Deutsches
Forschungszentrum
fur Kunstliche
Intelligenz GmbH

Research
Report
RR-94-17

Philosophical Logics—A Survey and a Bibliography

Georg Struth

May 1994

Deutsches Forschungszentrum fur Kunstliche Intelligenz GmbH

Postfach 20 80
67608 Kaiserslautern, FRG
Tel.: + 49 (631) 205-3211
Fax: + 49 (631) 205-3210

Stuhlsatzenhausweg 3
66123 Saarbrucken, FRG
Tel.: + 49 (681) 302-5252
Fax: + 49 (681) 302-5341

Deutsches Forschungszentrum für Künstliche Intelligenz

The German Research Center for Artificial Intelligence (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI) with sites in Kaiserslautern and Saarbrücken is a non-profit organization which was founded in 1988. The shareholder companies are Atlas Elektronik, Daimler-Benz, Fraunhofer Gesellschaft, GMD, IBM, Insiders, Mannesmann-Kienzle, Sema Group, Siemens and Siemens-Nixdorf. Research projects conducted at the DFKI are funded by the German Ministry for Research and Technology, by the shareholder companies, or by other industrial contracts.

The DFKI conducts application-oriented basic research in the field of artificial intelligence and other related subfields of computer science. The overall goal is to construct systems with technical knowledge and common sense which - by using AI methods - implement a problem solution for a selected application area. Currently, there are the following research areas at the DFKI:

- Intelligent Engineering Systems
- Intelligent User Interfaces
- Computer Linguistics
- Programming Systems
- Deduction and Multiagent Systems
- Document Analysis and Office Automation.

The DFKI strives at making its research results available to the scientific community. There exist many contacts to domestic and foreign research institutions, both in academy and industry. The DFKI hosts technology transfer workshops for shareholders and other interested groups in order to inform about the current state of research.

From its beginning, the DFKI has provided an attractive working environment for AI researchers from Germany and from all over the world. The goal is to have a staff of about 100 researchers at the end of the building-up phase.

Dr. Dr. D. Ruland
Director

Philosophical Logics—A Survey and a Bibliography

Georg Struth

DFKI-RR-94-17

This work has been supported by a grant from The Federal Ministry for Research and Technology (FKZ ITWM-9201).

Philosophical Logics—A Survey and a Bibliography

Georg Struth

November 2, 1994

Abstract

Intensional logics attract the attention of researchers from differing academic backgrounds and various scientific interests. My aim is to sketch the philosophical background of alethic, epistemic, doxastic and deontic logics, their formal and metaphysical presumptions and their various problems and paradoxes, without attempting formal rigor. A bibliography, concise on philosophical writings, is meant to allow the reader's access to the maze of literature in the field.

Contents

1 Introduction **3**

2 A Metaphysical Prelude **5**

2.1 Historical Remarks and Basic Concepts 5
2.2 Frege Semantics and General Intensionality 8

3 Propositional Modal Logics **12**

3.1 General Remarks 12
3.2 Alethic Modal Logics 15
3.3 Epistemic Modal Logics 24
3.4 Deontic Modal Logics 28

4 Quantificational Modal Logics **36**

4.1 General Remarks 36
4.2 The Maze of Metaphysics 37
4.3 Rigid Designators 40
4.4 Non-Rigid Designators 41
4.5 Further Alternatives—The Call for Free Logics 44

5 Conclusion **45**

1 Introduction

Only a few decades ago, the concept 'modality' was known only to philosophers, linguists and logicians with philosophical inclinations. Today, the situation is completely changed. Cognitive science and especially artificial intelligence are now the sciences which have predominant interest in modalities, at least when formal theories of modalities are concerned. Since Hilary Putnam proposed his *functionalism* thesis, many philosophers, among them Jerry Fodor, Fred Dretske, and Noam Chomsky, believe that the computer is the adequate paradigm for the study of complex entities like knowledge, belief, thought and language learning. One therefore might expect an intense interaction between different academic domains in favor of synergetic effects, but a closer examination of the AI literature reveals, that this is only partially the case [1]. I believe that this lack of cooperativity is mainly due to a difference in mentality—a difference of scope on the concept of a *scientific problem.* Scientists usually have a pragmatic attitude towards scientific problems. Those problems are believed to be obvious and meta-investigations about their status are avoided. For many philosophers, on the other hand, only to say, what a genuine philosophical problem is, lies at the core of philosophy. Scientists often have the feeling, that philosophical investigations only hamper the scientific enterprise or even that they are merely confusions. So they often even try to immunize their work against philosophical considerations. Obvious exceptions of this case are scientific revolutions—as Thomas Kuhn would call it—or changes of paradigms, when scientific theories become infected with doubt and meta-reasoning seems unavoidable.

I think that the tendency of immunization against philosophical considerations is not entirely misplaced in general, but in our particular case it inhibits effective work of a certain group of researchers. In order to characterize this group further, let me first introduce some rough distinctions. Depending on education and academic background, there seem to be at least two different general motivations to study modal theories.

The *formal tradition* studies the algebraic, model- and proof-theoretic features of modal theories, for instance metalogical properties as completeness, decidability, the characterization of non-trivial models. Interdependencies between various calculi or between algebraic and model theoretic properties, the interplay between syntactic and semantic characteristics and the relation between modal logic and higher-order-logic are investigated. This tradition is connected with Tarski, Segerberg, Fine, Gabbay, Cresswell, van Benthem, Fitting and others. A sub-tradition of the formal one is what I call the *implementation tradition* connected with the names Fariñas del Cerro, Wallen, Ohlbach, Gabbay, Fitting and others.

The *metaphysical tradition* is interested in modal metaphysics or intensionality in general. This tradition stems from analytic philosophy and philosophy

[1]The fact, that there is almost no transfer of ideas from philosophy to the sciences seems to be one of the most severe problems of philosophy in general and philosophy of science in particular. Almost no working physicist, for instance, is aware of the foundational attempts for Quantum Mechanics done by philosophers of science since more than half a century.

of language and is connected with names like Carnap, Kripke, von Wright, Hintikka, Church, Montague, Kaplan, Castañeda and others. A sub-tradition is the *cognitive science tradition*, attempting to model modal concepts like 'belief', 'obligation', 'permission' and 'knowledge' for cognitive science or AI purposes, connected with the names Halpern, Fagin, Konolige, Levesque, Lehmann, de Rijke and others.

This perhaps provoking distinction is intended mainly to point out two things: First, computer scientists are split into two groups, subsumed under mathematical and philosophical positions. This is not meant to be a rigorous distinction but only to depict a tendency. Second, not all researchers investigating modal logics are naturally confronted with philosophical issues. But I argue, that the cognitive science tradition is. To make my point, let me examine the relation between metaphysics and cognitive science more closely.

According to functionalism, computer models lead to adequate representations of human intelligence or thinking in general and modal concepts in particular. Adequacy may in principle be achieved in three stages, namely on a *physical* or *physiological*, a *cognitive* or *conceptual* and an *empirical with respect to a limited environment* basis.

Physical adequacy is beyond the scope of cognitive science. It is claimed that an investigation of the hardware of thought is not necessary to understand thought and intelligence. Thought and intelligence are rather seen as abstract concepts independent of a physical realization. Empirical adequacy on the other hand is too limited, it is merely an *engineering approach* heuristically trying to build limited systems which *behave like* intelligent agents. This approach may be methodologically appropriate to make progress in science, but researchers on an AI mission should always be aware of its limitations. It is a tautology, that cognitive scientists require cognitive adequacy, but by giving upper and lower bounds, I tried to depict its meaning without defining it.

Functionalism now is the bridge between cognitive science and metaphysics. Cognitive adequate models of modal concepts are believed to have metaphysical impact, depicting what modal concepts are by how they work. But beyond functionalism as a philosophical theory, I think, that it is almost common sense among AI researchers to be scientists and not engineers, to study thought and intelligence as abstract entities and not merely to imitate or approximate the structure of the brain with different means. So in a way, I believe, that functionalism is nothing but an attempt to precise mainstream AI ideology.

Cognitive science therefore is confronted with philosophy not only for foundational, but also for methodological reasons—as for instance Quantum Mechanics deals with philosophical questions like causality and holism. But there is one big difference to Quantum Mechanics, where philosophical questions arise as a consequence of the interpretation of the formal apparatus. In the case of modal logics, it is the philosophical decisions that determine the formal systems. But the philosophical decisions in this context are rather transparent and do not imply such obscure concepts like 'causality', 'scientific explanation' or 'realism', as it is usually the case in philosophy of science. Here they arise in a twofold way: First, there are *meta-decisions*, for instance about how to construct a formal modal system, its syntax and if relevant its semantics. Second there are *meaning decisions*, for example of how to formalize the modal concepts with regard to natural language discourse and metaphysical convenience. The name 'meaning decision' pays tribute to the Carnapian view, that expressions formalizing some concept can be seen as *meaning postulates*. The last point may not seem very philosophical to some readers, maybe they are right, but nevertheless philosophers investigate these problems since almost a century and I think that it would be convenient to share their experience.

