([FH98]) *Let* $\mathcal{S}^{QPL}$ *be the class of all qualitative plausibility structures and* $\mathcal{S}$ *an arbitrary class of plausibility structures.* $\mathcal{S} \subset \mathcal{S}^{QPL}$ *iff for all* Δ, ϕ, ψ, *if* $\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi$ *then* $\Delta \models_{\mathcal{S}} \phi \rightarrow \psi$.

Let $\mathcal{S}^{Poss}, \mathcal{S}^{\kappa}, \mathcal{S}^{p}, \mathcal{S}^{r}, \mathcal{S}^{\epsilon}, \mathcal{S}^{QPL}$ be the sets of plausibility structures corresponding to possibility structures, κ-ranking, preferential models, rational models (ranked models), ϵ-semantics and qualitative plausibility structures.

Theorem 5.27 ([FH96b]) *Each of* $\mathcal{S}^{Poss}, \mathcal{S}^{\kappa}, \mathcal{S}^{p}, \mathcal{S}^{r}, \mathcal{S}^{\epsilon}$ *is a subset of* $\mathcal{S}^{QPL}$.

Thus, for every model in the above systems, **P** is sound.

Theorem 5.28 ([FH96b]) *For* $\mathcal{S} \in \{\mathcal{S}^{Poss}, \mathcal{S}^{\kappa}, \mathcal{S}^{p}, \mathcal{S}^{r}, \mathcal{S}^{\epsilon}, \mathcal{S}^{QPL}\}$, *and all* Δ, ϕ, ψ,

$$\Delta \vdash_{\mathbf{P}} \phi \rightarrow \psi \iff \Delta \models_{\mathcal{S}} \phi \rightarrow \psi$$

One of the conclusions that Friedman and Halpern draw from their investigation is that to go beyond the system **P** of meta-rules, one should either consider interesting *non-rich*[29] classes of structures or to define a notion of entailment from Δ that does not consider all the structures of a given class. They comment that no work known to them is of the first kind, but many works are of the second (such as [Gef92a], [GP92b], [GMP93], [LM92], [Pea90], [BGHK93]).

With respect to Plausibility structures themselves, there are also earlier similar structures that were proposed by Weber [Web91] (Uncertainty Measures) and Greco [Gre87] (L-fuzzy measures). A comparison is given in [FH98]. [FHK96] discussed First-order extensions for Conditional Logic and showed that Plausibility structures provide semantics for it as well. This is an area of current research.

[29] The property of *richness* basically says that we do not have *a priori* constraints on the relative plausibilities of a collection of disjoint sets (thus any set of mutually exclusive formulas can be ordered by plausibilities in all possible ways).

5.2.6 Conditional Logic and Nonmonotonic Inference Relations

The typical approach to the use of conditionals for representing nonmonotonic inference relations says that the conditional operator $\rightarrow$ is the described nonmonotonic inference relation $\mid\sim$. Nevertheless, there is a difference between nonmonotonic entailment and conditional entailment. The first is interested in proving base-language[30] formulas from base-language knowledge base. The second is interested in proving conditional formulas (e.g., $\alpha \rightarrow \beta$) from conditional knowledge bases. Because of this difference, we shall (in this section) refer to the conditional operators as $\rightarrow$ and the nonmonotonic inference relation as $\mid\sim$.

There were many works on the relationships between conditional logics and nonmonotonic reasoning, including [Nej91], [dCHL94], [CL92] and [FH96a]. These comparisons take the stance that a nonmonotonic entailment relation is an instance of a conditional knowledge base but none actually addresses any of the nonmonotonic systems we discussed in section 3 on page 20.

While the conditional setting includes a (conditional) knowledge base $\mathbf{K}$, a premise α and a conclusion β, some nonmonotonic systems contain only the two components α and β. Default Logic is the exception here, as for a default theory $\langle W, D\rangle$ there is a clear distinction between the conditional knowledge base D, the premise W and the conclusion (defined according to the choice of caution or braveness).

To make the discussion more uniform, we shall resort to the entailment relations defined in section 4.2. It seems that there are two main approaches that one can take in representing nonmonotonic entailment relations using conditional logics.

1. For $\alpha \mid\sim \beta$, we write $\mathbf{K} \wedge \alpha \Rightarrow \beta$, i.e., unify the knowledge base and the premise. This is particularly useful if we wish to represent Circumscription or Nonmonotonic Logic in a conditional setting. A variant of this approach is to decide that $\mathbf{K}$ has a different language from that of α, creating a natural separation between the two. For example, $\mathbf{K}$ is the only part that can hold sentences with the predicate Ab (for "abnormal"), when using Circumscription. Another example restriction is for $\mathbf{K}$ to be the only part that can hold modalities, when using Autoepistemic Logic (Thanks to Daniel Lehmann for the proposal of this variant).

2. For each one of the systems, we identify a set of conditional statements that determine its behavior, conditionals that we will collect into $\mathbf{K}$. Then, $\alpha \mid\sim \beta$ corresponds exactly to $\mathbf{K} \models \alpha \rightarrow \beta$. Along this direction (although seemingly not intentionally), are the works of Satoh [Sat90] and Freund [Fre93], getting closer to representing Circumscription by means of conditional axioms.

[30] Propositional, First-order or other languages that serve as the basis on which the nonmonotonic inference "works".

5.2.7 Summary

The major properties of the the different semantics presented in sections 5.2, 5.1.2, 5.4 are summarized in table 12.

The material in the table below and its proofs can be found in [KLM90], [LM92], [Mak93], [Kon93] and [FH98]. Some of the properties (such as the rule D for preferential models) that were not explicitly stated in the literature but were easily provable (as consequences of other theorems (in this section and sections 5, 6)) were put in by the author himself without a proof[31].

	Reflexivity	LLE	Right Weakening	Cut	Cautious Monotonicity	Loop	OR	D	Consistency Preservation	Negation Rationality	Disjunctive Rationality	Rational Monotonicity	Monotonicity
Cumulative Models	+	+	+	+	+	—	—	—	+	—	—	—	—
Cumulative Ordered	+	+	+	+	+	+	—	=	+	—	—	—	—
Simple Cumulative	+	+	+	+	+	+	—	=	+	—	—	—	+
Preferential Models	+	+	+	+	+	+	+	+	+	–	–	–	–
Simple Preferential	+	+	+	+	+	+	+	+	+	+	+	+	+
Ranked Pref. Models	+	+	+	+	+	+	+	+	+	+	+	+	–
PPD model	+	+	+	+	+	+	+	+	=	=	=	=	—
Possibility structure	+	+	+	+	+	+	+	+	=	=	=	=	—
κ-structure	+	+	+	+	+	+	+	+	=	=	=	=	—
Plausibility structures	=	=	=	–	–	–	–	=	=	=	=	=	—
Qualitative Pl. structures	+	+	+	+	+	+	+	+	=	=	=	=	—

Figure 12: Meta-rules obeyed by the different semantics.

The Legend for this table is as follows:

+	Always holds.
–	Does not hold.
?	Open Problem.
=	Not Known.

By "Not known" we mean that the author could not find any theorem stating the property and no proof was immediate and trivial. Open Problems are problems that were declared as such by other authors.

5.3 Semantics for the different nonmonotonic systems

Here we show the specific semantics given to the systems we described earlier in section 3.

[31] The author made an effort not to put anything, that is not an "obvious" conclusion of current theorems, without a proof.

5.3.1 Circumscription

Given by Lifschitz [Lif85b], McCarthy [McC86] and Etherington [Eth86], the preferential semantics for Circumscription was the first such semantics discovered.

Definition 5.29 ([Lif85b]) *For any two models M and N of a theory T we write $M \leq_{P,Z} N$ if the models M, N differ only in how they interpret predicates from P and Z and if the extension of every predicate from P in M is a subset of its extension in N.*

The relation $\leq_{P,Z}$ is a pre-order and hence we can talk about minimal models M with respect to $\leq_{P,Z}$ in the class of all models of T. More precisely, we say that a model M of T is (P,Z)-*minimal* if there is no model N such that $N <_{P,Z} M$.

Theorem 5.30 ([Lif85b]: Semantics for Circumscription) *A structure M is a model of $Circ[T;P;Z]$ iff M is a (P,Z)-minimal model of T.*

A similar theorem is given with respect to the semantics for Prioritized Circumscription.

Theorem 5.31 ([Lif85b]: Semantics for Prioritized Circumscription) *A structure M is a model of the prioritized circumscription $Circ[T;P_1 > ... > P_n;Z]$ iff M is a model of T such that for every $i \leq n$, M is $(P_i, \{P_{i+1}, ..., P_n, Z\})$-minimal model of T.*

This last theorem is also a direct consequence of the previous theorem and theorem 3.1 on page 23.

5.3.2 Pointwise Circumscription

In [Lif87b], Pointwise Circumscription was given semantic interpretation.

Take, without loss of generality, $i_0 = 1$. For a model $\mathcal{M}$ of $A(S)$, let $|\mathcal{M}|$ be the associated universe, and for every term, function or predicate a, $a^{\mathcal{M}}$ is the realization of a in $\mathcal{M}$.

Definition 5.32 *Let $\mathcal{M}_1, \mathcal{M}_2$ have the same universe U, and let $\xi \in U^k$, where k is the arity of S_1. We say that $\mathcal{M}_1 \leq^{\xi} \mathcal{M}_2$ iff*

1. *$K^{\mathcal{M}_1} = K^{\mathcal{M}_2}$ for every function or predicate constant K that is not in S,*
2. *for any $i = 1, ..., n$, $S_i^{\mathcal{M}_1}$ and $S_i^{\mathcal{M}_2}$ coincide on $\{\eta \mid \neg V_i^{\mathcal{M}_1}(\xi, \eta)\}$*
3. *$S_1^{\mathcal{M}_1}(\xi) \implies S_1^{\mathcal{M}_2}(\xi)$.*

Proposition 5.33 ([Lif87b]) *Let $\mathcal{M}$ be a model of $A(S)$.*

$$\mathcal{M} \models C_{PW}[A; S_1; S_1/V_1, ..., S_n/V_n] \iff \forall \mathcal{M}' \in [\![A(S)]\!]\ \forall \xi \in |\mathcal{M}|^k \neg(\mathcal{M}' \leq^{\xi} \mathcal{M} \wedge \mathcal{M} \not\leq^{\xi} \mathcal{M}')$$

Other words, $\mathcal{M} \models C_{PW}[A; S_1; S_1/V_1, ..., S_n/V_n]$ iff for each $\xi \in |\mathcal{M}|^k$, $\mathcal{M}$ is minimal relative to $\leq^{\xi}$.

5.3.3 Closed World Assumptions

In [VK76] a semantical characterization of $CWA(T)$ was given in terms of minimal Herbrand models of T.

For $GCWA(T)$, Minker [Min82] provided a semantic characterization in terms of minimal Herbrand models for theories without functional symbols. He proved that for consistent T, $GCWA(T)$ is consistent (arbitrary T).

5.3.4 Default Logic

There are several semantics for Default Logic that were described in the literature. As seen in the sections above, every cumulative entailment relations is representable by a cumulative model (see section 5.1.1 on page 71). Thus, all special cases of Default Logic or other Default Logics that are cumulative are representable this way. Therefore, we shall restrict ourselves to the general case, referring the interested reader to Poole [Poo88a] on preferential semantics for the prerequisite free (open) normal defaults (for that kind of defaults, one can give semantics using Circumscription and its preferential semantics as well (see section 6.3 on page 89).

Stable Model Semantics In the case that a default theory $\langle W, D\rangle$ has W consisting of conjunctions of atoms and the defaults are of the form $\frac{a_1 \wedge ... \wedge a_n : \neg b_1, ..., \neg b_m}{c}$, where the a_i, b_j, ω are all atoms, the default rules can be put in correspondence with the logic program (section 7.2.1), with negation as failure (section 3.5.2) clause

$$a_1, ..., a_n, notb_1, ..., notb_m \Rightarrow c$$

where $notb_i$ means negation-as-failure.

Theorem 5.34 *Let P be the translation of a default theory T with the above properties. Then a model M is stable (see section 7.2.1) for the logic program produced by the translation above iff M is an extension of T.*

Maximal Models The following sequence of definitions was proposed by Etherington [Eth87b] to try to capture Default Logic in a semantics that uses preferential methods.

Definition 5.35 ($\mathcal{M}_1$ is preferred over $\mathcal{M}_2$) *Let $T = \langle W, D\rangle$ a default theory. If $\mathcal{M}_1$ and $\mathcal{M}_2$ are sets of models of W, we say that $\mathcal{M}_1 \succ_0^D \mathcal{M}_2$ if*

1. *there is a default of the form $\frac{\alpha : \beta_1, ..., \beta_n}{\gamma}$, an instantiation of an element of D, such that $\forall M \in \mathcal{M}_2\ M \models \alpha$ and*
2. *for all i such that $1 \leq i \leq n$, there exists $M \in \mathcal{M}_2$ such that $M \models \beta_i$ and*
3. *$\mathcal{M}_1 = \{M \in \mathcal{M}_2 \mid M \models \gamma\}$.*

We let $\succ^D$ be the transitive closure of $\succ_0^D$.

Intuitively, $\mathcal{M}_1 \succ^D \mathcal{M}_2$ implies that there is a sequence of defaults from D that leads leads from the set of models $\mathcal{M}_2$ to a subset $\mathcal{M}_1$ by removing those models that do not "comply".

If our definition ended here, we would have proved that Reiter's Default Logic is Cumulative. As it is not, there is something to be added to the recipe, and Etherington adds the notion of *stability* as that thing.

Definition 5.36 (Stability) *Let $\Delta = \langle W, D\rangle$ be a default theory, and let Γ be a $\succ$-maximal element. $\mathcal{I}$ is* stable *for Δ iff there is a $\mathbb{D}' \subset D$ such that $\mathcal{I} \succ^{D'} [\![W]\!]$ and for each $\frac{\alpha:\beta_1,\ldots,\beta_n}{\gamma} \in D'$, there are $M_i \in \mathcal{I}$ such that $\mathcal{M}_i \models \beta_i$.*

In other words, a set of models is stable for a default theory $\langle W, D\rangle$ if it is a maximal specialization of the set of models of W and does not refute the justifications of any of the defaults used in the specialization.

Theorem 5.37 ([Eth87b]: Soundness and Completeness) *E is an extension for Δ iff there is a set of models $\mathcal{I}$ that is stable for Δ such that $M \models E \iff M \in \mathcal{I}$.*

Corollary 5.38 *φ is true in every extension of Δ iff it is true in every maximal model of Δ.*

Epistemic Semantics Lin and Shoham [LS90] have developed an epistemic semantics for both Default Logic and Autoepistemic Logic. They use modal logic with two operators, K ("known") and A ("assumed"), a logic they call GK. Given a standard Kripke semantics of these symbols, they develop a preference over Kripke interpretations, and then choose among the preferred models those that satisfy a condition over the knowledge and assumptions.

Definition 5.39 *A Kripke interpretation M_1 is* preferred *over M_2, written $M_1 \prec M_2$, if for all base formulas φ, $M_1 \models A\varphi \iff M_2 \models A\varphi$ and $M_1 \models K\varphi \implies M_2 \models K\varphi$.*

A Kripke interpretation M is a preferred model *of a set of axioms T, if M is $\prec$-minimal and $M \models K\varphi \iff M \models A\varphi$.*

Thus, one can see that the "stability" of "fixed-point" we saw earlier with Default Logic is embodied in the requirement for the assumptions to be equivalent to the knowledge.

The default theory $\Delta = \langle W, D\rangle$ is translated into the modal theory Δ_{GK} by translating

$$\begin{array}{lcl} \varphi \in W & \implies & K\varphi \in \Delta_{GK} \\ \frac{\alpha:\beta_1,\ldots,\beta_n}{\gamma} \in D & \implies & \mathcal{K}\alpha \wedge \neg A\neg\beta_1 \wedge \ldots \wedge \neg\beta_n \Rightarrow K\omega \end{array}$$

Theorem 5.40 *A consistent set of sentences E is a default extension of Δ iff there is a preferred model M of Δ_{GK} such that $E = \{\varphi \mid M \models K\varphi\}$.*

One semantics goes along the lines of translating to Logic Programs (see section 7.2.1 on page 94). The relationship between DL and stable model semantics for LP was discovered by Bidoit and Froidevaux [BF91] and Marek and Truszczynski [MT89b].

5.4 Nonmonotonic Modal Logics

The "classic" semantics for Autoepistemic Logic (section 3.3.2) was given by Moore [Moo87] as follows: Each stable set S corresponds to a *universal S5 model* (the accessibility relation R is an equivalence relation) $\mathcal{M}_S$ that is the maximal such that S is the set of formulas it satisfies.

An *Autoepistemic Model* for Σ then refers to an S5 model whose set of theorems corresponds to a stable expansion for Σ in AE. The correspondence with Autoepistemic logic is that *T is a stable autoepistemic theory iff T is the set of formulas that are true in all the worlds of some complete-S5 structure.*

A different semantics for Autoepistemic Logic is presented in section 5.3.4 on the preceding page.

Extensions to the First-Order Case The First-Order case with no free variables in the modal formula is naturally treated like the propositional case (as far as defining the fixed points for the nonmonotonic logics we presented above). The case where variables that appear in the scope of the modal operators but are quantified outside is described for the monotonic case in section 2.3.6 on page 16.

The problem with extending AEL to the first-order case is that we conclude $L\neg\varphi$ by showing $\varphi \notin \Gamma$ (for Γ our fixed-point). To check $\varphi \in \Gamma$ we need a closed sentence φ', which various methods tried to suggest.

Konolige [Kon89], [Kon92b] and [Kon93] described three related extensions of AEL's *Belief-set Semantics* to FOL. All of his extensions try to give names to *all* the object in the domain, allowing definition of the semantics for variables that are quantified-in as the equivalent formula with the variables replaced by the appropriate ground term. The following are the three proposed semantics. We take $\mathcal{N}$ to be a countably infinite set of names (constant symbols) for elements that is added to our original language and Γ is our theory. We define the possible world semantics for First-Order AEL.

Any-name Semantics Let k be an element in the universe U, w a world in the kripke structure. $w, x/k \models L\phi(x)$ iff for some name $t \in \mathcal{N}$ such that $v(t) = k$, $\phi(t) \in \Gamma$.

All-name Semantics Let k be an element in the universe U, w a world in the kripke structure. $w, x/k \models L\phi(x)$ iff for all names $t \in \mathcal{N}$ such that $v(t) = k$, $\phi(t) \in \Gamma$.

Unique-name Semantics A specialization of both Any-name and All-name, adding unique-name axioms for all the symbols of $\mathcal{N}$.

Levesque [Lev90] described an extension to the possible-worlds semantics that is also based on names for elements in the universe. His language $\mathcal{OL}$ (section 3.3.3 on page 43) has no function symbols and constants are restricted to a set of standard names $\mathcal{N} = \{n_i\}$. The possible worlds semantics for his logic consists of a set of worlds W (no accessibility relation). The evaluation rules for $\mathcal{OL}$ are

$$\begin{array}{lcl} w \models_W L\varphi & \iff & \forall w' \in W.w' \models \varphi \\ w \models_W (n_i = n_j) & \iff & n_i \text{ is the same as } n_j in the world w \\ w \models_W \exists x\varphi & \iff & \exists n \in \mathcal{N}.w \models \varphi(x/n) \end{array}$$

Because of the presence of a set of standard names covering all individuals, this semantics gives the unique-names interpretation to quantifying-in.

6 Relating Principles: Translations

Nonmonotonic Reasoning research is saturated with works comparing nonmonotonic systems and finding generalizations of nonmonotonic systems and their principles. Nevertheless, the stream of works in this line of research continues in full steam even today. Why is this work of any further interest, one may (or should) ask. At least a partial answer lies in the use perceived for nonmonotonic reasoning.

It became clear in the mid-90's that, despite the promise that nonmonotonic systems held for the field of Commonsense Reasoning, there was no clear way for accounting for all the various potential uses a nonmonotonic reasoning system may have. There is no one system that is "best", and even the derivation of a methodology for the use of these nonmonotonic systems advanced pretty slowly. Now that there are several nonmonotonic systems and several methodologies for some applications (especially Inheritance Hierarchies and Theories of Action), the need for a framework, that will allow us to choose the right system and the right methodology for a given problem, is emphasized. With one system, that describes the rest of the systems as special cases, one can develop a methodology of choose&apply for the different problems. Wherever one system seems to do better than the others, we shall prefer that one.

The landmark work of Imielinski [Imi87] opened a path for much of the systematic work of investigating translations between nonmonotonic systems. As Imielinski puts it,

> The possibility of translation between two formalisms could help in combining their advantages; we could, for example, specify the system in default logic, making it conceptually simple and easy to update (modularity) and perform the reasoning using its circumscriptive translation.

We first present Imielinski's framework and then go on to describe the different translations investigated in the literature.

6.1 Governing Principles

Most translations found in the literature refer to the analysis of translations done by Imielinski [Imi87] (and various versions of that paper).

Imielinski suggested the notion of modularity. He splits every theory into two parts: general facts and specific facts. The general facts are the "defaults" and the specific facts are base-language sentences. In Default Logic the "specific facts" correspond to the "facts" and the "general facts" correspond to the "Defaults". In Circumscription, the general facts correspond to sentences in which we have *ab*'s or, more generally, minimized predicates. The "specific facts" are simply sentences in the base language. The situation for Modal-based approaches is similar.

The following are slight generalizations of the original definitions as presented in [Imi87] and [Got95].

Definition 6.1 ([Imi87]) *A translation is* modular *if it translates general facts independently of specific facts (the latter stay the same).*

Put differently, if we change the facts of our original theory, there is no need to recompute the translation of the defaults (the general facts). Some people refer to these translations as *context independent.*

Without the restriction for modularity, it is difficult to estimate the value of the translation, for, obviously, every system can be translated to any other system (limiting expressivity of the different systems) by simply computing all the inferences made in the original system and asserting these results in the target system (if the reader is not convinced, see Etherington [Eth87a]).

Definition 6.2 *A* faithful *translation is a mapping which transforms a theory in one nonmonotonic system into a theory in another such that the objective parts (the base language sentences) of the extensions/expansions/minimal-models of the original and the target are identical.*

In addition, there are other factors that one should pay attention for when devising a translations, such as the translation of fact-for-fact[32] (we would rather translate each fact independently of the rest), the computational complexity of the translation and the exact subset of theories that can be translated using such a translation.

6.2 Default Logic vs Autoepistemic Logic

The relationship between Reiter's Default Logic (section 3.2.1) and its variants and Moore's Autoepistemic Logic (section 3.3.2) and its variants has been investigated by many researchers among which are Konolige [Kon88b],[Kon92a], Marek and Trsuszcynski [MT89a],[MT90],[MT93a], Trsuszcynski [Tru91b], [Tru91a], Kaminski [Kam91], Lin and Shoham [LS90], Lifschitz [Lif91], Gottlob [Got95],[Got93], Chen [Che94], Janhunen [Jan96] and Shchwarz [Shv90], [Sch96].

[32] Translations that are one-to-one and modular are called *local* in [Kon88b].

6.2.1 Default Logic vs Strongly Grounded AEL

The first result is due to Konolige. He defines the AE transformation of a default theory to AEL as

$$\frac{\alpha : \beta_1, ..., \beta_n}{\omega} \mapsto (L\alpha \wedge \neg L\neg\beta_1 \wedge .... \wedge \neg L\neg\beta_n) \Rightarrow \omega \tag{13}$$

Notice that this translation is modular and we leave the facts W unchanged.

Theorem 6.3 ([Kon88b]) *Let A be the AE transform of a default theory $T = \langle W, D\rangle$. A set E is a default extension of T iff it is the kernel of a strongly grounded AE extension of A.*

So, the translation is faithful and thus, Strongly Grounded AEL can express Reiter's Default Logic theories using a modular translation.

Konolige also showed that the opposite translation is faithful, and since every sentence of the language has an equivalent to a sentence in extended normal form [Kon88b] (see definition 3.38 on page 43), after transforming to extended normal form, there is a faithful, modular translation from Strongly Grounded AEL to Reiter's DL. Take the translation

$$\neg L\alpha \wedge L\beta_1 \wedge .... \wedge L\beta_n \vee \omega \mapsto \frac{\alpha : \neg\beta_1, ..., \neg\beta_n}{\omega}$$

Theorem 6.4 ([Kon88b]) *For any sentence A of $\mathcal{L}$ in extended normal form, there is an effectively constructible default theory $\langle W, D\rangle$ such that E is a default extension of $\langle W, D\rangle$ iff it is the kernel of a strongly grounded extension of A.*

Later, Truszczyński [Tru91a] showed that the same translation can be used to translate to Strongly Grounded AEL with any of the range **N–KD45** of modal logics.

6.2.2 Default Logic vs Nonmonotonic Modal Logics

Marek and Truszczyński [MT90],[MT93a] and Truszczyński [Tru91b],[Tru91a] tried to find translations based on the concept of nonmonotonic modal logic as introduced by McDermott and Doyle [MD80] and McDermott [McD82]. They show that Default Logic is faithfully embeddable in a wide range of different nonmonotonic modal logics.

In the case of strongly grounded expansions for nonmonotonic logic (this is the same definition as 3.39 on page 43), they showed that the translation (13) is faithful for any modal logic between the logic of pure necessitation **N**and the logic KD45.

Another translation was devised by Marek and Truszczyński [MT90], where a default is translated to $L\alpha \wedge \neg LL\neg\beta_1 \wedge ... \wedge \neg LL\neg\beta_n \Rightarrow \omega$. Under this interpretations, consistent extensions of $\langle W, D\rangle$ correspond exactly to **N**-expansions of the translation of $\langle W, D\rangle$.

A different translation was devised by Truszczyński [Tru91b]:

$$W \mapsto LW \quad \frac{alpha : \beta_1, ..., \beta_n}{\omega} \mapsto L\alpha \wedge LM\beta_1 \wedge ... \wedge LM\beta_n \Rightarrow L\omega \qquad (14)$$

This translation interprets a default to say "If I know α and I know that $\beta_1, ..., \beta_n$ are possible, then I know ω". Let us write $emb(W, D) = LW \cup \{emb(d) | d \in D\}$ where $emb(d)$ is our translation of a default d as described in (14).

Theorem 6.5 ([Tru91b]) *Let $\langle W, D\rangle$ be a default theory and let $S \subset \mathcal{L}$ be consistent and closed under propositional consequence. Then, S is an extension of $\langle W, D\rangle$ if and only if $ST(S) = Cn_S(emb(W, D) \cup \{\neg L\varphi \mid \varphi \notin ST(S)\})$.*

Put in different words, every extension of a default theory is the kernel of *the* expansion of the embedding $emb(W, D)$, and vice versa.

Thus, NML (as introduced by [MD80],[McD82]) with any of the modal logics T^-–S4F can express Default Logic reasoning.

6.2.3 Default Logic vs Standard Autoepistemic Logic

Unfortunately (or not), the translation that allowed theorem 6.5 falls short of representing DL in standard[33] AEL, as AEL is NML_{KD45} and this logic does not include T^-.