Now there remain two main problems. The first is to show, to what extent metaphysical questions have impact on formalizations, the second one to overcome the language barrier between philosophy and computer science. I use Frege's semantic theory to face both problems: to give an introduction to the philosophical theories of intensionality and to show, how philosophy of language and metaphysics are intertwined. After this metaphysical prelude, I discuss propositional alethic modal logic—the mother of (almost) all modern modal logics with a view at meta- and meaning decisions. Then in further sections, the specific philosophical problems of epistemic and deontic propositional logics are considered. In a last section, various philosophical approaches towards a quantificational intensional logic are discussed.

Readers of AI literature might have the impression, that there are canonical modal logics for knowledge and belief—namely **S4** and **S5**—and that the transition from propositional to quantificational logics is straight forward, but from a philosophical point of view, this impression is not adequate. I argue, that philosophy of language provides a powerful general apparatus to handle modalities, that goes beyond the precision of many of the AI approaches. On the other hand, concrete AI systems may be seen as a testbed for philosophical theories as for example rigid and non-rigid designators or the transworld identity problem, which perhaps in the spirit of Kripke might turn out to be no problem at all.

So this survey has the fourfold aim to give a brief formal introduction to alethic, epistemic and deontic logics, to discuss the philosophical problems connected with intensionality in general and with the formalization of modal concepts in particular, and to provide a bibliography with emphasis on philosophical literature. If the reader becomes aware of the many (nontrivial) problems imported by a philosophical perspective or at least is guided to find exciting philosophical literature, I consider my mission accomplished.

2 A Metaphysical Prelude

2.1 Historical Remarks and Basic Concepts

In the introduction, I used some philosophical concepts, that may not be part of a computer scientist's vocabulary. These concepts shall be introduced in this section together with a brief exposition of the history of modal metaphysics.

Philosophical theories of modalities are more than thousand years older than the explicit use of the word. Aristotle, in his *prior analytics*, not only distinguishes four ways of specifying a statement: as *possible*, *contingent*, *impossible* or *necessary* [2], he was probably also the first to develop *modal syllogisms*. Unless Aristotle, who used modalities only in this *de dicto* way specifying expressions, other Greek philosophers employed also *de re* modalities specifying physical objects [3]. Consider for example the sentences

(i) Gödel proved that first order predicate calculus is complete.

(ii) Gödel proved that first order predicate calculus is necessarily complete.

(iii) Necessarily, Gödel proved that first order predicate calculus is complete.

Sentence (ii) is obviously true. It ascribes a *de re* modality, because the necessity of the predicate calculus's being complete is a matter of fact. But sentence (iii), ascribing a *de dicto* modality by claiming, that sentence (i) is necessarily true, is wrong, because Gödel might possibly have shifted his interests from logics to psychosomatics before trying the proof or the proof might have been anticipated by somebody else.

The linguistic transition from the adjective '*modalis*' to the noun '*modus*' was achieved by scholastic philosophers around Aquinus, who also obtained an elaboration of modal theory. According to the new theory, *modi* may be physical or logical, specifying an object or an expression by a concept, for example as an essential or an accidental property of a substance or specifying it through a syllogism. Aquinus for instance conceived of three different entities a mode can specify, namely *a subject* (Philosophical ideas are *important* for modeling cognitive concepts.), *a predicate* (Many researchers from the AI community *almost completely* ignore the philosophical background of modal logics.) and *the relation between a subject and a predicate* (I *believe* that the AI community should take philosophical ideas seriously.).

Note that only the last expression is modal in the modern sense. The modal metaphysics of Aquinus and Duns Scotus are excellent examples for the highly elaborated—but today almost forgotten—technical apparatus of scholastic logicians. This is not the right place to go into details, but philosophy of language is really not an invention of the twentieth century [4]. Another concept, the *propositional attitudes* of contemporary philosophy of language, expressing the state of mind that a person has related to an expression, can be found in Occam and *possible worlds*, which made Leibniz famous were not only discussed in Descartes some years earlier, their invention can probably be ascribed already to Gilbert de Poitiers and later Duns Scotus [5]. Nevertheless, possible worlds had a bad reputation in philosophy for the next centuries, perhaps due to their rejection by Hume, until philosophers such as Wittgenstein and Carnap prepared their renaissance in analytic philosophy early in this century.

[2] In fact, the etymology of the English concepts can be traced back to Boetius's translation of corresponding Greek ones. Boetius also introduced the attribute '*modalis*' in connection with these specifications.

[3] The explicit terminology for the *de re-de dicto* dichotomy was reintroduced centuries later by Thomas Aquinus.

[4] For an introduction to medieval logics consider for instance [114].

[5] c.f. the discussion in [4], especially in footnote 30) of this text.

Despite the rejection of possible worlds, modal concepts still played an important rôle in metaphysics, they lie for instance at the heart of Kant's '*Kritik der reinen Vernunft*' [111]. The functions of thinking in a judgment can, according to Kant, be subsumed under four concepts: Quantity, quality, relation and modality (Krv A70) [6]. Modality is defined as that function of a judgment, which does not contribute to the judgment's content, but expresses—as in Aquinus's third clause—the relation of the subject and the predicate of a sentence. Kant—in contrast to Aristotelian tradition—distinguishes three kinds of judgments with respect to modality: possible (*problematische Urteile*), true or real (*assertorische Urteile*) and necessary (*apodiktische Urteile*) ones (KrV A75).

Another influential idea of Kant is that modalities do not contribute internal information about the object a statement is about, but express the epistemic relation between the speakers ratio (and his empirical application) and the statement (KrV A219)—today one would again speak of propositional attitudes. Definitions of modalities are meant to be explanations of the empirical use of the concepts 'possibility', 'necessity' or 'reality' (KrV A219), the postulate of possibility of things requires, for example, that its concept coincides with the formal conditions of an experience (KrV A220).

Today, following von Wright [205], one would give a whole catalogue of different modalities:

- (Truth or Falseness) (the majority of philosophers would opt for omission),
- alethic modalities (necessity, contingency, and possibility),
- epistemic modalities (knowledge, belief—sometimes considered a proper, a doxastic modality),
- temporal modalities (always, sometimes etc...),
- boulomaic modalities (desire, wishes etc...),
- deontic modalities (duty, permission, obligation, commitment, juristic norms etc...),
- evaluative modalities (good, bad, ethical norms...),
- physical or causal modalities (causal consequences etc...),
- modalities of action (somebody does, did, will do, starts doing, stops doing something etc...),
- mellontic modalities (modalities of process, of becoming).

These modalities can be further structured by defining alethic, epistemic, boulomaic, causal and temporal modalities as *ontic* modalities—modalities of being, by subsuming evaluative modalities under the class of *deontic* modalities and

[6] Traditionally, Kant's **K***ritik der* **r***einen* **V***ernunft* is not referred to by page-numbers, but by a special index. The meaning of 'KrV' is obvious, the letters 'A' or 'B' stand for editorial variants and the number gives the location in the text

modalities of action under the class of mellontic modalities. So much for the history of the concept of modality, let us now return to Occam's and Kant's idea of propositional attitudes. Those are today treated in the context of *intensionality*, a concept that I already have used without a definition. Theories of intensionality will be the content of the next section.

2.2 Frege Semantics and General Intensionality

For readers familiar with formal semantics, the semantical theory of Frege is not hard to understand modulo a translation of Frege's terminology to the modern standard one. Consider a formal language together with a domain of interpretation. There are now two relevant levels: a *symbolic* and a *physical* one. Symbols *denote* objects, or the latter are the *denotation* or *extension* of the former. But according to Frege, this is not enough for a formal semantics as a *theory of meaning*.

In his influential article '*Über Sinn und Bedeutung* [70]', Frege asks why the equations '$a = a$' and '$a = b$' differ in cognitive content (*Erkenntniswert*). He argues that this difference can neither be a matter of symbols, as notation is arbitrary, nor of extension, because in case '$a = b$' is true, then 'a' and 'b' must denote the same object (the extension of the symbol is the same) and then the difference disappears. Frege, in order to incorporate the cognitive content, therefore calls for an intermediate third level, where the arbitrariness of symbols vanishes, but their difference is conserved. In a more formal way in extensional semantics, there is a valuation function mapping symbols to objects of the domain. Now, if a is equal to b, the valuation function maps both symbols to the same object. Analogically, assume a function, that maps the symbols to the new level, such that the values of this new function may be different although the symbols may denote the same object. But how could we give an intuition to this third level? Frege tries in different ways and it was sometimes argued, that he mixes up what should best remain separated [112, 32]. I leave it to the reader to follow Frege's arguments [71, 66] in detail and sketch a modern paraphrase of his account.

The entities that are supposed to be situated at the new level and which I roughly brought in connection with cognitive content, are usually called *intensions*, *senses* or *meanings*. Before considering these concepts more closely, let us take a closer look at the ontological side of the problem, i.e., at the bearers of intensions etc. Here again the situation parallels formal semantics. Frege doesn't analyze sentences in terms of *subject* and *predicate*, as philosophical tradition had done since Aristotle, he employs applications or saturations (*Sättigung*) of a functional part—a concept (*Begriff*)—and an object (*Gegenstand*) [68, 70]. These ideas lead to the following scheme:

symbol	(proper) name	concept symbol	sentence
intension	intension of (proper) name	intension of concept symbol	proposition
extension	object	concept	truth value
impression	impression connected with (proper) name	impression connected with concept symbol	impression connected with sentence

Names and their rôle in language are an important subject in analytic philosophy. Frege and Russell gave ideas on how to treat them. As we just saw, names according to Frege have intensions as well as extensions. Russell employs a much smaller ontology than Frege. For him, only sentences have intensions, the other entities *referring directly*. But what both theories have in common is that names can be eliminated in favor of *definite descriptions*, for example the name 'Armstrong' by the definite description 'the man who first walked on the moon'. In fact, all names can be eliminated and the information is then carried by a predicate, identity and bound variables. But in general, one must be careful. The expression 'The Holy Roman Empire', for instance is a name but not a definite description, because it was neither holy, nor Roman nor an empire. Today, it seems that many philosophers, among them Kripke and Kaplan favor the direct reference approach. I will return to this subject in connection with objectual quantification and rigid designators.

There is perhaps one oddity in the table above: Sentences are nothing but names for truth values. This, on the one hand, is exactly the situation of propositional logic, on the other hand, it shows how little information is carried by extension, such that another level of language seems indeed necessary to make communication possible. Let me close this argument with a little deviation to Frege's general ontology. Frege is a Platonist, postulating three realms of existence: a *realm of physical objects* (*Dinge der Außenwelt*), a *realm of subjective impressions* (*Vorstellungen*), and a *realm of non-physical objects*, existing independent of human thought (e.g. intensions like propositions (*Gedanken*), relations, or properties, truth values etc.)

The objects of the first realm are physical in the sense that they can be 'measured' with a physical device or are directly accessible to human sense organs. The objects of the second realm are the purely subjective colorings and illuminations (*Färbungen und Beleuchtungen*), that every cognitive agent connects with the objects of thought. The objects of the third realm are non-physical in the aforementioned sense, but they are assumed to exist without any relation to a human being. Mathematical theorems, for instance, are assumed to exist independently of humans proving them. The third realm is also populated by intensions, even by intensions of extensionless sentences, extensionless proper-names like 'Pegasus', and of extensionless concepts like 'moves faster than the speed of light'. In Frege's ontology, even all higher order concepts and their intensions explicitly exist, for instance, not only the existence of functions, but also of functions of functions etc., i.e., of higher order functions [68] is assumed. Those readers familiar with λ-calculus will note a striking similarity of ideas, and it is no wonder that Alonzo Church who developed this calculus later became 'the great modern champion of Frege's semantical theories', as David Kaplan [112] calls him.

From the argument presented above, it follows that Frege's intension shall be independent of any cognizers and psychological influences, which explains the separate last line in the table above—the private aspects coming with any cognition are accidental and do not contribute. Now Frege gives two definition of intension. First, he defines the intensions of (proper) name symbols, concept symbols and sentences (I should better say 'declarative sentences', of course) as *the way they are given* [71] or specified—so he defines them with reference to modalities in the medieval meaning of the word—but he still gives a different one for sentences, namely as *that which—if anything—may come into question for truth* [66]. Here the twofold meaning that the term 'intension' has in Frege's philosophy becomes evident. Note that the asymmetry between the two definitions is only superficial: Name and concept symbols have only an indirect connection with truth. They have no truth values, but determine the truth value of any sentence in which they occur. In more modern terms, intensions in the Fregean sense can be ascribed the following features [32]: they are *objects of thought*, they have *cognitive significance*, allowing for an individuation of propositional attitudes and thoughts, they *bear the meaning of declarative sentences*, and they *determine the truth-values*.

It is the last feature that first proved accessible to formalizations: Carnap, referring to ideas of Wittgenstein, took the statement 'to come into question for truth' literally and defined a proposition as the set of those possible state descriptions in which the corresponding sentence is interpreted as true. The idea behind this is that in order to understand the meaning of a sentence, one must be able to predict its 'behavior' in changing situations, to know under what circumstances it will turn out true, false or without denotation. Today, most philosophers formalize propositions as characteristic functions of sets of possible worlds [75, 113, 81, 9], 'that is, as functions, which assign to each possible world one of the two truth values' [113], but this is clearly in the same spirit [7]. Another modification is that today the standard terminology has shifted from 'state description' to 'possible world'. Maybe this terminology is a bit mistaken, because the idea behind possible worlds is really not that of some mysterious universes in hyperspace, but coincides with the commonplace supposition, that

[7] Readers familiar with λ-calculus might perhaps find some pleasure in verifying the analogy, that intensions are in principle λ-abstractions over possible worlds.

things could be different. Unfortunately, the concept of possible worlds, if taken literally in the original Leibnizian sense, opens the door to metaphysical speculation and a touch of science fiction. Actually, there is no conceptual difference to the *sample space* of probabilistic calculus and many philosophical pseudo-problems would disappear, if one had chosen a more neutral terminology right from the beginning [8]. For more elaborate theories of intensionality in this spirit, consider for instance the writings of Kaplan [112], Gallin [75] and Anderson [9].

All approaches to a general logic of intensions characterized so far share one important feature: Intensions are reducible to possible worlds and truth values. It is doubtful, whether this is still a reconstruction of Frege's ideas and therefore only natural, that logicians like Church [41, 40, 42, 43], Bealer [19] and Anderson [9] proposed taking intensions as irreducible entities, thus dispensing with possible worlds. For a detailed discussion see [9]; Church explicitly claims that his systems are reformulations of Fregean semantics.

The above alternatives are still heatedly debated. Anderson gives a series of arguments in favor of the latter approach, arguing that a possible worlds approach is appropriate only for alethic modalities [9]. On the opposite Kaplan [112] believes that for his (intensional) logic of demonstratives, a possible world approach is more convenient. There are even authors who see possible worlds as an historic error and a possible worlds metaphysics simply as counterintuitive [119], but I do not want to go into that fundamental discussion.

Let us now turn to a second Fregean perspective in intensionality [71]: the perspective of *indirect speech* (indirekte Rede)—in modern standard terminology *intensional* or *referentially opaque contexts*—closely related to the problem of formalizing propositional attitudes. To give an example, consider the sentence

> Laurence Sterne was ill-reputed for his frivolity.

This context tolerates *salva veritate* substitution of 'Laurence Sterne' by any expression with the same extension, for example by 'The author of 'Tristram Shandy''. Contexts, in which such substitutions of equiextensional expressions *salva veritate* are possible, are called *extensional contexts.* Now on the other hand the sentence

> It is necessary, that the author of 'Tristram Shandy' is the author of 'Tristram Shandy'.

is true, but its substitutional variant

> It is necessary, that Laurence Sterne is the author of 'Tristram Shandy'.

[8] We will remember this fact in the context of transworld-identity.

[9] These arguments depend on the intensional/hyperintensional dichotomy developed in the sequel and are similar to arguments against formalizing epistemic logics analogously to alethic modal logics.

is false. Which is the class of substitutional invariants for this kind of sentence? It is often stated, that the distinction between necessary and contingent statements rests on the distinction between *analytic* and *synthetic* truths, the latter being true by virtue of facts, the former dividing into the class of truths of logic and into the class of those non-logical truths, which can be decided alone with the help of a dictionary. As often in philosophy, there is much controversy about the analytic/synthetic dichotomy and there are even philosophers like Quine, who are sceptical of this dichotomy in general. But there is one important point for us. The class of logical truths (or in general the class of analytic truths) is exactly the class of sentences which is invariant *salva veritate* under substitution in contexts like the above, which are called intensional contexts. (This can be taken as a definition of an intensional context.) Note that this definition coincides with Carnap's definition of intensions—logical truths have the same truth value in all possible circumstances.

Beside intensional and extensional contexts, there is a third kind of context which is sometimes called *hyper-intensional*, where not even substitutions of equiintensional expressions are tolerated. Examples for hyper-intensional contexts are propositional attitudes like statements about beliefs, knowledge and convictions. For example, a beginner with logic may believe the logical truth of the formula

$$(\phi \vee \neg\phi)$$

without believing the logical truth of the equivalent statement

$$(\phi \to \psi) \vee (\psi \to \phi)$$

which is one expression of the paradox of material implication.

In Fregean semantics, reason for the non-extensionality of some sentences is easily explained: The denotation of an intensional sentence is a proposition, i.e., an intension and that of a hyper-intensional sentence even is a relation between an agent and a proposition. A deeper analysis of these semantical features would lead right to theories of 'that-clauses' or indexicals of Kaplan or Castañeda, but we stop here and draw the following conclusions:

- Alethic and deontic contexts are intensional,
- epistemic contexts are hyper-intensional.

These facts play a crucial rôle for the construction of epistemic logics.