Other attempts for translating DL into close formalisms to AEL have been developed by Niemelä [Nie91], Lin and Shoham [LS90] and Kaminski [Kam91]. All approaches introduced a different version of AEL which captures Default Logic, while the question of the existence of a modular translation remained open.

Finally, in 1995, Gottlob showed that there is no modular translation of Default Logic into AEL.

Theorem 6.6 ([Got95]) *There exists no faithful modular translation from default logic into AEL (nonmonotonic KD45). Such a translation does not exist even if the defaults are restricted to be normal.*

This result holds even if we restrict our default theory to be justification-free (i.e., defaults are of the form $\frac{\alpha:}{\gamma}$).

For prerequisite-free defaults, Konolige's translation (13) works, as can be easily proved by combining theorems 3.41 and 6.4 and noticing that every prerequisite-free default will be essentially translated to a simple normal form.

As far as non-modular translations are concerned, [Got95] showed how can such a faithful translation be constructed by combining a translation from DL to $NML_{\mathbf{N}}$ (modular) with a polynomial-time (nonmodular) translation from $NML_{\mathbf{N}}$ to AEL. Later, [Sch96] described a one-step translation that is somewhat more intuitive.

[33] By Standard AEL we mean the original AEL.

6.3 Circumscription vs Default Logic

Early works of Grosof [Gro84] and Imielinski [Imi85], [Imi87] gave clear boundaries on the ability to modularly translate from Default Logic to Circumscription.

Theorem 6.7 ([Imi85]) *Let A and B be independent propositional formulas. The set of defaults $D = \{\frac{A:B}{B}\}$ is not modularly translatable into Circumscription (with abnormality theory), no matter how many abnormality variables are introduced.*

The positive part is summarized by combining the results of Grosof and Imielinski.

Theorem 6.8 ([Gro84],[Imi87]) *Every propositional, prerequisite-free, semi-normal default theory is modularly translatable to propositional Circumscription (simple abnormality theory).*

Imielinski gave sufficient and necessary conditions for default theories to be modularly translatable to circumscription, but we do not present it here as it introduces some extra notation and does not help our intuition.

The first order case of semi-normal defaults without preconditions is translatable to circumscription, if the facts of the default theory imply domain-closure axioms and unique names axioms (see the opposite below).

For the opposite direction, translating Circumscription to a default theory, Etherington [Eth87a] showed that, if the background theory T logically entails domain closure axioms, then Circumscription (with all symbols varied) is equivalent to a prerequisite-free default theory. Here, the translation is

$$Circ[T; P; Z] \mapsto W = T; D = \{\{\frac{\neg Px}{\neg Px}, T\}$$

Theorem 6.9 ([Eth87a]) *Let T be a theory, $\alpha_1, ..., \alpha_n$ are ground terms such that $T \vdash \forall x(x = \alpha_1 \vee ... \vee x = \alpha_n)$ and that $T \vdash \alpha_i = \alpha_j$ or $T \vdash \alpha_i \neq \alpha_j$. Let Z be the set of all constant symbols of $\mathcal{L}$. Then, those formulas true in every extension of $D = \{\{\frac{\neg Px}{\neg Px}, T\}$ are precisely those entailed by $Circ[T; P; Z]$.*

This theorem has several limitations. The first and foremost is the restriction to finite domains. This is a reasonable restriction sometimes, but it is still disturbing. Etherington also showed that the same translation does not work if T does not entail a domain-closure axiom or if Z does not include all the constant symbols of $\mathcal{L}$.

At this point the reader may be convinced that there is no translation in the case that there are some fixed constant symbols in the Circumscription[34].

[34] I believe this is a fabulous example of a case where a result/example may make us (researchers) jump to (the wrong) conclusions.

Theorem 6.10 ([dKK89]) *Let A be a set of sentences in $\mathcal{L}$, P a sequence of predicate constant symbols in $\mathcal{L}$ and Z a sequence of constant symbols in $\mathcal{L}$. If $Q \notin Z$ (held "fixed") a predicate, and Q' is a new predicate constant symbol, then*

$$Circ[A; P; Z] \equiv Circ[A \wedge (Q' = \neg Q); P, Q, Q'; Z]$$

Thus, a circumscription formula with some fixed constant symbols can be substituted for a formula with no fixed constant symbols, extending the applicability of theorem 6.9 on the preceding page.

For a comparison with a slightly more general theorem (than theorem 6.9 on the page before) for AEL, see theorem 6.11.

6.4 Autoepistemic Logic vs Circumscription

In trying to express AEL or DL in Circumscription, researches have found that the basic problem is that a minimal-model or even preferred-model semantics does not have the capability of representing the required proof-theoretic or epistemic concepts such as the difference between LPx and Px (see section 5 on page 70). This seems to be the reason most researchers (e.g., [Kon89], [Kon93], [Imi87]) concluded that Circumscription cannot represent AEL or any other Modal Nonmonotonic Logic.

On the other direction, Konlige showed in [Kon89] that AEL, properly extended to the case of First-Order Modal Logic (with quantifying-in), can capture (in some cases) all the first-order sentences entailed by Circumscription. His extension and others are described in section 5.4 on page 84.

Konolige defined the following translation. For A a set of first-order sentences, leaving a countably infinite number of names from $\mathcal{N}$ unused ($\mathcal{N}$ is a set of constant symbols used in Konolige's FOL-extension to AEL to reference all the individuals) and P, Q sequences of constant symbols (P is all predicates),

$$M(A; P; Q) \stackrel{def}{=} A \bigcup_i \{\forall x(M\neg P_i x \Rightarrow \neg P_i x)\} \bigcup_i \{\forall x(M\neg Q_i x \Rightarrow \neg Q_i x)\} \bigcup_i \{\forall x(MQ_i x \Rightarrow Q_i x)\} \tag{15}$$

This formula says that for all x, if it is possible that $\neg P_i x$, then $\neg P_i x$ is the case. Same for $Q_i x$ and $\neg Q_i x$ (recall that M is the invariant modality of L (i.e., $M = \neg L \neg$)). We now translate $Circ[A; P; Q] \mapsto M(A; P; Q)$.

Theorem 6.11 ([Kon89]) *Let A, P, Z as above and let φ be a first order sentence. If $Circ[A; P; Q]$ has a set of countable minimal models that characterize*[35] *it, then φ is true in every stable expansion of $M(A; P; Z, =)$ iff $Circ[A; P; Q] \models \varphi$.*

A result due to Schlipf [Sch86] is that entailment over minimal countable structures (this is $M(A; P; Q)$ and the entire FOL-extension of Konolige) is not equivalent to entailment over all structures, so this theorem is not true in the general

[35] I.e., $Circ[A; P; Q] \models \varphi$ iff for every M that has a countable universe $M \models \varphi$.

case. Notice, however, that this is an improvement over the result of Etherington regarding Default Logic and Circumscription, where only finite domains are allowed (theorem 6.9 on page 89).

The relationship between Introspective Circumscription (section 3.4.2) and Autoepistemic Logic was discussed by [Lif89b] and then expanded on by [MS94].

6.5 Relationships between other systems

6.5.1 Model-Preference Defaults vs Default Logic

Selman and Kautz [SK90] showed that their Model-Preference Default Rule systems can sometimes be translated to a prerequisite-free semi-normal default theory in Reiter's Default Logic. In this section, we will use the term MP-default to refer to their default rules, whereas plain "defaults" will refer to Reiter's defaults.

They define their translation for MP-default rules that are horn and are specificity-ordered (see section 3.4.3 on page 48). Their translation aggregates different MP-default rules in D to creates default theories that correspond to their meaning. For an MP-default $\alpha \Rightarrow x \in D$, we create the default

$$\frac{x \wedge \alpha \wedge \neg\beta_1 \wedge ... \wedge \neg\beta_k}{x}$$

where each β_i corresponds to $(\alpha \cup \beta_i) \Rightarrow \neg x \in D$.

Also, for each literal x (positive or negative) such that $TRUE \Rightarrow \neg x \notin D$, we create the default

$$\frac{x \wedge \neg\delta_1 \wedge ... \wedge \neg\delta_l}{x}$$

where each δ_i corresponds to $\delta_i \Rightarrow \neg x \in D$.

Theorem 6.12 ([SK90]) *Let D be a set of specificity-ordered Horn MP-defaults (i.e., $\mathcal{DH}^+$) that does not contain a mutually blocking pair of MP-defaults. Let $\mathcal{D}$ be its translation into a set of defaults in default logic. If the preference relation induced by D is a partial order*[36]*, then M is a maximal model of $\leq_D$ iff $Th(M)$ is an extension of the default logic theory $\langle \mathcal{D}, \emptyset \rangle$.*

Selman and Kautz say that the first restriction (the non-mutually-blocking defaults) can be removed, but this requires a more complicated translation.

6.5.2 Circumscription vs Closed World Assumptions

The relationship between Circumscription and the various forms of the CWA was studied in [GPP86], [GPP89], [Lif85a], [Prz88], [Rei82]. We discuss the relationship in detail in section 3.5.2.

[36] All MP-default sets D define a pre-order. Thus, all we need to get a partial order is anti-symmetry.

6.6 Other Comparison Works

Other comparison works include [GL89] (Circumscription and Logic Programs), [GPP89] (comparing Circ, Prioritized Circ, CWA, GCWA, ECWA, Iterated CWA), [Lif89b][MS94] (Introspective Circumscription and Autoepistemic Logic), [She84] (comparing CWA and NAF), [BDKT97] (Logic Programming and AEL and DL (semantic representation), [MLR91] talks about the relationship between Circ and disjunctive LP and [YY93] talks about Autocircumscription and Logic Programs.

7 Other Issues

In this work we did not touch many topics that are very relevant to research in nonmonotonic reasoning and some that are in the core of research. The first and foremost is the issue of algorithms and computational complexity. Others are related areas such as Logic Programming, Belief Revision and Commonsense Reasoning. In this section we will try to give some of the main references to these topics and more and fill in the gaps on some essentials.

7.1 Computational Aspects

Each of the systems has its own literature of computational complexity and algorithms. We describe only a few of them here, and mostly give point of reference.

One of the most comprehensive surveys of nonmonotonic computational complexity is the work of Cadoli [Cad95]. A shorter paper by Cadoli and Schaerf [CS93] is also a useful resource.

7.1.1 Computing Circumscription

Roughly speaking, there are two kinds of computation systems for circumscription. The first is an equivalent transformation system of circumscription into first-order sentences ([DLS97], [GO92], [KP90], [Lif85b], [Rab89], [CEG92]) and logic programs ([GL89], [Prz89b]). The second is a deductive computation system for circumscription ([Gin89], [Prz89b]).

The following theorem allows us to remove the varied constant symbols from the circumscription formula, in price of incrementing the order of the formula (i.e., the formula may become second-order if it was originally first-order).

Theorem 7.1 ([Lif93]) *Let A be a set of sentences in $\mathcal{L}$, P a sequence of predicate constant symbols in $\mathcal{L}$ and Z a sequence of constant symbols in $\mathcal{L}$. Then,*

$$Circ[A(P,Z);P;Z] \equiv Circ[\exists z.A(P,z);P;]$$

Another interesting theorem allows us to reduce the circumscription formula to a slightly less complicated circumscription formula.

Theorem 7.2 *$Circ[A \wedge B; P; Z] \equiv Circ[A; P; Z] \wedge B$ if P, Z do not appear in B.*

Other works on reducing Circumscription to FOL include the works of [Lif86a] and [Mot87] on the satisfiability of circumscription and Cadoli et al [CS92] on approximate inference for circumscription and default logic. Also, Ginsberg [Gin89] used an ATMS to supply a circumscriptive theorem prover. Finally, Cadoli [Cad92] wrote about the complexity of model checking for circumscriptive formulae.

7.1.2 Reducing Second Order Logic to FOL: SCAN

The problem of reducing Second Order Logic to FOL was tackled already by Ackerman ([W. 35], [HA50b]) who devised some methods for the reduction. SCAN ,devised by Gabbay and Ohlbach ([GO92]), is an algorithm generalizing over the reduction proposed by Ackermann.

SCAN reduces existentially quantified second-order formulas to FOL. It does that in a way somewhat similar to resolution. The idea is to resolve the second-order Skolem Constant out, and so to remain only (an equivalent) first-order formula. The process is called C-resolution, as we resolve, but maintain constraints. If we resolve $P(x) \vee Q(x)$ with $\neg P(y) \vee R(y)$, we will have $x = y \vee Q(x) \vee R(y)$ in the resolvant, because that will enable us to specify P independently (if $x \neq y$ we will just specify P appropriately).

The result is an algorithm that tries to get to an equivalent first-order formula by resolving all the clauses it can. Evidently, this process might never end. In particular, there are some cases that are easily detectable as *loops* (e.g., $P(x, y) \vee \neg P(y, x)$). In the case of circumscription, we negate the circumscription formula, and use SCAN (the circumscription formula is universally quantified).

The reduction algorithm proposed by Doherty, Lukaszewicz and Szalas ([DLS95], [DLS97]) is another powerful method for eliminating second-order existential quantifiers. Unlike SCAN, this algorithm always terminates, and either returns an equivalent first-order formula or reports failure.

7.1.3 Computing Default Logic

[KS89],[KS91] analyzed Default Logic in the propositional case with the following results

- NP-hard for unary defaults
- n^2 for disjunction free, ordered theories
- NP-hard for disjunction free normal theories or unary ordered theories
- n to determine if a given literal appears in any extension of a Horn Default theory.
- co-NP-hard to determine if a given literal appears in every extension for Horn default theories.

- n^3 to determine if a given literal appears in every extension of a normal unary theory.

Later, Amati et al [AAGP96] describe a general approach of tableaux for various Default Logics. They apply their method for Reiter's DL (section 3.2.1), Brewka's (section 3.2.3) and Lukaszewicz's (section 3.2.2) Default Logic. Schlechta [Sch95a] talked about a two-stage approach to computing First-Order Default Logic.

[RDB89] and [JK90] used ATMSs to compute extensions of Autoepistemic Logic and Default Logic.

7.2 Closely Related Areas

7.2.1 Logic Programs

The field of Logic Programs is an entire field in its own right, and we do not intend to give a survey of it in the short space we have here. The interested reader should refer to [AE82] or the book [Llo87] for an extensive review of the logic programming field. Here, we describe the issues that we consider important from the nonmonotonic reasoning perspective, omitting the well-founded semantics (see [GRS91], [Prz94], [PAA91] for more details).

Among the different semantics given for Logic Programs, the most prominent ones are the Well-Founded semantics, the Stable Semantics, the two-valued and the three-valued program completion semantics. In this section we deal only with the Stable Semantics. A comparison between the four different semantics we just mentioned is given by Schlipf [Sch95b].

Propositional Programs A *propositional program* is a finite set of formulas of the form

$$A_1 \wedge ... \wedge A_m \Rightarrow B, \tag{16}$$

where $m \geq 0$, each of $A_1, ..., A_m$ is a propositional literal (i.e., a propositional symbol possibly preceded by negation[37]) and B is a propositional symbol. These formulas are called the *rules* of the propositional logic program Π. If $m = 0$, then we call the clause a *fact*. The antecedent of the rule is called the *body* and the consequent is called the *head*.

Stable Models Semantics for Propositional Programs [GL88] gave the Stable Models Semantics to Logic programs. Let Π be a propositional logic program whose propositional symbols are $P_1, ..., P_n$. For any $Y \in \{false, true\}^n$ (i.e., Y is a truth assignment for the P_i's), let Π_Y be the program obtained from Π by deleting (i) each rule that has a negative literal $\neg P_i$ in its body such that $Y_i = true$, and (ii) all negative literals in the bodies of the remaining rules. Clearly, Π_Y is negation-free, so that it has a unique minimal Herbrand model.

[37] This is the Negation-as-Failure negation.

This model is denoted by $S_\Pi(Y)$. Thus, S_Π is a mapping of $\{false, true\}^n$ into itself. We say that Y is a *stable vector* of Π if it is a fixed-point of S_Π (i.e., $S_\Pi(Y) = Y$).

Theorem 7.3 ([GL88]) *Any stable vector of* Π *is a model of* Π.

Thus, stable vectors can also be called *stable models*.

A propositional program Π is meaningful if it has exactly one stable model. The value of P_i in that model is declared to be the value assigned to P_i by Π.

General (First-Order) Logic Programs A *general logic program* is a finite set of formulas of the form (16) where each of $A_1, ..., A_m, B$ is an atomic formula. If all the formula involved in the clause are ground, then the clause is called *ground*. If there are no negative atoms, then the clause is called a *definite clause*. A *definite logic program* (or *Horn logic program*) is a logic program which consists solely of definite program clauses.

Both the autoepistemic approach to LPs and the Stable Model approach reduce general programs to the propositional case by replacing every rule with variables by its ground instances. This amounts to ignoring any models of a logic program other than its Herbrand models.

The *predicate logic interpretation* of logic program P is denoted by $pc(P)$ and is the first-order theory which consists of the predicate logic interpretation of all program clauses in P.

Let P be a logic program. Every model of the first-order $pc(P)$ is called a *model* of P. Every Herbrand model of $pc(P)$ is called a *Herbrand model of* P.

Theorem 7.4 *Every logic program possesses at least one minimal Herbrand model. A definite logic program has a unique least Herbrand model.*

The semantics for first-order logic programs with negation (negation-as-failure) is then given the same way we described it above.

Logic Programs with Classical Negation One limitation of the discussion so far is that there is only one type of negation. Gelfond and Lifschitz continued their work by extending the language to include classical negation in [GL90] and [GL91].

An *extended program clause* is an expression of the form

$$A_1, ..., A_n, notB_1, ..., notB_m \Rightarrow L$$

where $A_1, ..., A_n, B_1, ..., B_m, L$ are literals, that is, atomic formulae or their negations. An *extended logic program* is a set of extended program clauses.

According to this definition, we are allowed to write the extended clause $\neg train \Rightarrow cross$ to express that a car should only cross the tracks if on train is approaching (classic negation). This is different from writing $nottrain \Rightarrow cross$ which says that a car should cross the road if there is no information signaling that a train is coming (including the fact that a car blocks your view).

Circumscriptive theories were used to give semantics to logic programs by Lifschitz [Lif87a]. Introspective Circumscription is also used to give semantics to general logic programs (see section 3.4.2 on page 47).

7.2.2 Theories of Action

In this section we chose to demonstrate one of the applications of nonmonotonic reasoning to Commonsense Reasoning, by displaying a simulation of prototypical chronological minimization (PCM) using pointwise circumscription.

Doherty and Lukaszewicz [DL94] examined various solutions to the Frame Problem, implementing them using either McCarthy's Circumscription or Lifschitz's Generalized Pointwise Circumscription (formula 7 on page 24). We briefly review some of the theory of Features and Fluents [San94] and one of the solutions that was examined in [DL94], namely, PCM, for which Pointwise Circumscription was used.

The general theory for PCM The language $\mathcal{L}(FL)$ is a sorted first-order language with equality. There are two domain independent sorts, $\mathcal{T}$ for time points (t, s are variables of that sort) and $\mathcal{F}$ for propositional fluents (f, g are variables of that sort), which are functions from time to truth values. We include the predicate symbols $Holds, Clip$ of type $\mathcal{T} \times \mathcal{F}$ and the predicate symbols $<, \leq$ of type $\mathcal{T} \times \mathcal{T}$ (total orders on time points).

Let $\Gamma_C = \Gamma_{OBS} \cup \Gamma_{SCD} \cup \Gamma_{UNA}$ denote the set of formulas in $\mathcal{L}(FL)$ corresponding to the given scenario description, where Γ_{OBS} are the observation axioms (e.g., $Holds(f, s)$), Γ_{SCD} are the schedule axioms (a form of "effect axioms") and Γ_{UNA} are the appropriate unique name axioms for the fluent constants. Γ_{PER} denotes the set containing the two persistence axioms

$$\forall ft.Holds(t, f) \Rightarrow [Holds(t+1, f) \bar{\vee} Clip(t+1, f)]$$
$$\forall ft.\neg Holds(t, f) \Rightarrow [\neg Holds(t+1, f) \bar{\vee} Clip(t+1, f)]$$

PCM is defined roughly as follows (see [SS94a] for more details): Interpretations $I \ll_{pcm} I'$ iff $|I| = |I'|$ (the objects are the same) and there is t_0 such that

1. for all $t < t_0$, $R(t) = R'(t)$ (the fluents are identical in I and I' before t_0) and
2. $breakset(I, t_0) \subset breakset(I', t_0)$ (the fluent-value changes from $t_0 - 1$ to t_0 in I are a subset of the appropriate changes in I').

Then the set of models for PCM is $S_pcm(T) = Min(\ll_{pcm}, [\![T]\!])$.

For the logic of PCM, we include the following. Let $\Gamma(Clip, Holds)$ denote the theory $\Gamma_C \cup \Gamma_{PER}$. Let $V(t, f) = \lambda t' f'.[\langle t, f\rangle = \langle t', f'\rangle \vee t < t']$ and $U(t, f) = \lambda t' f'.[t \leq t']$. Then, $C_{PW}[\Gamma(Clip, Holds); Clip; Clip/V, Holds/U]$

is

$$\begin{array}{l}\Gamma(Clip, Holds)\wedge \\ \forall t, f.\forall clip, holds.\neg[Clip(t,f)\wedge\neg clip(t,f)\wedge \\ \quad EQ_{V(t,f)}(Clip, clip)\wedge EQ_{U(t,f)}(Holds, holds)\wedge \\ \quad \Gamma(clip, holds)]\end{array} \tag{17}$$

The intention is that when minimizing $Clip$ at time t for fluent f, the only part of $Clip$'s extension that will vary is the point $\langle t, f\rangle$ itself or $\langle t', f'\rangle$ for $t' > t$ and arbitrary f'. In addition, $Holds$ is varied only for time points $t' \geq t$.

7.2.3 Nonmonotonic Reasoning and Belief Revision

[DPN94] (ELP for revising KB), [MG89] (the postulates of [KLM90] translate to the postulates of Belief Revision in [AGM85] and vice versa), [GM94a][MG89][Bou94b] (generating a preference relation (in the sense of [Sho88],[KLM90]) from a set of beliefs - views both BR and NMR with the same machinery), [GP92b] (ranking systems for BR and NMR), [BG93] (BR and conditional logic), [PL97] (Belief Revision and Circumscription).

7.3 Applications

In [McC86] a typology of intended uses of nonmonotonic reasoning was presented. This typology included:

1. A communication convention (e.g., CBCL of [McC90a]).
2. A database convention (now days depicted in Prolog, Datalog, LP in general and Deductive Databases in particular).[38]
3. A rule of conjecture (a rule of probabilistic nature ("most birds fly") or standard cases (as in Minsky's *Frames*)).
4. A representation of policy ("The meeting will be on Wednesday, unless other decision is made").
5. A very streamlined expression of probabilistic information, when numerical probabilities as unobtainable.
6. Auto-epistemic reasoning - reasoning about our own knowledge (Moore 1985).
7. Commonsense Physics and commonsense Psychology.

Only a few of these were exploited so far. An attempt was done in KIF [GF92] to include nonmonotonic information, but since no consensus was reached as to the nature of these nonmonotonic rules in KIF, it is no longer given as part of the language.

Anecdotal applications include:

[38] Reiter has talked about it as well, discussing the Domain-Closure assumption and the Negation-as-Failure used in Databases and Logic Programs.

- [Fal95] uses circumscription for common-sense reasoning about spatial information.
- [KL95] formalized action using circumscription (and some filtering). [DL94] used Circumscription and Pointwise-Circumscription for that.
- [Nar96] achieves Elaboration Tolerance of KBs that are results of knowledge discovery in DBs (using "inheritance").

In an invited talk in KR'92, Reiter describes the places where Nonmonotonic Reasoning contributed to AI and Computer Science in general.

- Semantics for logic programs.
- Logic programming implementations of nonmonotonic reasoning
- Solutions to the frame and ramification problems in the situation calculus and their relevance to database update transactions, program verification and software specification.
- The remarkable convergence of probabilistic and nonmonotonic formalisms for default reasoning.
- Abduction and diagnostic reasoning.
- Foundations for TMSs.
- Nonmonotonicity and theories of Belief Revision.
- Introspection and reasoning about Belief.
- Educating the next generation of AI students.

[Mor90], [MG92] worked on NMR with multiple agents.

7.4 Dictionary

DCA (Domain Closure Axioms)

$$\forall x.(x = t_1 \vee ... \vee x = t_n)$$

This axiom entails that the domain is always finite.

UNA (Unique Names Axioms)

$$\forall i \neq j.(t_i \neq t_j)$$

This axiom entails that all of the symbols refer to distinct objects in the domain.

Together with the DCA, UNA implies the exact cardinality of the domain.

8 Conclusions

Now days it seems that the major use of nonmonotonic reasoning systems comes in *formalizing* different aspects of common-sense reasoning and belief revision. One may ask: "Why do we need to formalize these issues?" and the answers are split. One argument comes to say that by formalizing a topic we understand it better. An example of this is the Frame Problem. Before we go and write an algorithm we would better be clear about what we want. Another argument says that we actually want to use tools that are similar to theorem provers to drive our intelligent agents.

Regardless of the argument, it seems that the field of nonmonotonic reasoning advanced in several different directions, some outlined in this paper. The most obvious directions are extending and modifying earlier nonmonotonic systems, semantics for nonmonotonic reasoning, meta-rule systems and finding relationships between the systems and between nonmonotonic reasoning and other fields, such as belief revision, conditional logic, logic programs and databases. Many of the researchers in these areas seem to fail to see their work in a larger context, leading to separation and lack of flow of information between the different sub-communities.

Many challenges are standing in front of the nonmonotonic reasoning community. Aside from a few works on methodology, not many have focused on this difficult area ([EKP91] is an exception). The work on relationships between the systems is not done, and in my opinion is not even mature yet. Works like those of Imielinski [Imi87] and Gottlob [Got95] are in demand, as they shape the field and allow precise measurement of the limitations of our reasoning with a given system. The topic of applications is dominating in our times, and the nonmonotonic community does not escape its merit. This demand for applications should focus the field on evaluating what was learned so far and the directions that should be taken. The areas of learning and cognitive robotics are now emerging as possible directions nonmonotonic reasoning may take in the near future.

On a smaller scale, this work identified many open problems. The relationship between conditional logic and nonmonotonic entailment reltations is not understood yet, there is yet to be found one semantic framework that will encompass all of the nonmonotonic systems that we described in the beginning of the article, there is a need to account for the complete set of meta-rules (properties) for all the nonmonotonic systems, find a set of meta-rules that represents a given nonmonotonic system, account for the set of meta-rules true in a sublanguage of a given system (a step in that direction is [Eng98]), find extensions of current nonmonotonic systems that satisfy all kinds of properties, identify what kinds of entailment relations (and properties) can be defined using a certain nonmonotonic system and more. Emerging and current sibling-fields, such as probabilistic reasoning and economics, can also contribute and benefit from nonmonotonic theories.