3 Propositional Modal Logics

3.1 General Remarks

In the sequel, I want to keep the general remarks about intensionality in mind and specialize to the less ambitious alethic, deontic and epistemic modal calculi in order to discuss their applicability to AI purposes. Following Bull and Segerberg [30], Kreiser et al. [120], Chellas [35], Lenzen [132], Meyer et al. [145, 146] and Åqvist [10], we introduce the propositional variants before proceeding to quantificational theories. Deontic and epistemic modal logics grew up in the shadows of alethic modal logics [10], so I treat the alethic logics first, too. For the following paragraphs, I assume the reader's familiarity with non-modal first order logic and the basics of modal logics, their syntax and semantics. Therefore, I do not attempt rigor in formal details but concentrate on an informal discussion of motivations and philosophical background of the formal concepts.

According to van Benthem [21], three pillars of wisdom support the edifice of modal logic—*completeness theory*, *correspondence theory* and *duality theory*. Completeness theory studies the interplay between syntax and semantics of modal systems, correspondence theory the interdependence between axioms and frame constraints on accessibility, and duality theory the relation between model theoretic and algebraic properties of modal logic. This is the typical point of view of a mathematician; philosophers would add (at least) two further pillars—the pillar of application and the pillar of meta-reflection. Applications in this context do not include implementation issues, because these problems seem technical rather than structural, they are rather to be understood in connection with questions like internal adequacy for modeling modal concepts or propositional attitudes, what I called 'meaning decisions' in the introduction. In opposition to that, meta-reflections consider external adequacy for example by investigating, what general features a theory formalizing intensional concepts should possess.

In this spirit, it is only natural to start with two meta-reflections: First, should modalities be incorporated as *(meta-linguistical) predicates* or else as *sentence-forming operators*? Second, what kind of approach towards modal logic does one prefer—the *syntactic*, the *model theoretic* or the *algebraic approach*?

The first decision is that between an approach which leads to a hierarchy of metalanguages and another one, which remains at one level. A metalinguistic predicate acts on metatheoretical names of sentences, as in Tarski's theory of truth and—as for example in some early systems of deontic logic, for instance [204] or Gödel's provability interpretation—iterations of modalities are not necessary or even meaningless. Another significant feature of the predicate approach is that applications of modal predicates are not restricted to names of sentences, one is free to insert names of entities like proofs or act-types [204].

[10] Bull and Segerberg, in the historical part of their paper [30], do not only give references to the historically interested reader, they also mention the standard textbooks on alethic logics, some of the more recent survey papers, collection of important papers and bibliographies. Besides this work, I suggest reading Hughes and Cresswell [102, 103], Chellas [35] and the survey paper by Fitting [64]. For a discussion of the older epistemic systems and historical items, I recommend Lenzen [131], bibliographies may be found in Lenzen [131, 132]. Good surveys for the more recent development and especially the applications to computer science are the papers of Halpern and Moses [89, 90] and Meyer et al. [145, 146]. As introductions to deontic logics, consider Åqvist [10] and van Eck [54]. At the beginning of his references, Åqvist gives indications of bibliographies (the one in [46] approaching 1500 items!), reference to standard publications or selection of important articles as [94, 95] can be found in Åqvist and van Eck. A reader interested in the history of deontic logics should take a look at von Wright's survey [211].

The sentence-forming operator approach may be called the standard approach. I will concentrate on the latter in this text.

The second decision between a syntactic, a model-theoretic and an algebraic approach is more controversial. There are several sub-traditions, which can be subsumed under the syntactic approach:

The *intuitive semantics approach* (Lewis, Becker, von Wright, Mally, Grzegorczyk) tries to formalize real-world semantical intuitions of modal concepts through syntactic rules and axioms, like Euclidean geometry yielded a syntactic axiomatization of the properties of classical space. Especially in the case of epistemic and deontic logics, this approach became infected with paradoxes—how can one guarantee that the theory is sound with respect to intuitions? One major problem seems to be that material implication does not really conform to our natural language intuitions. But, however, intuitive semantical considerations must precede any inquiry in modal logics, if they are meant to be more than formal games.

The *relevance and entailment approach* (c.f. Anderson and Belnap [8]) constitutes an alternative to classical logic, trying to cope with certain paradoxes (e.g. the paradox of material implication) and inconveniences of classical logics.

The *proof theoretical approach* (Zeman [212], Prawitz [165], Bull and Segerberg [30], Boolos, Wallen [202], Fitting [62], Ohlbach [155]) may be subdivided into two branches. One branch being the *provability interpretation* of modal logics, the other one concentrating on *proof methods* like sequent calculi, natural deduction or resolution.

The *model-theoretic approach* is discussed at large in the introductory section of Bull and Segerberg [30], including a brief discussion and some references concerning the birth of possible world semantics, so I don't want to go into details. An influential predecessor of this tradition was Carnap [31] with his theory of state descriptions influenced by Wittgenstein's concept of logical space. The contemporary model theoretic tradition can be subsumed under the slogan '*relational semantics*', beginning with the work of Kanger [109, 110], Hintikka [96] and Kripke [121, 122]. Recent years witnessed the emergence of some non-standard possible world semantics [23]. Within the standard theory, many refined methods (for an overview, consider [30, 21]) have been developed, *correspondence theory*—studying the relation between the formal properties of the accessibility relation and the axioms of the theory—was developed [21, 156] and definability features in connection with higher-order logic [21, 22] were investigated.

The outstanding figure of the algebraic tradition is Alfred Tarski. In his joint paper with Jónsson [106], Stone's representation theorem [171] relating boolean algebras and algebras of sets is generalized to boolean algebras with functions, implying that the algebraic theory for propositional logic could be lifted to propositional modal logics. Bull and Segerberg [30] discuss, how natural features of the modal algebras reflect those of the corresponding modal logics and first of all their model theoretic properties. One may draw the conclusion, that Tarski and his collaborators anticipated the early model theoretic results; it is indeed not difficult to use ideas from the proof of Stone's theorem to obtain the crucial model theoretic structures—frames and models—from the algebraic ones without postulating possible worlds (one rather considers the set of ultrafilters constructed on the power set of the carrier of the algebra). These algebraic results apply even to such modal logics, where no standard Kripke-models exist (definitions follow). The algebra then induces *general models* which correspond in a way to the models of higher order logics [21]. For more information on these algebraic features, which I will not consider further in this text, I recommend any standard work on algebraic logic or the writings of Bull and Segerberg [30], van Benthem [21], Goldblatt [84] and Fine (c.f. the bibliography of [30].). Because of these fundamental results, many logicians favor an algebraic approach as fundamental. The fact that the algebraic results—of course without the explicit transfer to Kripke-frames—were published six years before the first model theoretic results without being recognized by the philosophic community, must be considered a historic peculiarity.

Resuming these ideas, possible worlds today do not have the predominant status that they had in the sixties and it is questionable whether they really are the most natural approach to modalities. I have already pointed out that there are general intensional logics without possible worlds and now we have seen, that also modal logics can do without it. So possible worlds in the literal sense of the word are only *possible* approaches to our purposes, but not *necessary* ones.

3.2 Alethic Modal Logics

This section introduces to some of the most fundamental concepts of alethic modal logic with intended applications to epistemic and deontic systems, starting with the basic syntactic notions. The *language* L_M *of modal logic* is a language L of non-modal predicate logic supplied with the additional symbol '$\Box$': $L_M = L \cup \{\Box\}$. Finite strings of L_M-symbols are called *expressions*, the set of L_M*-formulas* is the smallest set of expressions containing all L-formulas, closed under the formation rule $\phi \in L_M \Rightarrow \Box\phi \in L_M$.

A *propositional modal logic*, is any non-modal propositional logic, for example the smallest subset of L_M containing *all tautologies* and closed under *modus ponens* and the *substitution rule*. One could use any non-modal propositional calculus [11] over L_M in order to obtain this logic. Then, all syntactic notions like 'derivation' or 'proof' can be transfered from non-modal propositional logic.

As modal logics are intended to be theories of specific modal concepts and therefore conceived to be richer than non-modal logics not only in vocabulary, we must extend the calculus by new axioms and rules, thus providing *meaning postulates* for the modal concepts to be explicated. Following Bull and Segerberg, we call a modal logic *classical*, if it contains the axioms

$$\begin{array}{ll} \mathbf{K} & \Box(\phi \to \psi) \to (\Box\phi \to \Box\psi), \\ \mathbf{t} & \Box\top, \end{array}$$

[11] Here and later we follow Fitting [64] by using the notion of a *logic* for a set of sentences closed under certain conditions, a *calculus* is a specific set of axioms and rules leading to a logic.

but a logic with only these additional axioms is still too weak for most applications. A further step is to provide the calculus with new rules

$$
\begin{array}{ll}
\textbf{nec} & \dfrac{\phi}{\Box\phi}, \\
\textbf{mon} & \dfrac{\phi \rightarrow \psi}{\Box\phi \rightarrow \Box\psi}, \\
\textbf{cgr} & \dfrac{\phi \leftrightarrow \psi}{\Box\phi \leftrightarrow \Box\psi}.
\end{array}
$$

A modal logic including **K** is called *congruential* when closed under the congruence rule **cgr**, *regular* when closed under the monotonicity rule **mon**, and *normal* when closed under the necessitation rule **nec**. Normality implies Regularity implies Congruence.