A future version of this work will contain more on algorithms and computational complexity, methodologies, applications and other important results that

managed to elude us in this version.

References

[AAGP96] G. Amati, L. Carlucci Aiello, D. Gabbay, and F. Pirri. A proof theoretical approach to default reasoning i: Tableaux for default logic. *Journal of Logic Computation*, 6(2):205–231, 1996.

[Ada75] Ernest Wilcox Adams. *The Logic of Conditionals*. D. Reidel, Dordrecht, 1 edition, 1975.

[AE82] Krzysztof R. Apt and M. H. Van Emden. Contributions to the theory of logic programming. *Journal of the ACM*, 29(3):841–862, July 1982.

[AGM85] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[AK88] D. E. Appelt and K. Konolige. A nonmonotonic logic for reasoning about speech acts and belief revision. In M. Reinfrank, J. de Kleer, M. L. Ginsberg, and E. Sandewall, editors, *Non-Monotonic Reasoning — Proceedings of the Second International Workshop (LNAI Volume 346)*, pages 164–175. Springer-Verlag: Heidelberg, Germany, 1988.

[Ami98a] Eyal Amir. Nonmonotonic Reasoning as SAT Search. Unpublished note. Available at http://www-formal.stanford.edu/eyal/nmr/nmr-search.ps, 1998.

[Ami98b] Eyal Amir. Pointwise circumscription revisited. In *Principles of Knowledge Representation and Reasoning: Proc. Sixth International Conference (KR '98)*, pages 202–210, 1998.

[Ant96] Grigoris Antoniou. A comparative survey of default logic variants. In Dov M. Gabbay and Hans Jürgen Ohlbach, editors, *Proceedings of the International Conference on Formal and Applied Practical Reasoning (FAPR-96)*, volume 1085 of *LNAI*, pages 15–28, Berlin, June3–7 1996. Springer.

[Ant97] Grigoris Antoniou. *Nonmonotonic Reasoning*. MIT Press, Cambridge, Massachusetts, 1997.

[BDK97] Gerd Brewka, Jürgen Dix, and Kurt Konolige. *Nonmonotonic Reasoning: An Overview*, volume 73 of *CSLI Lecture Notes*. CSLI Publications, 1997.

[BDKT97] A. Bondarenko, P.M. Dung, R.A. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93(1-2):63–101, 1997.

[Ben89] Johan Van Benthem. Modal logic as a theory of information. In *Proceedings of the Prior Memorial Conference*, 1989.

[BF78] Bruce G. Buchanan and Edward A. Feigenbaum. Dendral and metadendral, their applications dimension. *Artificial Intelligence: An International Journal*, 11:5–24, 1978.

[BF91] N. Bidoit and C. Froidevaux. General logic databases and programs: Default logic semantics and stratification. *Information and Computation*, 91:85–112, 1991.

[BG93] Craig Boutilier and Moisés Goldszmidt. Revision by conditional beliefs. In *Proc. National Conference on Artificial Intelligence (AAAI '93)*, pages 649–654. Morgan Kaufmann, Washington, D.C., July 1993.

[BGHK93] Fahiem Bacchus, Adam J. Grove, Joseph Y. Halpern, and Daphne Koller. Statistical foundations for default reasoning. In Ruzena Bajcsy, editor, *Proc. Thirteenth International Joint Conference on Artificial Intelligence (IJCAI '93)*, pages 563–569. Morgan Kaufmann, Chambéry, France, August 1993.

[BMM89] Philippe Besnard, Yves Moinard, and Robert E. Mercer. The importance of open and recursive circumscription. *Artificial Intelligence*, 39(2):251–262, June 1989.

[BMP97] Hassan Bezzazi, David Makinson, and Ramón Pino Pérez. Beyond rational monotony: Some strong non-Horn rules for nonmonotonic inference relations. *Journal of Logic and Computation*, 7(5):605–631, October 1997.

[Bon91] P. A. Bonatti. A family of three-valued autoepistemic logics. *Lecture Notes in Computer Science*, 549:28–??, 1991.

[Bon95] Piero A. Bonatti. Autoepistemic logics as a unifying framework for the semantics of logic programs. *Journal of Logic Programming*, 22(2):91–149, February 1995.

[Bou94a] Craig Boutilier. Conditional logics of normality: a modal approach. *Artificial Intelligence*, 68:87–154, 1994.

[Bou94b] Craig Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68(1):33–85, 1994.

[BQQ83] Philippe Besnard, Rene Quiniou, and Patrice Quinton. A theorem-prover for a decidable subset of default logic. In Michael A. Genesereth, editor, *Proceedings of the National Conference on Artificial Intelligence*, pages 27–30, Washington, D.C., August 1983. AAAI Press.

[Bre91a] Gerhard Brewka. Cumulative Default Logic: In Defense of Nonmonotonic Inference Rules. *Artificial Intelligence*, 50:183–205, 1991.

[Bre91b] Gerhard Brewka. *Nonmonotonic Reasoning: Logical Foundations of Common Sense*. Cambridge University Press, 1991.

[Bre94] Gerhardt Brewka. Reasoning about priorities in default logic. In Barbara Hayes-Roth and Richard Korf, editors, *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 940–945, Menlo Park, CA, 1994. American Association for Artificial Intelligence, AAAI Press.

[BS85] G. Bossu and P. Siegel. Saturation, nonmonotonic reasoning and the closed world assumption. *Artificial Intelligence*, 25(1):13–65, 1985.

[BS88] Bruce Buchanan and Reid Smith. Fundamentals of expert systems. *Annual Review of Computer Science*, 3:23–58, 1988.

[Cad92] M. Cadoli. The complexity of model checking for circumscriptive formulae. *Information Processing Letters*, 44(3):113–??, November 1992.

[Cad95] Marco Cadoli. *Tractable reasoning in artificial intelligence*, volume 941 of *Lecture Notes in Artificial Intelligence and Lecture Notes in Computer Science*. Springer-Verlag Inc., New York, NY, USA, 1995.

[CEG92] M. Cadoli, T. Eiter, and G. Gottlob. An efficient method for eliminating varying predicates from a circumscription. *Artificial Intelligence*, 54(3):397–410, April 1992.

[Che94] Jianhua Chen. The logic of only knowing as a unified framework for non-monotonic reasoning. *Fundamenta Informaticae*, 21(3):205–220, 1994.

[CL92] Gabriella Crocco and Philippe Lamarre. On the connection between non-monotonic inference systems and conditional logics. In William Nebel, Bernhard; Rich, Charles; Swartout, editor, *Proc. Third International Conference on Principles of Knowledge Representation and Reasoning (KR '92)*, pages 565–571, Cambridge, MA, October 1992. Morgan Kaufmann.

[Cla78] K. L. Clark. Negation as failure. In H. Gallaire and J. Minker, editors, *Logic and Databases*, pages 293–322, New York, 1978. Plenum Press.

[Coh63] P. J. Cohen. The independense of the continuum hypothesis. *Proceedings of the National Academy of Science*, 50:1143–1148, 1963.

[Coh88] J. Cohen. A view of the origins and development of PROLOG. *Communications of the Association for Computing Machinery*, 31:26–36, 1988. Cohen, J.

[CPS93] M. Caryol, O. Palmade, and T. Schiex. A fixed-point semantics for the atms. *Journal of Logic and Computation*, 3(2):115–130, 1993.

[CS92] M. Cadoli and M. Schaerf. Approximate inference in default logic and circumscription. In Bernd Neumann, editor, *Proceedings of the 10th European Conference on Artificial Intelligence*, pages 319–323, Vienna, Austria, August 1992. John Wiley & Sons.

[CS93] Marco Cadoli and Marco Schaerf. A survey of complexity results for non-monotonic logics. *Journal of Logic Programming*, 17(2, 3 and 4):127–160, November 1993.

[dCHL94] L. Farinas del Cerro, A. Herzig, and J. Lang. From Expectation-Based Nonmonotonic Reasoning to Conditional Logics. *Artificial Intelligence*, 66:375–393, 1994.

[de 86] Johan de Kleer. Problem solving with the ATMS. *Artificial Intelligence*, 28:197–224, 1986.

[DG93a] J.A.Robinson D.M. Gabbay, C.J.Hogger, editor. *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 1: Logical Foundations.* Oxford University Press, Great Britain, 1993.

[DG93b] J.A.Robinson D.M. Gabbay, C.J.Hogger, editor. *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning.* Oxford University Press, Great Britain, 1993.

[DG93c] J.A.Robinson D.M. Gabbay, C.J.Hogger, editor. *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 4: Epistemic and Temporal Reasoning.* Oxford University Press, Great Britain, 1993.

[DG93d] J.A.Robinson D.M. Gabbay, C.J.Hogger, editor. *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 5: Logic Programming.* Oxford University Press, Great Britain, 1993.

[DJ91] James P. Delgrande and W. Ken Jackson. Default logic revisited. In James F. Allen, Richard Fikes, and Erik Sandewall, editors, *KR'91: Principles of Knowledge Representation and Reasoning*, pages 118–127. Morgan Kaufmann, San Mateo, CA, 1991.

[dK86a] Johan de Kleer. An assumption-based truth maintenance system. *Artificial Intelligence*, 28(1):127–162, 1986.

[dK86b] Johan de Kleer. Extending the ATMS. *Artificial Intelligence*, 28(1):163–196, 1986.

[dKK89] Johan de Kleer and Kurt Konolige. Eliminating the fixed predicates from a circumscription. *Artificial Intelligence*, 39(3):391–398, July 1989.

[DL92] Patrick Doherty and Witold Lukaszewicz. NML3 — A non-monotonic logic with explicit defaults. *Journal of Applied Non-Classical Logics*, 2:9–48, 1992.

[DL94] Patrick Doherty and Witold Lukaszewicz. Circumscribing features and fluents. In D. Gabbay and H.J. Ohlbach, editors, *Proceedings of the 1st int'l conf. on Temporal Logic*, pages 82–100, 1994.

[DLS95] Patrick Doherty, Witold Lukaszewicz, and Andrzej Szalas. Computing circumscription revisited: Preliminary report. In *Proceedings of the International Joint Conference on AI (IJCAI-95)*, pages 1502–1508, 1995.

[DLS97] Patrick Doherty, Witold Lukaszewicz, and Andrzej Szalas. Computing circumscription revisited: A reduction algorithm. *Journal of Automated Reasoning*, 18(3):297–336, June 1997.

[DMT98] M. Denecker, W. Marek, and M. Truszczyński. Fixpoint 3-valued semantics for autoepistemic logic. In *Proc. National Conference on Artificial Intelligence (AAAI '98)*, pages 840–845, 1998.

[Doy79] Jon Doyle. A truth maintenance system. *Artificial Intelligence: An International Journal*, 12(3):231–272, 1979.

[Doy85] Jon Doyle. Circumscription and implicit definability. *Journal of Automated Reasoning*, 1(4):391–405, 1985.

[DP86] Didier Dubois and Henri Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York, 1986.

[DP88] Didier Dubois and Henri Prade. Default reasoning and possibility theory. *Artificial Intelligence*, 35(2):243–257, June 1988.

[DP91] Didier Dubois and Henri Prade. Possibilistic logic, preferential models, non-monotonicity and related issues. In Ray Myopoulos, John; Reiter, editor, *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 419–425, Sydney, Australia, August 1991. Morgan Kaufmann.

[DPN94] Carlos Viegas Damásio, Luís Moniz Pereira, and Wolfgang Nejdl. REVISE: An extended logic programming system for revising knowledge bases. In Pietro Torasso Jon Doyle, Erik Sandewall, editor, *Principles of Knowledge Representation and Reasoning: Proc. Fourth International Conference (KR '94)*, pages 607–618, Bonn, FRG, May 1994. Morgan Kaufmann.

[DS94] James P. Delgrande and Torsten Schaub. A preference-based approach to default reasoning: Preliminary report. In *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94)*, 1994.

[DS97a] James P. Delgrande and Torsten Schaub. Compiling reasoning with and about preferences into default logic. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI-97)*, 1997.

[DS97b] James P. Delgrande and Torsten Schaub. Reasoning with sets of preferences in default logic. In *Second IJCAI Workshop on Nonmonotonic Reasoning, Action and Change*, 1997.

[DSJ94] James P. Delgrande, Torsten Schaub, and W. Ken Jackson. Alternative appraoches to default logic. *Artificial Intelligence*, 70:167–237, 1994.

[dW87] Johan de Kleer and Brian C. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32:97–130, 1987.

[EKP91] David E. Etherington, Sarit Kraus, and Donald Perlis. Nonmonotonicity and the scope of reasoning. *Artificial Intelligence*, 52:221–262, 1991.

[Elk90] Charles Elkan. A rational reconstruction of nonmonotonic truth maintenance systems. *Artificial Intelligence*, 43:219–234, 1990.

[EMR85] David W. Etherington, R. E. Mercer, and Ray Reiter. On the Adequacy of Predicate Circumscription for Closed World Reasoning. *Computational Intelligence*, 1, 1985.

[Eng98] Joeri Engelfriet. Monotonicity and persistence in preferential logics. *Journal of Artificial Intelligence Research*, 8:1–21, 1998.

[ER83] David Etherington and Raymond Reiter. On inheritance hierarchies with exceptions. In *Proceedings of the Third National Conference on Artificial Intelligence*, pages 104–108, 1983.

[ET96] J. Engelfriet and J. Treur. Specification of nonmonotonic reasoning. In *Practical Reasoning, Proceedings FAPR'96*, pages 111–125. Springer-Verlag, 1996. vol. 1085 of Lecture Notes in Artificial Intelligence.