Normal logics, because of their semantic and syntactic niceties, are the most familiar systems of modal logic. For reasons that will become clear later, consider now the relation between **nec** and **mon**. One difference is, that the formula '$\Box\top$' can be derived in normal, but not in regular logics. Another fact is that addition of '$\Box\top$' to a regular logic allows for a deduction of **nec**. So the axiom '$\Box\top$' is the crucial difference between a normal and the corresponding regular logic.

Even at this stage, the logics constructed so far are too weak for applications. Let us follow Bull and Segerberg in giving a catalogue of modal systems using the *Lemmon code*, which concatenates named formulas. But first, for the sake of brevity, let me define the operator dual to '$\Box$':

$$\Diamond\phi \leftrightarrow \neg\Box\neg\phi.$$

The following list contains some of the most popular formulas leading to axiom systems with real-world applications:

$$
\begin{array}{ll}
\textbf{D} & \Box\phi \rightarrow \Diamond\phi, \\
\textbf{T} & \Box\phi \rightarrow \phi, \\
\textbf{4} & \Box\phi \rightarrow \Box\Box\phi, \\
\textbf{E} & \Diamond\phi \rightarrow \Box\Diamond\phi, \\
\textbf{B} & \phi \rightarrow \Box\Diamond\phi, \\
\textbf{Tr} & \Box\phi \leftrightarrow \phi, \\
\textbf{V} & \Box\phi, \\
\textbf{M} & \Box\Diamond\phi \rightarrow \Diamond\Box\phi, \\
\textbf{G} & \Diamond\Box\phi \rightarrow \Box\Diamond\phi, \\
\textbf{4.3} & \Box(\Box\phi \rightarrow \Box\psi) \rightarrow \Box(\Box\psi \rightarrow \Box\phi), \\
\textbf{4.4} & \phi \rightarrow (\Diamond\Box\phi \rightarrow \Box\phi), \\
\textbf{4.2.1} & \Box(\Box(\phi \rightarrow \Box\phi) \rightarrow \phi) \rightarrow (\Diamond\Box\phi \rightarrow \phi), \\
\textbf{4.3.2} & \Box(\Box\phi \rightarrow \psi) \vee (\Diamond\Box\phi \rightarrow \psi), \\
\textbf{H} & (\Diamond\phi \wedge \Diamond\psi) \rightarrow (\Diamond(\phi \wedge \psi) \vee \Diamond(\phi \wedge \Diamond\psi) \vee \Diamond(\psi \wedge \Diamond\phi)), \\
\textbf{Grz} & \Box(\Box(\phi \rightarrow \Box\phi) \rightarrow \phi) \rightarrow \phi, \\
\textbf{W} & \Box(\Box\phi \rightarrow \phi) \rightarrow \Box\phi.
\end{array}
$$

The name **D** stands for 'deontic', **T** is a name invented by Feys, **4** signifies the characteristic axiom of Lewis' **S4** system, **E** stands for 'Euclidean', **B** for

'Brouwer', **Tr** for 'trivial', **V** for 'verum', **M** for 'MacKinsey', **G** for 'Geach', **H** for 'Hintikka', **Grz** for 'Grzegorczyk', **Dum** for 'Dummett' and **W** for '(anti)-well-ordered'. The name **K**, defined some lines above, honors Kripke. We now list the names of the systems to the left and the corresponding Lemmon code to the right. The systems given below are defined to be the smallest normal systems containing the formulas on the right side of the equation as axioms.

K	:=	**K**,
T	:=	**KT**,
S4	:=	**KT4**,
S5	:=	**KT4B** = **KT4E**,
deontic **T**	:=	**KD**,
deontic **S4**	:=	**KD4**,
deontic **S5**	:=	**KD4E**,
Brouwer system	:=	**KTB**,
S4.1	:=	**KT4M**,
S4.2	:=	**KT4G**,
S4.3	:=	**KT44.3.2**,
S4.4	:=	**KT44.4**
S4.2.1	:=	**KT44.2.1**
S4.3.	:=	**KT44.3.2**
Prior's Diodorean logic	:=	**KT4Dum** = **D**,
Grzegorczyk's system	:=	**KT4Grz** = **KGrz**,
Löb's system	:=	**K4W** = **KW**,
the trivial system	:=	**KTr** = **KT4MB**,
the verum system	:=	**KV**

The Geach formula $\Diamond\Box\phi \rightarrow \Box\Diamond\phi$ can easily be expanded to the generalized Geach formula

$$\mathbf{G}_{(m,n,p,q)}) \qquad \Diamond^m\Box^n\phi \rightarrow \Box^p\Diamond^q\phi.$$

Then, many of the formulas given above can be seen as special cases:

$$\begin{array}{rlrl} \mathbf{D} := & \mathbf{G}_{(0,1,0,1)}, & \mathbf{T} := & \mathbf{G}_{(0,1,0,0)}, \\ \mathbf{4} := & \mathbf{G}_{(0,1,2,0)}, & \mathbf{E} := & \mathbf{G}_{(1,0,1,1)}, \\ \mathbf{B} := & \mathbf{G}_{(0,0,1,1)}, & \mathbf{Tr} := & \mathbf{G}_{(0,1,0,0)} \wedge \mathbf{G}_{(0,0,1,0)}, \\ \mathbf{G} := & \mathbf{G}_{(1,1,1,1)}. & & \end{array}$$

The normal logics form a lattice. Pictures are given in [83, 35, 102].

Besides these *Hilbert calculi*, there are also *natural deduction*, *sequent*, *tableaux* and *resolution systems*. For natural deduction, the classical reference is Prawitz [165], consider also Fitch [61] and Siemens Jr. [188]. For tableaux, consider Fitting [62], for a matrix calculus Wallen [202] and Ohlbach [155] for resolution. General introductions are of course Bull and Segerberg [30] and Fitting [64], where also some of the problems connected with these methods are discussed [12]. In general, for the most popular systems, there are non-axiomatic calculi, but there are other systems, for which only axiomatic calculi exist.

[12] For natural deduction, for instance, there is a symmetry break between the elimination and the introduction rules.

Without going further into the details, let us now turn to the *semantics* of propositional modal logics. I want to concentrate on model theoretic semantics, the references for the other kinds of semantics are given above.

A (Kripke) *frame* is a pair $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$ of a nonempty set $\mathcal{W}$—the *set of possible worlds* and a binary relation $\mathcal{R}$—the *accessibility relation*—on $\mathcal{W}$. Strictly spoken, frames are just directed graphs. A (Kripke) model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, v \rangle$ is a triple of a frame together with a *valuation function* v, which assigns a truth value to every propositional letter in every world. So $v(\phi, w) = \top$ is the formal expression of the fact that the sentence ϕ holds in the possible world w. Note the difference to classical logics, where models are true interpretations. Here, a model is merely an assignment of meaning. But of course, the concept of a possible worlds model is used to define truth or semantical implication; we have to distinguish between truth in a possible world, in all possible worlds and in all models (of a certain class). The notion of truth in a possible world is—paralleling the construction of modal formulas—simply the non-modal truth definition relativized to possible worlds plus an additional rule for the modal operator. With the definitions given above

$$\mathcal{M}, w_0 \models \Box\phi \quad \Leftrightarrow \quad \text{f.a. } w_1 \in \mathcal{W}(\mathcal{R}w_0w_1 \Rightarrow \mathcal{M}w_1 \models \phi).$$

By duality, the rule for the second modal operator is

$$\mathcal{M}, w_0 \models \Diamond\phi \quad \Leftrightarrow \quad \text{ex.} w_1 \in \mathcal{W}(\mathcal{R}w_0w_1 \text{and} \mathcal{M}w_1 \models \phi).$$

In words: *Under the alethic interpretation, a formula is necessarily true in a possible world if it is true in all possible worlds accessible from the first one and possibly true in a possible world if it is true in at least one possible world accessible from the first world.* Furthermore a formula is called *valid in a model* $\mathcal{M}$ if it is true in all possible worlds of this model. Note, that thereby all logical truths are necessary truths, as desired. Let **C** be a non-empty collection of models. Then a formula is called **C**-*valid*, if it is valid for all models in this collection, it is **F**-*valid* in a collection of frames **F**, if it is valid in the collection of all models based on this frame. Note that frames are already sufficient to structure validity, because the quantification over valuation functions is merely a combinatorics of truth values. The concept of **C**- or **F**-validity plays exactly the rôle that unrestricted validity plays in non-modal logic, the restriction being due to the fact, that one is not free in the choice of the accessibility relation. Given these definitions, it is straightforward to define the concept of semantic consequence and related properties. But the deduction theorem for example, a standard theorem of first order logic, obtains only for a restricted class of frames, namely that closed under generated subframes and ultraproducts [64, 30]. This comes down to the question of first-order definability [21]. For the above set of frames, completeness proofs are also straightforward. The most popular systems—among them **D**, **T**, **B**, **S4** and **S5**—have the abovementioned properties and thus all the nice properties of non-modal propositional logics, as completeness, compactness, Löwenheim-Skolem and decidability, as can be verified by considering any standard text. For completeness results and decidability proofs, see for example Chellas [35]. He gives them for the normal systems (in Lemmon code) **K**, **K4**, **KE**, **K4E**, **KB**, **KB4**, **KD**, **KD4**, **KDE**, **KD4E**, **KDB**, **KT**, **KT4**, **KTB4** and **KTB**.

The frame-constraints on the accessibility relations for the several axioms are also not difficult to obtain. They can be calculated from algebraic semantics [30], guessed from the possible worlds semantics [102, 35, 21, 64] or calculated by an algorithm based on resolution [156]. I will only give the result generalized Geach formula, sloppily using object-language quantifiers for metalanguage expressions:

$$\mathbf{G}_{(m,n,p,q)} \qquad \forall w_0, w_1, w_2 \exists w_3 (w_0 R^m w_1 \wedge w_0 R^p w_2 \rightarrow w_1 R^n w_3 \wedge w_2 R^q w_3)$$

This frame-constraint obviously describes a *diamond property*, remembering the Church-Rosser-Property of λ-calculus or the confluence property of rewrite systems. It enables easy deduction of the constraints for the most popular formulas:

T:	reflexivity,	**D**:	seriality or idealization,
B:	symmetry,	**E**:	euclidity,
4:	transitivity,	**Grz**:	inductivity,
W:	finite irreflexive trees,	**K**:	no restrictions,
G:	diamond property or directedness.		

Again, from first order definability, there are formulas—as for example the McKinsey formula—for which no first order frame constraint exists [21]. But instead of going further into details, let us rather discuss the specific *meaning postulates*, introduced by the various axioms thus simulating the procedure of informal semantics. We consider this as the natural way of obtaining modal logics. As van Benthem writes: *... limits for the useful application of a correspondence perspective... are to be found in philosophical relevance, rather than technical impossibility* [21].

- The necessitation rule states that every proof from propositional logic of a formula can be extended to a proof of its necessity. Therefore all tautologies are necessary. **nec** is really built into the possible worlds semantics, it is even a motivation for the latter, as, from a metaphysical point of view, at least the analytical sentences of logic should be necessary and, if one wants to formalize these logical necessities only, the necessitation rule gives a full account of this fact.

- The distribution axiom **K** expresses the closure of modalities under consequences: for instance necessary truths entail only necessary truths. This is—in a way—the strengthening of non-modal material implication to modal contexts, what seems quite natural, because necessity is nothing but quantification over possible worlds. The example, to conclude the sentence 'It is necessary that electrons are fermions' from the premises 'It is necessary that if electrons have non-integer spin, then they are fermions' and 'It is necessary that electrons have non-integer spin' seems to be in perfect coincidence with our intuitions.

- In alethic contexts **T** is reasonable, too. If something is necessary, then it must be true in a deliberate world.

- **4** says that something is necessarily necessary, if it is necessary. This also seems plausible, because, for example, a necessary fact holds not only possibly necessarily. But—contrary to the examples given so far— one is not forced to accept this formula. One may accept it as formalizing one's intuitions or choose another axiom, as well.

- **B** seems plausible, too, because something true, is true—by a law of metalogic—in at least one possible world. But as this is a metalogical *definition*, one may express it as a logical analytical sentence of the object language as well.

Similar arguments are easily given for most of the other formulas—including **D**, **E**, **G** and the axioms leading to the logics between **S4** and **S5**.

A look at the frame constraints reveals further information about the corresponding modal concept. *Reflexivity* for example requires only very limited exchange of information from one world to another, on the other hand, this allows for much freedom for the individual worlds, i.e., for *situationalism* where the modal concept may change wildly from one world to another. *Symmetry* allows for a return from any vertex in the graph. The modal behavior in one world now influences that in its neighbors without determining it. Information is only available from neighbors. *Transitivity* establishes unidirectional information flow on paths. Two of these constraints still allow for situationalism, it is only all three constraints together, that give a partition of the set of possible worlds into equivalence classes. This situation, where one has independent clusters of equivalent possible worlds is—modulo equivalence—the situation that Leibniz had in mind, the situation of 'truth in all possible worlds'. This finally motivates the accessibility relation. It allows for a weakening of the metaphysical concepts and for situationalism (what might be useful in connection with non-logical analytic expressions—there might be a world, where electrons have non-integer spin and are bosons (*per definitionem*).

Up to now I—like most of the textbooks—have mentioned only positive results, now here are some negative ones:

- There are axioms without first order characterizable frame constraints,

- the McKinsey axiom violates the Löwenheim-Skolem property,

- irreflexivity is not modally definable,

- the relation between correspondence and completeness is only partially understood [21],

- the relation between modal and higher order logic contains many open questions,

- not all modal logics are complete [57], not even all first order axiom sets [21].

This may suffice to show, why modal propositional logic is not at all trivial. On the other hand it may give one of the reasons, why some modal systems are more popular than others.

There is even one more item, which makes the popular systems more popular (some modal logics have all the nice properties). For the weak modal systems, there exist countably many distinct modal prefixes for non-modal formulas, because the axioms only allow for accumulation and not for reduction of modal operators. In the stronger systems, those reduction rules are available and therefore only finitely many distinct modal prefixes exist. For a broad discussion of these facts, consider any of the standard textbooks [102, 35, 120], here are only some results [35]:

- **T**, deontic **T**, the Brouwer system, the deontic Brouwer system and deontic **S4** possess a countably infinite number of distinct modal prefixes.
- **S4** possesses exactly fourteen distinct modal prefixes, namely the empty prefix, $\Box$, $\Diamond$, $\Box\Diamond$, $\Diamond\Box$, $\Box\Diamond\Box$, $\Diamond\Box\Diamond$ and their negations. The reduction laws are
$$\begin{aligned} \Box\phi &\leftrightarrow \Box\Box\phi, & \Diamond\phi &\leftrightarrow \Diamond\Diamond\phi, \\ \Diamond\Box\phi &\leftrightarrow \Diamond\Box\Diamond\Box\phi, & \Box\Diamond\Box\phi &\leftrightarrow \Box\Diamond\Box\Diamond\phi. \end{aligned}$$
- **S5** possesses exactly six distinct modal prefixes, namely the empty prefix, $\Box$, $\Diamond$ and their negations. The reduction laws are
$$\begin{aligned} \Box\phi &\leftrightarrow \Box\Box\phi, & \Box\phi &\leftrightarrow \Diamond\Box\phi, \\ \Diamond\phi &\leftrightarrow \Diamond\Diamond\phi, & \Diamond\phi &\leftrightarrow \Box\Diamond\phi. \end{aligned}$$

For some applications, normal logics seem to strong. One way of weakening is to consider *regular* or even *congruent* logics. But the shift from **nec** to **mon** has some impact on semantics. The Kripke models considered so far are too strong for regular systems—in regular systems, the formula $\Box\top$ is Kripke-valid, but no longer derivable. To weaken the semantics, one assumes *queer worlds* in which logical truths come out false. But then, from **mon**, in particular from

$$\frac{\phi \rightarrow \top}{\Box\phi \rightarrow \Box\top},$$

it follows, that $\Box\phi$ must be false for deliberate ϕ in any queer world, too. So, nothing is necessary in queer worlds, but by duality, everything possible. These ideas given, the construction of frames is obvious [62, 64, 120]:

An *augmented frame* is a triple $\langle\mathcal{W},\mathcal{Q},\mathcal{R}\rangle$, where $\langle\mathcal{W},\mathcal{R}\rangle$ is a frame and $\mathcal{Q}\subseteq\mathcal{W}$ is the (possibly empty) set of *queer worlds*. The set $\mathcal{N}=\mathcal{W}-\mathcal{Q}$ is called the set of *normal worlds*. The models based on augmented frames are constructed as usual, except the definition of the necessity operator:

$$\mathcal{M}, w_0 \models \Box\phi \quad\Leftrightarrow\quad w_0 \in \mathcal{N} \text{ and f.a. } w_1 \in \mathcal{W}(\mathcal{R}w_0w_1 \Rightarrow \mathcal{M}w_1 \models \phi),$$

the rule for the dual operator is obvious.

For regular systems, all the meta-theorems of normal logics still hold, but correspondence theory is more complicated, because now frame constraints can relate normal worlds, normal worlds with queer worlds or even queer worlds such that one has to fix the constraints in an appropriate way, for example by using the formula $\Box\top$ as a selector between normal and queer worlds. The restriction of transitivity to normal worlds, for instance, leads to the characteristic formula $\Box\phi \to \Box(\Box\top \to \Box\phi)$ instead of $\Box\phi \to \Box\Box\phi$. The modification of the formulas is obvious, but expressions may become a bit lengthy during calculation.