[Eth86] David Etherington. *Reasoning with incomplete Information.* PhD thesis, University of British Columbia, 1986.

[Eth87a] David W. Etherington. Relating default logic and circumscription. In John McDermott, editor, *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 489–494, Milan, Italy, August 1987. Morgan Kaufmann.

[Eth87b] David W. Etherington. A semantics for default logic. In John McDermott, editor, *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 495–498, Milan, Italy, August 1987. Morgan Kaufmann.

[Fal95] Boi Faltings. Qualitative spatial reasoning using algebraic topology. In *International Conference on Spatial Information Theory, COSIT-95 Semmering (Austria) September 21-23, 1995*, 1995.

[Fef84] Solomon Feferman. Towards useful type-free theories. *Journal of Symbolic Logic*, 49(1):75–111, March 1984.

[Fei92] Edward Feigenbaum. A personal view of expert systems: Looking back and looking ahead. Technical Report KSL 92-41, Knowledge Systems Lab, Stanford University, April 1992.

[Fey37] Robert Feys. Les logiques nouvelles des modalités. *Revue Néoscholastique de Philosophie*, 40:517–553, 1937.

[FH96a] N. Friedman and J. Y. Halpern. Belief revision: A critique. In *Principles of Knowledge Representation and Reasoning, Proceedings of the fifth International Conference (KR '96)*. Morgan Kaumann, 1996.

[FH96b] Nir Friedman and Joseph Y. Halpern. Plausibility Measures and Default Reasoning. In *Proceedings AAAI-96*, pages 1297–1304. AAAI Press, 1996.

[FH98] Nir Friedman and Joseph Y. Halpern. Plausibility Measures and Default Reasoning. *Journal of the ACM*, 1998. To Appear. Available at http://www.cs.huji.ac.il/ nir/Papers/FrH5Full.ps.

[FHK96] Nir Friedman, Joseph Y. Halpern, and Daphne Koller. First-order conditional logic revisited. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference*, pages 1305–1312, Menlo Park, August4–8 1996. AAAI Press / MIT Press.

[FHMV95] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about knowledge.* MIT Press, 1995.

[Fit93] Melvin Fitting. Basic modal logic. In J.A.Robinson D.M. Gabbay, C.J.Hogger, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 1: Logical Foundations*. Oxford University Press, 1993.

[FL96] Michael Freund and Daniel Lehmann. On negation rationality. *Journal of Logic and Computation*, 6(2):263–269, 1996.

[Fla98] Peter A. Flach. Comparing consequence relations. In *Principles of Knowledge Representation and Reasoning: Proc. Sixth International Conference (KR '98)*, pages 180–189. Morgan Kaufmann, 1998.

[FLM91] Michael Freund, Daniel Lehmann, and Paul H. Morris. Rationality, transitivity, and contraposition. *Artificial Intelligence*, 52(2):191–203, December 1991. Research Note.

[For86] Kenneth D. Forbus. The qualitative process engine. *University of Illinois-Urbana - Dept. of Comp. Sci. TR UIUCDCS-R-86-1288, Revised version in Qualitative Reasoning about Physical Systems, ed. Daniel Weld and J. De Kleer, Morgan Kaufmann, 1990*, pages 220–235, 1986.

[Fre79] G. Frege. *Begriffsschrift, eine der Arithmetischen Nachgebildete Formelsprache des Reinen Denkens*. Halle, 1879. English translation in [JvH67], pp. 1–82.

[Fre93] Michael Freund. Injective models and disjunctive relations. *Journal of Logic and Computation*, 3(3):231–247, 1993.

[Fre98] Michael Freund. Preferential reasoning in the perspective of poole default logic. *Artificial Intelligence*, 98(1–2):209–235, 1998.

[Gab85] Dov M. Gabbay. Theoretical foundations for non-monotonic reasoning in expert systems. In Krzysztof R. Apt, editor, *Proc. of the NATO Advanced Study Institute on Logics and Models of Concurrent Systems*, pages 439–457, La Colle-sur-Loup, France, October 1985. Springer-Verlag.

[Gär78] Peter Gärdenfors. Conditionals and changes of belief. *Acta Philos. Fennica*, 30:381–404, 1978.

[Gär92] Peter Gärdenfors, editor. *Belief Revision*. Cambridge University Press, New York, 1992.

[Gef91] Hector Geffner. Beyond negation as failure. In James Allen, Richard Fikes, and Erik Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 218–229, San Mateo, CA, USA, April 1991. Morgan Kaufmann Publishers.

[Gef92a] Hector Geffner. *Default Reasoning.* MIT Press, Cambridge, Massachusetts, 1992.

[Gef92b] Hector Geffner. High-probabilities, model-preference and default arguments. *Minds and Machines*, 2(1):51–70, February 1992.

[Gel87] Michael Gelfond. On stratified autoepistemic theories. In Howard Forbus, Kenneth; Shrobe, editor, *Proceedings of the 6th National Conference on Artificial Intelligence*, pages 207–211, Seattle, WA, July 1987. Morgan Kaufmann.

[GF92] Michael Genesereth and Richard Fikes. Knowledge interchange format version 3.0 reference manual. Technical Report Logic-92-1, Dept Computer Science, Stanford University, June 1992.

[Gin87a] M. L. Ginsberg. *Readings in Non-monotonic Reasoning.* Morgan Kaufmann, 1987.

[Gin87b] Matthew L. Ginsberg. *Readings in Nonmonotonic Reasoning*, chapter 1, pages 1–23. Morgan Kaufmann, Los Altos, CA, 1987.

[Gin89] Matthew L. Ginsberg. A circumscriptive theorem prover. *Artificial Intelligence*, 39(2):209–230, June 1989.

[GL88] Michael Gelfond and Vladimir Lifschitz. The stable model semantics for logic programming. In *Logic Programming: Proceedings of the Fifth International Conference and Symposium*, pages 1070–1080, 1988.

[GL89] Michael Gelfond and Vladimir Lifschitz. Compiling Circumscriptive Theories into Logic Programs. In Reinfrank, de Kleer, Ginsberg, and Sandewall, editors, *Non-Monotonic Reasoning*, pages 74–99. Lecture Notes in Artificial Intelligence 346, Springer-Verlag, January 1989.

[GL90] Michael Gelfond and Vladimir Lifschitz. Logic Program with Classical Negation. In David H. D. Warren and Peter Szeredi, editors, *Proceedings of the 7th Int. Conf. on Logic Programming*, pages 579–597. MIT, June 1990.

[GL91] Michael Gelfond and Vladimir Lifschitz. Classical Negation in Logic Programs and Disjunctive Databases. *New Generation Computing*, 9, 1991.

[GLPT91] Michael Gelfond, Vladimir Lifschitz, Halina Przymusinska, and Miroslow Truszczyński. Disjunctive Defaults. In J. Allen, R. Fikes, and B. Sandewall, editors, *Proceedings of the second Conference on Principles of Knowledge Representation and Reasoning, Cambridge, Massachusetts.* Morgan Kaufmann, 1991.

[GM94a] Peter Gärdenfors and David Makinson. Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65(1):197–245, January 1994.

[GM94b] L. Giordano and A. Martelli. On cumulative default reasoning. *Artificial Intelligence*, 66:161–180, 1994.

[GMP90] M. Goldszmidt, P. Morris, and J. Pearl. A maximum entropy approach to nonmonotonic reasoning. In *Proc. National Conference on Artificial Intelligence (AAAI '90)*, volume 2, pages 646–652, Boston, Massachusetts, August 1990. MIT Press. Goldszmidt, M.Pearl, J.Morris, P.

[GMP93] M. Goldszmidt, P. Morris, and J. Pearl. A maximum entropy approach to nonmonotonic reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):220–232, March 1993. Goldszmidt, M.Pearl, J.Morris, P.

[GO92] Dov M. Gabbay and Hans Jürgen Ohlbach. Quantifier elimination in second–order predicate logic. In Bernhard Nebel, Charles Rich, and William Swartout, editors, *Principles of Knowledge Representation and Reasoning (KR92)*, pages 425–435. Morgan Kaufmann, 1992.

[Göd31] Kurt Gödel. Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.

[Göd67] K. Gödel. On formally undecidable propositions of principia mathematica and related systems. In J. Van Heijenoort, editor, *From Frege to Gödel: A Source Book in Mathematical Logic*, pages 592–617. Harvard University Press, 1967. A translation of [Göd31].

[Got93] Georg Gottlob. The power of beliefs or translating default logic into standard autoepistemic logic. In *IJCAI-93*, 1993.

[Got95] Georg Gottlob. Translating default logic into standard autoepistemic logic. *Journal of the ACM*, 42(4):711–740, 1995.

[GP86] M. Gelfond and H. Przymusinska. Negation as failure: Careful closure procedure. *Artificial Intelligence*, 30:273–287, 1986.

[GP92a] M. Goldszmidt and J. Pearl. System-Z^+: A formalism for reasoning with variable-strength defaults. In *Tenth National Conference on Artificial Intelligence*, pages 399–404, San Jose, California, 1992. American Association for Artificial Intelligence.

[GP92b] Moisés Goldszmidt and Judea Pearl. Rank-based systems: A simple approach to belief revision, belief update, and reasoning about evidence and actions. In William Nebel, Bernhard; Rich, Charles;

Swartout, editor, *Proc. Third International Conference on Principles of Knowledge Representation and Reasoning (KR '92)*, pages 661–672, Cambridge, MA, October 1992. Morgan Kaufmann.

[GPP86] M. Gelfond, H. Przymusinska, and T. C. Przymusinski. The extended closed world assumption and its relation to parallel circumscription. *Proc. Fifth ACM SIGACT-SIGMOD Symposium on Principle of Database Systems*, pages 133–139, 1986.

[GPP89] Michael Gelfond, Halina Przymusinska, and Teodor C. Przymusinski. On the relationship between circumscription and negation as failure. *Artificial Intelligence*, 38(1):75–94, February 1989.

[Gre69] Cordell Green. Theorem-proving by resolution as a basis for question-answering systems. In Bernard Meltzer, Donald Michie, and Michael Swann, editors, *Machine Intelligence 4*, pages 183–205. Edinburgh University Press, Edinburgh, Scotland, 1969.

[Gre87] G.H. Greco. Fuzzy integrals and fuzzy measures with their values in complete lattices. *Journal of Mathematical Analysis and Applications*, 126:594–603, 1987.

[Gro84] Benjamin N. Grosof. Default reasoning as circumscription. In *Workshop on Nonmonotonic Reasoning*, 1984.

[Gro91] Benjamin N. Grosof. Generalizing prioritization. In James Allen, Richard Fikes, and Erik Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 289–300, San Mateo, CA, USA, April 1991. Morgan Kaufmann Publishers.

[GRS91] Allen Van Gelder, Kenneth A. Ross, and John S. Schlipf. The well-founded semantics for general logic programs. *Journal of the ACM*, 38(3):620–650, July 1991.

[HA28] D. Hilbert and W. Ackermann. *Grundzüge der theoretischen Logik*. Springer, 1928.

[HA50a] D. Hilbert and W. Ackermann. *Principles of Mathematical Logic*. Chelsea, New York, 1950. A translation into English of [HA28].

[HA50b] D. Hilbert and W. Ackermann. *Principles of Mathematical Logic*. Chelsea, New York, 1950.

[Hal95] Joseph Y. Halpern. Reasoning about Knowledge. In J.A.Robinson D.M. Gabbay, C.J.Hogger, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol.4*. Oxford University Press, 1995.

[HB34] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*. Springer, Berlin, 1934.

[HB93] Bernhard Hollunder and Franz Baader. How to prefer more specific defaults in terminological default logic. In *IJCAI-93*, 1993.

[HC84] G.E. Hughes and M.J. Cresswell. *A Companion to Modal Logic*. Methuen, 1984.

[Hew72] Carl F. Hewitt. Description and theoretical analysis (using schemata) of planner: a language for proving theorems and manipulting models in a robot. Technical Report TR-258, MIT AI Laboratory, 1972.

[Hin62] K. Jaako J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell Press, Ithica, 1962.

[Hin88] Jaakko Hintikka. Model minimization – an alternative to circumscription. *Journal of Automated Reasoning*, 4(1):1–14, March 1988.

[HM85] J. Y. Halpern and Y. Moses. Towards a theory of knowledge and ignorance: preliminary report. In *Logics and Models of Concurrent Systems*, pages 459–476. Springer-Verlag, 1985. vol. F13 of NATO ASI series.

[Imi85] Tomasz Imielinski. Results on translating defaults to circumscription. In Aravind Joshi, editor, *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pages 114–120, Los Angeles, CA, August 1985. Morgan Kaufmann.

[Imi87] Tomasz Imielinski. Results on translating defaults to circumscription. *Artificial Intelligence*, 32(1):131–146, April 1987.

[IO96] Koji Iwanuma and Kazuhiko Oota. An extension of pointwise circumscription. *Artificial Intelligence*, 86(2):391–402, 1996.

[Jan96] T. Janhunen. Representing autoepistemic introspection in terms of default rules. In *Proceedings ECAI-96*, 1996.

[JK90] Ulrich Junker and Kurt Konolige. Computing the extensions of autoepistemic and default logics with a TMS. In *Proc. National Conference on Artificial Intelligence (AAAI '90)*, Boston, August 1990. presented also at the 3rd International Workshop on Non-Monotonic Reasoning, South Lake Tahoe, USA, May 1990.

[JvH67] Editor J. van Heijenoort, editor. *From Frege to Gödel, a Source Book in Mathematical Logic*. Harvard University Press, Cambridge, Massachusetts, 1967.

[Kam91] M. Kaminski. Embedding a default system into nonmonotonic logic. *Fundamenta Informaticae*, 14:345–354, 1991.

[Kan57] S. Kanger. On the characterization of modalities. *Theoria*, 23:152–155, 1957.

[KL95] G. N. Kartha and V. Lifschitz. A simple formalization of actions using circumscription. In *Proc. Thirteenth International Joint Conference on Artificial Intelligence (IJCAI '95)*, 1995.

[KLM90] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, 44(1):167–207, 1990.

[Kon84] Kurt Konolige. *A Deductive Model of Belief and its Logics.* PhD thesis, Stanford University, 1984.

[Kon88a] Kurt Konolige. Hierarchic autoepistemic theories for nonmonotonic reasoning. In *Proc. National Conference on Artificial Intelligence (AAAI '88)*, 1988.

[Kon88b] Kurt Konolige. On the relation between default and autoepistemic logic. *Artificial Intelligence*, 35(3):343–382, July 1988.

[Kon89] Kurt Konolige. On the Relations between Autoepistemic Logic and Circumscription – preliminary report. In *Proc. Eleventh International Joint Conference on Artificial Intelligence (IJCAI '89)*, pages 1213–1218, 1989.

[Kon92a] Kurt Konolige. Ideal introspective belief. In William Swartout, editor, *Proceedings of the 10th National Conference on Artificial Intelligence*, pages 635–641, San Jose, CA, July 1992. MIT Press.

[Kon92b] Kurt Konolige. Quantification in autoepistemic logic. *Fundamenta Informaticae*, 1992.

[Kon93] Kurt Konolige. Autoepistemic logic. In J.A.Robinson D.M. Gabbay, C.J.Hogger, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning.* Oxford University Press, 1993.

[Kow78] R. A. Kowalski. Logic for data description. In H. Gallaire J. Minker, editor, *Logic and Data Bases*, pages 77–102. Plenum Press, New York, 1978.

[Kow79] R. Kowalski. Algorithm = logic + control. *Communications of the ACM*, 22:424–436, 1979.

[KP90] P. G. Kolaitis and C. H. Papadimitriou. Some computational aspects of circumscription. *Journal of the ACM, JACM*, 37(1):1–14, January 1990.

[Kri63] Saul Kripke. Semantical considerations of model logic. *Zeitschrift fur Mathematishe Logik und Grundlagen der Mathematik*, 9:67–96, 1963.

[Kri65] S. A. Kripke. Semantical analysis of intuitionistic logic I. In J. N. Crossley and M. Dummett, editors, *Formal Systems and Recursive Functions*, pages 92–130. North-Holland, Amsterdam, 1965.

[KS89] Kautz and Bart Selman. The Complexity of Model-Preference Default-Theories. In Reinfrank, de Kleer, Ginsberg, and Sandewall, editors, *Non-Monotonic Reasoning*. Lecture Notes in Artificial Intelligence 346, Springer-Verlag, January 1989.

[KS91] H. A. Kautz and B. Selman. Hard problems for simple default logics. *Artificial Intelligence*, 49(1-3):243–279, January 1991.

[LB85] Levesque and Brachman. A fundamental tradeoff in knowledge representation and reasoning. In Ronald Brachman and Hector Levesque, editors, *Readings in Knowledge Representation*, pages 41–70. Morgan Kaufmann, 1985. Revised from a paper that appeared with the same title in Proc. Fourth Conference of the Canadian Society for Computational Studies of Intelligence.

[Lev84a] Hector Levesque. *A Formal Treatment of Incomplete Knowledge Bases*. PhD thesis, Department of Computer Science, University to Toronto, Ontario, 1984.

[Lev84b] Hector J. Levesque. Foundations of a Functional Approach to Knowledge Representation. *Artificial Intelligence*, 24:155–212, 1984.

[Lev90] H. J. Levesque. All i know: a study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.

[Lew73] David Lewis. *Counterfactuals*. Basil Blackwell, 1973.

[Lif85a] Vladimir Lifschitz. Closed-world databases and circumscription. *Artificial Intelligence*, 27:229–235, 1985.

[Lif85b] Vladimir Lifschitz. Computing circumscription. In *Proc. of IJCAI-85*, pages 121–127, 1985.

[Lif86a] Vladimir Lifschitz. On the Satisfiability of Circumscription. *Artificial Intelligence*, 28:17–27, 1986.

[Lif86b] Vladimir Lifschitz. Pointwise circumscription: Preliminary report. In *Proc. AAAI-86*, pages 406–410, 1986.

[Lif87a] V. Lifschitz. On the declarative semantics of logic programs with negation. In J. Minker, editor, *Foundations of deductive databases and logic programming*, pages 177–192, Los Altos, CA, 1987. Morgan Kaufmann.

[Lif87b] Vladimir Lifschitz. Pointwise circumscription. In Matthew Ginsberg, editor, *Readings in nonmonotonic reasoning*, pages 179–193. Morgan Kaufmann, San Mateo, CA, 1987.

[Lif89a] Vladimir Lifschitz. Benchmark problems for formal nonmonotonic reasoning. In *Nonmonotonic Reasoning: 2nd International Workshop (Lecture Notes in Artificial Intelligece 346)*, pages 202–217, 1989.

[Lif89b] Vladimir Lifschitz. Between circumscription and autoepistemic logic. In Hector J. Levesque Ronald J. Brachman and Raymond Reiter, editors, *Proc. First International Conference on Principles of Knowledge Representation and Reasoning (KR '89)*, pages 235–244, Toronto, Canada, May 1989. Morgan Kaufmann.

[Lif89c] Vladimir Lifschitz. Circumscriptive theories: A logic-based framework for knowledge representation. In R. H. Thomason, editor, *Philosophical Logic and Artificial Intelligence*, pages 109–159. Kluwer Academic Publishers, 1989.

[Lif91] V. Lifschitz. Nonmonotonic databases and epistemic queries: Preliminary report. In *Proceedings of International Joint Conference on Artificial Intelligence*, Sydney, Australia, 1991.

[Lif93] Vladimir Lifschitz. Circumscription. In J.A.Robinson D.M. Gabbay, C.J.Hogger, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford University Press, 1993.

[Lin88] F. Lin. Circumscription in a modal logic. *Proceedings*, TARK-88:113–127, 1988.

[LL59] C. Lewis and C. Langford. *Symbolic Logic*. Dover, second edition, 1959. First edition published in 1932.

[LL92] C. J. Liau and B. I. Lin. Abstract minimality and circumscription. *Artificial Intelligence*, 54(3):381–396, April 1992.

[Llo87] J. W. Lloyd. *Foundations of Logic Programming, Second Edition*. Springer-Verlag, 1987.

[LM92] Daniel Lehmann and Menachem Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

[LS90] Fangzhen Lin and Yoav Shoham. Epistemic Semantics for Fixed-Points Nonmonotonic Logics. In R. Kowalski and K. Bowen, editors, *Proceedings Theoretical Aspects of Reasoning About Knowledge*, 1990.

[Luk88] Witold Lukaszewicz. Considerations on Default Logic. *Computational Intelligence*, 4, 1988.

[Luk90] Witold Lukaszewicz. *Nonmonotonic Reasoning: Formalization of Common Sense Reasoning.* Ellis Harwood, 1990.

[Mak89] David Makinson. General Theory of Cumulative Inference. In Reinfrank, de Kleer, Ginsberg, and Sandewall, editors, *Non-Monotonic Reasoning.* Lecture Notes in Artificial Intelligence 346, Springer-Verlag, January 1989.

[Mak93] David Makinson. General patterns in nonmonotonic reasoning. In J.A.Robinson D.M. Gabbay, C.J.Hogger, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning.* Oxford University Press, 1993.

[McA80] D. A. McAllester. An outlook on truth maintenance. (MIT) Artificial Intelligence Memo 551, Department of Computer Science, Massachusettes Institute of Technology, August 1980.

[McA90] David McAllester. Truth maintenance. In *Proc. National Conference on Artificial Intelligence (AAAI '90)*, pages 1109–1116, 1990.

[McC58] John McCarthy. Programs with Common Sense. In *Mechanisation of Thought Processes, Proceedings of the Symposium of the National Physics Laboratory*, pages 77–84, London, U.K., 1958. Her Majesty's Stationery Office. Reprinted in McC90.

[McC77] John McCarthy. Epistemological problems in artificial intelligence. In *Proc.* 5th *International Conference on Artificial Intelligence*, pages 1038–1044, 1977.

[McC79] John McCarthy. First Order Theories of Individual Concepts and Propositions. In Donald Michie, editor, *Machine Intelligence*, volume 9. Edinburgh University Press, Edinburgh, 1979. Reprinted in [McC90b].

[McC80] John McCarthy. Circumscription—A Form of Non-Monotonic Reasoning. *Artificial Intelligence*, 13:27–39, 1980. Reprinted in [McC90b].

[McC86] John McCarthy. Applications of Circumscription to Formalizing Common Sense Knowledge. *Artificial Intelligence*, 28:89–116, 1986. Reprinted in [McC90b].

[McC90a] John McCarthy. Common business communication language. In Vladimir Lifschitz, editor, *Formalizing Common Sense.* Ablex Publishing Corporation, Norwood, New Jersey, 1990.

[McC90b] John McCarthy. *Formalization of common sense, papers by John McCarthy edited by V. Lifschitz.* Ablex, 1990.

[McD82] D. V. McDermott. Nonmonotonic logic II: nonmonotonic model theories. *Journal of the ACM*, 29(1):33–57, 1982.

[MD80] D. McDermott and J. Doyle. Non-monotonic logic I. *Artificial Intelligence*, 13(1–2):41–72, April 1980. McDermott, D.Doyle, J.

[MG89] David Makinson and Peter Gärdenfors. Relations between the logic of theory change and nonmonotonic logic. In André Fuhrmann and Michael Morreau, editors, *Proceedings of the Workshop on The Logic of Theory Change*, volume 465 of *LNCS*, pages 185–205, Konstanz, FRG, October, 13–15 1989. Springer-Verlag.

[MG92] L. Morgenstern and R. Guerreiro. Epistemic logics for multiple-agent nonmonotonic reasoning I. In *Symposium on Formal Reasoning about Beliefs, Intentions, and Actions*, 1992.

[MH69] John McCarthy and Patrick J. Hayes. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In B. Meltzer and D. Michie, editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, 1969.

[Min75] M. Minsky. A framework for representing knowledge, 1975. Winston.

[Min82] Jack Minker. On indefinite databases and the closed world assumption. In D. W. Loveland, editor, *6th Conference on Automated Deduction*, volume 138 of *Lecture Notes in Computer Science*, pages 292–308. Springer-Verlag, New York, June 1982.

[Min92] Grigori E. Mints. *A Short Introduction to Modal Logic*, volume 30 of *CSLI Lecture Notes.* CSLI Publications, 1992.

[MLR91] Jack Minker, Jorge Lobo, and Arcot Rajasekar. Circumscription and Disjunctive Logic Programming. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation. Papers in Honor of John McCarthy*, pages 281–304. Academic Press, 1991.

[Moi88] Yves Moinard. Pointwise circumscription is equivalent to predicate completion (sometimes). In *Proceedings 5th International symposium on Logic Programming*, pages 1097–1105, 1988.

[Moi94] Yves Moinard. Reasoning by cases without contraposition in default logic. In A. G. Cohn, editor, *Proceedings of the Eleventh European Conference on Artificial Intelligence*, pages 381–385, Chichester, August8–12 1994. John Wiley and Sons.

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

[Moo87] Robert C. Moore. Possible-world semantics for autoepistemic logic. In Matthew Ginsberg, editor, *Readings in nonmonotonic reasoning*, pages 137–142. Morgan Kaufmann, San Mateo, CA, 1987.

[Mor89] P. H. Morris. Autoepistemic stable closure and contradiction resolution. In M. Reinfrank, J. de Kleer, M. L. Ginsberg, and E. Sandewall, editors, *Non-monotonic reasoning*, pages 176–186. Berlin: Springer-Verlag, 1989. Lecture Notes in Artificial Intelligence, 346.

[Mor90] L. Morgenstern. A theory of multiple agent nonmonotonic reasoning. In *aaai90*, pages 538–544, 1990.

[Mor96] Leora Morgenstern. Inheriting well-formed formulae in a formula-agumented semantic network. In Luigia Carlucci Aiello, Jon Doyle, and Stuart Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proc. Fifth International Conference (KR '96)*, pages 268–279, San Francisco, November 5–8 1996. Morgan Kaufmann.

[Mor98] Leora Morgenstern. Inheritance comes of age: Applying nonmonotonic techniques to problems in industry. *Artificial Intelligence*, 103(1–2):237–271, 1998.

[Mot87] Peter L. Mott. A theorem on the consistency of circumscription. *Artificial Intelligence*, 31(1):87–98, January 1987.

[MP84a] J. Minker and D. Perlis. Applications of protected circumscription. *Proc. Conference on Automated Deduction*, May 1984.

[MP84b] J. Minker and D. Perlis. Protected circumscription. *Proc. Workshop on Non-Monotonic Reasoning*, pages 337–343, October 17-19, 1984.

[MS94] Halina Przymusinskca Michael Gelfond, Vladimir Lifschitz and Grigori Schwarz. Autoepistemic logic and introspective circumscription. In Ronald Fagin, editor, *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 197–207, San Francisco, CA, USA, March 1994. Morgan Kaufmann Publishers Inc.

[MST91] Wiktor Marek, Grigori F. Shvarts, and Miroslaw Truszczyński. Modal nonmonotonic logics: Ranges, characterization, computation. In James Allen, Richard Fikes, and Erik Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 395–404, San Mateo, CA, USA, April 1991. Morgan Kaufmann Publishers.