Augmented models are extensions of the models for normal calculi—just set $\mathcal{N} = \emptyset$ to obtain the latter. On the other hand, if a logic has theorems that are necessary, the set of queer worlds must be empty. So again (and as required), $\Box\top$ acts as a mediator between regular and normal logics, but now on the semantic side. Instead of assuming a formula *globally* as in the case above, one can assume it *locally*: Let $\mathbf{F}$ be a collection of frames, $\{G\}$ and $\{L\}$ sets of sentences and ϕ a formula. Then we write $\{G\} \models_{\mathbf{F}} \{L\} \to \phi$ iff for every model $\mathcal{M}$ based on every frame $\mathcal{F} \in \mathbf{F}$ where $\{G\}$ is valid and every possible world w from $\mathcal{F}$, in which $\{L\}$ is true, $\mathcal{M}, w \models \phi$. $\{G\}$ is called the set of *global assumptions*, $\{L\}$ the set of *local assumptions*. The difference is, that global assumptions are *per definitionem* necessary, while local assumptions may hold locally in some worlds, but not in others. A local introduction of $\Box\top$ amounts to the introduction of this formula to normal worlds. For instance, the logic, where $\Box\top$ is locally assumed in the regular counterpart of **T** is called **S2**, the assumption to the regular counterpart of **S4** is called **S3**. These systems are not closed under regularity, but under the somewhat stronger *Becker Rule*:

$$\frac{\Box(\phi \to \psi)}{\Box(\Box\phi \to \Box\psi)}$$

These systems might be appropriate candidates for epistemic logics [64], but I am not sure whether there is a deeper analysis of these ideas in the literature. We return to this question in the context of logical omniscience and rational belief.

Another type of frame are the so-called *general frames* [30, 21, 64]. These frames have an algebraic and an intensional perspective. The algebraic perspective can be taken from Bull and Segerberg [30] and van Benthem [21] or from the more detailed articles of Goldblatt [84] and the numerous publications of Fine (c.f. the bibliography of [30].). As already mentioned, frames arise in the context of duality theory in a natural way.

The idea behind duality theory is to study translations between model theoretic and algebraic modal structures. The embedding of a frame in a modal algebra is easy. Roughly, one takes the power set of the set of possible worlds to build a set algebra, with some special provisions on accessibility to construct the modal operators of the algebra. The general frames are needed for the other direction. From ideas connected with Stone's theorem, it seems appropriate to use the set of all ultrafilters on the carrier of the algebra as the set of possible worlds of the corresponding frame. But now there are possibly much more ultrafilters than elements in the carrier of the algebra, such that the set of possible worlds contains more falsifiers than the carrier of the algebra. In order to balance the two systems, one has to modify the concept of frame, by restricting the frame $\mathcal{F}$ to the *general frame* $\mathcal{G} = \langle \mathcal{W}, \mathcal{R}, \mathcal{P} \rangle$, where $\mathcal{P} = \{p \in \mathcal{W} : a \in p\}$ and a is some element from the carrier of the algebra. Then$\mathcal{P}$ is the part of the frame corresponding to the carrier of the algebra. Note however as a detail that starting with a general frame, transforming it to a modal algebra and then back to a general frame, results in frame still not isomorphic to the original one. Further restrictions or *refinements* on the general frames have to be imposed [84].

There are also subtle correspondences between higher order logic and general frames, but want finish with these superficial remarks on algebraic models and turn to the intensional perspective.

As stated above, propositions can be seen as functions from possible worlds to truth values. These functions can be constructed in models, but not in frames, where valuations are not yet specified. To built intensions into frames, on may use the structural properties of propositions, given by closure conditions on the domain of possible worlds. These closure conditions can be shifted on the sets of possible worlds, which determine the proposition. So again, one can define a *general frame* somewhere in the middle between frames and models with $\mathcal{P}$ closed under complement, union and a specific modal projections [21, 64].

Fitting [64] draws the general conclusion:

> *If one is merely interested in characterizing logics, general frames are of no more use than frames. The advantage lies in their metamathematical features—their algebraic properties are more natural. But the additional machinery of general frames is natural and may turn out to be useful for applications. It is good to be aware of it.*

With the idea of propositions in the background, one can define an even more general type of semantics, which is not relational anymore, but encompasses the other ones. Why not declare for every possible world the set of propositions necessary in it? This leads to the so called *neighborhood semantics* or *Montague-Scott semantics*. A neighborhood frame is a pair $\langle \mathcal{W}, \mathcal{N} \rangle$, where $\mathcal{W}$ is a set of possible worlds and $\mathcal{N}$ is a mapping from possible worlds to the power set of the set of possible worlds—to the domain of propositions. The name stems from the fact, that in Graph Theory, all vertices connected to a vertex are called its neighbors, such that the mapping $\mathcal{N}$ indeed is a mapping to the neighborhood of a vertex. The semantics of the necessity operator are then straight forward. [120, 64, 35], the relevant rule is of course as usual the necessity operator

$$\mathcal{M}, w_0 \models \Box\phi \leftrightarrow \{w_1 \in \mathcal{W} : \mathcal{M}, w_1 \models \phi\} \in \mathcal{N}(w_0)$$

Neighborhood frames are the characteristic frames for *congruent* logics. As in the case of normal and regular logics, frame constraints can be calculated when supplementing the minimal congruent logic with the usual axioms [120]. Congruency is an important property. It is the presupposition for replacement of tautological equivalents in a formula.

Finally, here is again a negative result together with some general remarks: There are decidable normal calculi without a possible worlds semantics [120, 20, 48]. So from a mathematical point of view, relational semantics seem not entirely satisfactory. But from a philosophical point of view, one still may doubt how far these pathological calculi, constructed only to demonstrate the insufficiency of possible worlds semantics, are important with regards to a formalization of metaphysical modal concepts. For at least the alethic concepts, a possible worlds semantics seems sufficient and satisfactory. But as already mentioned, there are alternatives also from the philosophical side.

3.3 Epistemic Modal Logics

In our century, the investigation of epistemic logics began with the work of Łoś [138] on assertions and Pap [159] on belief. The incorporation of relational semantics is due to Hintikka [97] and was extended by von Kutschera [125]. Lenzen gives a detailed analysis of the most important systems for knowledge and belief until the end of the seventies [131] and proposes a series of new systems [132]. My impression is that philosophical endeavor for epistemic logics was rather limited, such that the number of publications in the eighties and nineties in philosophical journals is rather small and there seems to be not much development since Lenzen. On the other hand, the AI community became more and more involved, starting with the work of Halpern and Moses [89], Levesque [134], Moore [100] and Kraus and Lehmann [118]. Good surveys from this perspective are Halpern and Moses [89] and Meyer et al. [145, 146].

Approaches to epistemic logics usually oscillate between two extreme positions: The first pays tribute to the fact, that epistemic expressions are *de facto* hyperintensional [120] and therefore require a modeling deviant from alethic logics because hyperintensionality destroys the closure under modal consequences and substitution in modal contexts is forbidden as well. Further, the concept of belief needs an appropriate embedding into a probabilistic context. The second one conceives epistemic logic as nothing but a reinterpretation of alethic logic.

Most existing systems try to combine the nice formal features of the normal alethic calculi with the philosophic and linguistic requirements of hyperintensionality. As far as I know, there actually exists no satisfying hyperintensional approach, most logicians—for the sake of formal simplicity—tend to intensionalize the epistemic predicates or attitudes according to one of the following ways:

First, with the assumption of *ideal epistemic predicates*, as in Gödels proof system, which uses an idealized provability predicate. The transfer to epistemic predicates is far from obvious [120], second, with the assumption of *omniscient epistemic subjects*, i.e. the introduction of *rational epistemic agents* [132], whose propositional attitude is closed under consequences (Hintikka [97], Lenzen [132]) and third with the assumption of epistemic predicates or attitudes not referring to sentences or propositions (*explicit epistemic predicates*), but to equivalence classes of equiintensional sentences or propositions (Church [41]) or to the class of sentences or propositions logically implied by the

sentence addressed by the epistemic attitude (von Kutschera [125]) (*implicit epistemic predicates*).

The second way is the most popular one, immediately leading to a formalization that parallels alethic logics. I want to concentrate on this approach, starting with the naive assumption of equivalence of alethic and epistemic operators. But the problem of hyperintensionality lurks right at the beginning.

Normality immediately induces the problem of *logical omniscience*:

First, from $\Box(\phi \to \psi)$ and $\Box\phi$ one can deduce $\Box\psi$ by **K** and *modus ponens*. So, if $\Box$ is interpreted as a knowledge or belief operator, then someone who knows or believes ϕ and $\phi \to \psi$ must believe ψ, too. In the case of the doxastic interpretation, Lenzen for instance replaces **K** by a weaker rule, which I want to call $\mathbf{K}_K$, namely $\Box(\Box\phi \to \psi) \to (\Box\phi \to \Box\psi)$. The reader may translate this formula to natural language and decide, whether this seems reasonable. Second, **nec** forces know or believe all tautologies.