[MT89a] W. Marek and M. Truszczyński. Relating autoepistemic and default logic. In *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning (KR-89)*, pages 276–288, Toronto, Ontario, 1989.

[MT89b] W. Marek and M. Truszczyński. Stable semantics for logic programs and default theories. In *Proceedings of the North American Conference on Logic Programming*, pages 243–259. MIT Press, 1989.

[MT90] W. Marek and M. Truszczyński. Modal logic for default reasoning. *Annals of Mathematics and Artificial Intelligence*, 1:275–302, 1990.

[MT93a] Viktor M. Marek and Mirek Truszczyński. *Nonmonotonic Logics; Context-Dependent Reasoning*. Springer Verlag, Berlin-Heidelberg-New York, 1st edition, 1993.

[MT93b] A. Mikitiuk and M. Truszczyński. Rational default logic and logic programming with consistency constraints. In A. Nerode and L. Pereira, editors, *Logic programming and non-monotonic reasoning*. MIT Press, 1993.

[Naq86] S. A. Naqvi. Some extensions to the closed world assumption in databases. In G. Ausiello and P. Atzeni, editors, *Proceedings of the International Conference on Database Theory*, pages 341–348, Rome, Italy, September 1986. Springer-Verlag.

[Nar96] A. Narayanan. Revisable knowledge discovery in databases. *International J. Of Intelligent Systems*, 11:75–96, 1996.

[Nej91] Wolfgang Nejdl. The P-systems: A systematic classification of logics of nonmonotonicity. In Kathleen Dean, Thomas L.; McKeown, editor, *Proceedings of the 9th National Conference on Artificial Intelligence*, pages 366–372. MIT Press, July 1991.

[Nie91] Ilkka N. F. Niemelä. Constructive tightly grounded autoepistemic reasoning. In Ray Myopoulos, John; Reiter, editor, *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 399–405, Sydney, Australia, August 1991. Morgan Kaufmann.

[PA95] M. Pakkan and V. Akman. Issues in Commonsense Set Theory. *Artificial Intelligence Review*, 8:279–308, 1995.

[PAA91] L. Moniz Pereira, J. J. Alferes, and J. N. Aparicio. A practical introduction to well-founded semantics. In *Proceedings of the Third Scandinavian Conference on Artificial Intelligence*, 1991.

[Pea89] Judea Pearl. Probabilistic semantics for nonmonotonic reasoning: A survey. In Hector J. Levesque Ronald J. Brachman and Raymond Reiter, editors, *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 505–516, Toronto, Canada, May 1989. Morgan Kaufmann.

[Pea90] J. Pearl. System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning. In Rohit Parikh, editor,

Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge, pages 121–136, Pacific Grove, CA, March 1990. Morgan Kaufmann.

[Per85] D. Perlis. Languages with self-reference I: Foundations. *Artificial Intelligence*, 25(3):301–322, 1985.

[Per87] Donald Perlis. Circumscribing with sets. *Artificial Intelligence*, 31:201–211, 1987.

[Per88] D. Perlis. Autocircumscription. *Artificial Intelligence*, 36:223–236, 1988.

[PL97] Marco Schaerf Paolo Liberatore. Reducing belief revision to circumscription (and vice versa). *Artificial Intelligence*, 93(1-2):261–296, 1997.

[PM86] D. Perlis and John Minker. Completeness Results for Circumscription. *Artificial Intelligence*, 28:29–42, 1986.

[Poo88a] D. Poole. Representing knowledge for logic-based diagnosis. In Institute for New Generation Computer Technology (ICOT), editor, *Proceedings of the International Conference on Fifth Generation Computer Systems. Volume 3*, pages 1282–1290, Berlin, FRG, November28 December–2 1988. Springer Verlag.

[Poo88b] David Poole. A logical framework for default reasoning. *Artificial Intelligence*, 36(1):27–47, August 1988.

[Poo93] David Poole. Default logic. In J.A.Robinson D.M. Gabbay, C.J.Hogger, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford University Press, 1993.

[Pop94] Sally Popkorn. *First Steps in Modal Logic*. Cambridge University Press, 1994.

[Pro88] G. Provan. *Complexity Analysis of Truth Maintenance Systems, with Application to High Level Vision*. PhD thesis, Department of Mathematics, University of Oxford, 1988.

[Prz88] T. C. Przymusinski. On the declarative semantics of deductive databases and logic programming. In J. Minker, editor, *Foundations of Deductive Databases and Logic Programming*, pages 193–216. Morgan Kaufmann Pub., Washington, D.C., 1988.

[Prz89a] T. Przymusiński. Three-valued nonmonotonic formalisms and logic programming. In *Proceedings of the 1st international conference on principles of knowledge representation and reasoning, KR '89*, pages 341–348, San Mateo, CA., 1989. Morgan Kaufmann.

[Prz89b] Teodor C. Przymusinski. An algorithm to compute circumscription. *Artificial Intelligence*, 38(1):49–73, February 1989.

[Prz94] Teodor C. Przymusinski. Well-founded and stationary models of logic programs. *Annals of Mathematics and Artificial Intelligence*, 12:141–187, 1994.

[Rab89] Arkady Rabinov. A generalization of collapsible cases of circumscription. *Artificial Intelligence*, 38(1):111–117, February 1989.

[RC81] Raymond Reiter and Giovanni Criscuolo. On interacting defaults. In Ann Drinan, editor, *Proceedings of the Seventh IJCAI Conference*, pages 270–276, Vancouver, Canada, August 1981.

[RDB89] M. Reinfrank, O. Dressler, and G. Brewka. On the relation between truth maintenance and non-monotonic logics. In *Proceedings of IJCAI-89*, pages 1206–1212, San Mateo, CA., 1989. Morgan Kaufmann.

[RdK87] Raymond Reiter and Johan de Kleer. Foundations of assumption-based truth maintenance systems: Preliminary report. In *Proc. AAAI-87*, pages 183–188, 1987.

[Rei78] Raymond Reiter. On closed world data bases. In H. Gallaire and J. Minker, editors, *Logic and Data Bases*, pages 55–76. Plenum, New York / London, 1978.

[Rei80] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13 (1–2):81–132, 1980.

[Rei82] R. Reiter. Circumscription implies predicate completion (sometimes). *Proc. National Conference on Artificial Intelligence (AAAI '82)*, pages 418–420, 1982.

[Rei87] Ray Reiter. Nonmonotonic reasoning. *Annual Review of Computer Science*, pages 147–187, 1987.

[Rei98] Raymond Reiter. *Knowledge in Action: Logical Foundations for Describing and Implementing Dynamical Systems*. National Academy Press, 1998. in preparation.

[RLM89] Arcot Rajasekar, Jorge Lobo, and Jack Minker. Weak generalized closed world assumption. *Journal of Automated Reasoning*, 5(3):293–308, September 1989.

[Rob65] J.A. Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, 1965.

[Roo98] Nico Roos. Reasoning by cases in default logic. *Artificial Intelligence*, 99, 1998.

[RWM90] P. Rathmann, M. Winslett, and M. Manasse. Equality and circumscription. Technical report, University of Illinois at Urbana-Champaign, 1990.

[Ryc91] R. Rychlik. Some variations on default logic. In *Proceedings AAAI 91*, pages 183–188, 1991.

[San94] E. Sandewall. *Features and Fluents*. Oxford University Press, 1994.

[Sat90] Ken Satoh. A probabilistic interpretation for lazy nonmonotonic reasoning. In *Proceedings of AAAI-90*, pages 659–664, Boston, August 1990.

[Sat92] Ken Satoh. A probabilistic interpretation for lazy nonmonotonic reasoning. *Journal of Information Processing*, 15:29–39, 1992.

[Sch86] John S. Schlipf. How uncomputable is general circumscription? (extended abstract). In *Proceedings, Symposium on Logic in Computer Science*, pages 92–95, Cambridge, Massachusetts, 16–18 June 1986. IEEE Computer Society.

[Sch92] Torsten Schaub. On constrained default theories. In *Proceedings of the 10th European Conference on Artificial Intelligence*, pages 304–308, 1992.

[Sch95a] K. Schlechta. A two-stage approach to first order default reasoning. In Christine Froidevaux and Jürg Kohlas, editors, *Proceedings of the ECSQARU European Conference on Symbolic and Quantitive Approaches to Reasoning and Uncertainty*, volume 946 of *LNAI*, pages 379–386, Berlin, July 1995. Springer Verlag.

[Sch95b] J. S. Schlipf. The expressive powers of the logic programming semantics. In ACM, editor, *PODS '95. Proceedings of the Fourteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, volume 14 of *Proceedings of the ACM SIGACT SIGMOD SIGART Symposium on Principles of Database Systems*, pages 64–86, New York, NY 10036, USA, 1995. ACM Press.

[Sch96] Grigori Schwarz. On embedding default logic into moore's autoepistemic logic. *Artificial Intelligence*, 80(2):345–359, 1996.

[Sch97] Karl Schlechta. *Nonmonotonic logics: basic concepts, results, and techniques*, volume 1187 of *Lecture Notes in Computer Science and Lecture Notes in Artificial Intelligence*. Springer-Verlag Inc., New York, NY, USA, 1997.

[Sha76] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.

[Sha90a] G. Shafer. Perspectives on the theory and practice of belief functions. *International Journal of Approximate Reasoning*, 1990.

[Sha90b] Glenn Shafer. Belief functions. In Glenn Shafer and Judea Pearl, editors, *Readings in Uncertain Reasoning*. Morgan Kaufmann, San Mateo, California, 1990.

[Sha97] Murray Shanahan. *Solving the Frame Problem, a mathematical investigation of the common sense law of inertia*. MIT press, Cambridge, MA, 1997.

[She84] John C. Shepherdson. Negation as finite failure: A comparison of Clark's completed database and Reiter's closed world assumption. *Journal of Logic Programming*, 1(15):51–79, 1984.

[Sho67] Joseph R. Shoenfield. *Mathematical Logic*. Addison-Wesley, Duke University, 1967.

[Sho87] Y. Shoham. Nonmonotonic logics: Meaning and utility. In *Proc. Tenth International Joint Conference on Artificial Intelligence (IJCAI '87)*, pages 388–392. Morgan Kaufmann, 1987.

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

[Shv90] G. F. Shvarts. Autoepistemic modal logics. In R. Parikh, editor, *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the Third Conference*, pages 97–109, San Francisco, 1990. Morgan Kaufmann.

[SK90] Bart Selman and Henry A. Kautz. Model-preference default theories. *Artificial Intelligence*, 45(3):287–322, October 1990.

[Spo87] W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In W. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change and Statistics*, volume volume 2, pages 105–134. D. Reidel, Dordrecht, 1987.

[SS94a] Erik Sandewall and Yoav Shoham. Nonmonotonic Temporal Reasoning. In D.M. Gabbay, C.J.Hogger, and J.A.Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 4: Epistemic and Temporal Reasoning*. Oxford University Press, 1994.

[SS94b] Leon Sterling and Ehud Shapiro. *The Art of Prolog*. The MIT Press, Cambridge, Mass., second edition, 1994.

[Ten97] Choh Man Teng. Many Default Logics. In NRAC-97' and in http://u2.newcastle.edu.au/nrac/Teng.ps, 1997.

[THT86] Richmond Thomason, Jeff Horty, and D. S. Touretzky. A Calculus for Inheritance in Monotonic Semantic Nets. Research Note CMU-CS 86-138, Carnegie Mellon, 1986.

[Tou86] D. S. Touretzky. *The Mathematics of Inheritance.* PhD thesis, Carnegie Mellon, 1986.

[Tru91a] M. Truszczyński. Embedding default logic into modal nonmonotonic logics. In W. Marek A. Nerode and V. S. Submarahmanian, editors, *Logic programming and non-monotonic reasoning. Proceedings of the first international workshop*, pages 151–165, Cambridge, MA., 1991. MIT Press.

[Tru91b] Miroslaw Truszczynski. Modal interpretations of default logic. In Ray Myopoulos, John; Reiter, editor, *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 393–398, Sydney, Australia, August 1991. Morgan Kaufmann.

[VK76] M. H. Van Emden and R. A. Kowalski. The semantics of predicate logic as a programming language. *Journal of the ACM*, 23(4):733–742, October 1976.

[vW51] G. von Wright. *An Essay in Modal Logic.* North-Holland, 1951.

[W. 35] W. Ackermann. Zum Eliminationsproblem der mathematischen Logik. *Mathematische Annalen*, 111:61–63, 1935.

[Wai92] Jacques Wainer. Combining circumscription and modal logic. In Paul Rosenbloom and Peter Szolovits, editors, *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 648–653, Menlo Park, CA, 1992. American Association for Artificial Intelligence, AAAI Press.

[Web91] S. Weber. Uncertainty measures, decomposability and admissibility. *Fuzzy Sets and Systems*, 40:395–405, 1991.

[Win85] Terry Winograd. Frame representations and the declarative/procedural controversy. In Ronald Brachman and Hector Levesque, editors, *Readings in Knowledge Representation*, pages 357–370. Morgan Kaufmann, 1985.

[Woo75] William A. Woods. What's in a link: Foundations for semantic networks. In D. G. Bobrow and A. M. Collins, editors, *Representation and Understanding.* Academic Press, 1975.

[YH85] Adnan H. Yahya and Lawrence J. Henschen. Deduction in non-Horn databases. *Journal of Automated Reasoning*, 1(2):141–160, 1985.

[YY93] Li Yan Yuan and Jia-Huai You. Autoepistemic circumscription and logic programming. *Journal of Automated Reasoning*, 10(2):143–160, April 1993.

This is /u/eyal/ongo/nmr/compar/nmr-compar.tex.