I think it is worth while discussing the consequences of the assumption of normality, which has consequences obviously deviating from common sense and the use of epistemic concepts in natural language. Maybe rational belief or logical omniscience have applications in AI contexts, but it should always be explicitly verified, that this application is sound. Lenzen [132] defends the necessitation rule for epistemic systems; they cannot describe, what some fool believes, because then, they would merely model mental states and have no deductive power at all. He gives no explicit reasons for his preference of normality over the weaker systems, but I suppose him to believe that **nec** gives the clearest concept of rationality. In fact this is far from obvious. It seems reasonable that a rational believer believes some short tautologies, but one may have a machine, that writes down a tautology, which is so long, that our rational believer cannot even read it in his lifetime—is it still *rational* to believe it? But on the other hand, where should be the cutoff for rational beliefs, should one approximate believers by requiring, that he or she must know or believe all consequences that one can derive within n steps for some small n and all tautologies of a certain form and a certain length? And what is a rational belief at all? Lenzen argues, that the fact, that rationality in this context must handle only logical truths and not the larger class of analytic truths, leaves space for prejudice and erroneous behavior, it only makes beliefs *coherent*. Without going into details concerning the concepts of belief and rationality I only want to point out, that there are problems for doxastic logic right from the beginning and there is really no idea how to avoid it. At least it is not clear why Lenzen does not try rules weaker than necessitation; this might be a way into the right direction.

But before looking for alternatives to the rational believer, let us consider further axioms of alethic logic after some general remarks about the nature of knowledge and belief. Knowledge and belief both are propositional attitudes, relating a cognizing subject to a proposition. So compared to alethic logics, epistemic logics require further dimensions or degrees of freedom, namely a *bearer of knowledge or belief*, and maybe also time coordinates, because persons learn new things and forget them and they revise their beliefs. So the complete informal account of an epistemic expression would be something like '*x believes that ϕ in the time interval between t_0 and t_1*'. The bearer of beliefs or knowledge is the most important coordinate, the crucial case is, when operators are iterated. It clearly makes a difference whether I know that I know that I deserve a better position or whether my boss knows that. So it is convenient to index the operators, but the relevant cases with operator iterations are those, where all operators carry the same index and such none.

Another question that should be discussed before any formalization is about the structural properties of beliefs and knowledge. For belief, there is not much general to say: in principle it is a subjective attitude. I have already mentioned the problems of logical omniscience, further insights will be gained along with the discussion of the axioms. For knowledge the case is different, because knowledge has some objective features. The minimal requirement is that knowledge implies truth—one cannot know things that are false—and this of course leads immediately to axiom **T**. It is standard terminology to distinguish two different kinds of knowledge, namely the *weak* concept of knowledge—knowledge as *true belief*—and the *strong* concept of knowledge—knowledge as *justified true belief*.

The stronger concept can be found in Plato's *Menon* or the *Theätet* [132], Lenzen gives separate formalizations of the strong and the weak concept. But let us now discuss the axioms:

- **T** is not convenient under the doxastic interpretation, because the fact that some people believe in God does not imply his existence. Under the knowledge interpretation, as already mentioned, it is acceptable.

- **D** can take over **T**'s rôle in doxastic contexts, because the belief of one proposition implies that one does not believe the negated proposition—if I believe in God, then I do not believe that God does not exist. **D** is weaker than **T**, it can be deduced with the help of **T** and the fact that non-modal implication is transitive.

- **4**, under the knowledge interpretation, means that if I know something, I know that I know it. This principle is called *positive introspection* and is fairly reasonable for knowledge. But if I believe something, is it reasonable to say, that then I believe that I believe it? Lenzen [132] gives arguments in favor of this.

- **E** is equivalent to $\neg\Box\phi \rightarrow \Box\neg\Box\phi$. Under the knowledge interpretation, this adds the *principle of negative introspection*: if I do not know something, I know that I do not know it—the same with belief. It is open to discussion, if one is willing to accept the principles of positive and negative introspection. Lenzen argues, that it is too strong for both knowledge and belief, but in many AI publications, it is considered adequate.

- **4.4** is equivalent to the formula $(\phi \wedge \neg\Box\neg\Box\phi) \rightarrow \phi$, sounding reasonable for weak knowledge: I know something, if it is true and it is not the case that I know that I do not know it. Lenzen argues, that **S4.4** is exactly the right system for weak knowledge, thereby deviating from Hintikka, who takes **S4** for weak knowledge.

- **G** is relevant for strong knowledge, according to Lenzen. Therefore the logic of strong knowledge is **S4.2**. But this demission is highly controversial in the literature.

- Lenzen further discusses the characteristic formulas of **S4.2.1**, **S4.3** and **S4.3.2** with respect to their applicability. He concludes, that it is merely a matter of taste to make a decision, but in general it is very difficult to judge the more or less complex axioms between **S4.2** and **S5**—as already stated, Lenzen himself prefers **S4.2** as the logic of strong knowledge [13].

To summarize the above considerations, note that there are many concurring models for belief and knowledge and it is far from clear, which one is correct. But perhaps one should take a pragmatic pluralistic position and relativize the concept of correctness to applicability to certain contexts. Lenzen, which as far as I know gives the most elaborate discussion of epistemic logic so far, decides for **S4.2** for strong and **S4.4** for weak knowledge. His calculi for belief are not normal, because they use the doxastic version $\mathbf{K}_K$ instead of $\mathbf{K}$. Lenzen altogether gives eight different calculi, not only for knowledge and belief, but also for conviction and combined epistemic modalities. In multi-agent applications in AI it is standard to consider concepts like *common knowledge*. These concepts do not present any philosophical complications and are no further investigated by Lenzen. Possible worlds semantics for epistemic logics are analogous to the alethic case: in the doxastic case, it is appropriate to use a kind of neighborhood semantics with a probability measures instead of accessibility functions. We do not want to go into details and refer the reader to Lenzen [132].

But now to weakenings of normal systems. Fitting [64], in opposition to Lenzen believes, that $\mathbf{K}$ is too strong even for knowledge. As we already pointed out, the crucial difference between normal and regular systems is the formula $\Box\top$—in regular systems one is not forced to know or believe all consequences of tautologies, but one can show that if anything is known, then so is every tautology. So also regular logics are not at all satisfactory. Fitting pleads for **S2** or **S3** as being close to reasonable logics of knowledge.

Finally, I want to mention the approaches of Levesque [134] and Fagin and Halpern [56] for doxastic logics. Both approaches employ standard syntax—deontic **S5**, a normal logic, but change the semantics.

Levesque, following the terminology of Church and Carnap, distinguishes between *implicit* and *explicit* belief, submitting only implicit belief to logical omniscience. Explicit belief is treated with a *situation semantics* à la Barwise and Perry [18]. This approach is not appreciated by all researchers, because it has strong resemblances with relevance logics [145]. On the other hand, there are—with the success of linear logic—many people, who believe, that relevance logic is even more convenient to natural language deduction than material implication, which is designed first of all for reasoning within mathematics.

Fagin and Halpern supply the frames with an *awareness function* from possible worlds to sets of formulas, of which the agent is aware. Now in their approach, explicit belief is implicit belief plus awareness.

In general there is promising current research on these items in AI, but my

[13] Meyer et al. [145], in opposition to Lenzen, prefer **S5** as the appropriate system. They seem to be aware of the philosophical controversy about **S5** but their main argument in favor of **S5** are its *nicer technical properties*!

overall impression is, that there is almost no coordination with philosophical investigations. The problem of logical omniscience remains and it is even unclear whether it is a problem at all. Beliefs are first of all private propositional attitudes, without commitment to truth or laws of nature or logic. But this concept of belief has nothing to do with deduction. On the other hand, if one looks for a principle of rationality, it is not at all clear that it must be as strong as the one, that Lenzen supposes. In fact, rationality is a vague and somewhat obscure philosophical concept, and standards for rationality convenient to AI contexts seem difficult to derive—if one wants to go further than defining a cut-off in the way described above. On the other hand, it is questionable whether, at the current state of the art, it is not better to employ heuristics to proceed.

3.4 Deontic Modal Logics

Deontic logic is the most problematic and the least developed of the modal logics. Though deontic problems were already discussed by the ancient Greeks, it was only in our century that formal theories were developed. Historically interested readers should consider Kalinowski [107], Rescher [174] or Knuuttila [115]. For an introduction to the elder systems consider Føllesdal and Hilpinen [65] and for an exposition of formal systems Åqvist [10]. The first prominent modal theory of our century is due to Ernst Mally [141]. Von Wright [205, 204] developed a highly influential system **OS**, the so called **O***ld* **S***ystem*. But before going into details, let us again make some general remarks.

Deontic logic is the logic of *obligation* (prescription), *prohibition* (forbiddance), *permission* and *commitment* (conditional obligation). The corresponding natural language terms are

obligation:	shall, ought, must,
prohibition:	shall not, ought not, must not,
permission:	may, is allowed to,
commitment:	if ϕ is the case, then it shall (must, ought to) be, that ψ.

Now the question is, what kind of sentences are built with the help of these terms—are they *prescriptive sentences*, leading to legal, ethical or social *norms* or*descriptive sentences*, where the deontic predicates express a relation between a *norm addressee*, a *proposition* and a *norm source* or an *authority*?

Only descriptive sentences have a truth value and therefore there cannot be a semantics—at least not in the usual model theoretic sense—for prescriptive modal logics. So it is not at all clear, if a deontic logic should have a (possible worlds) semantics [14].

[14] According to Åqvist [10], this goes back to Bentham and Hedenius. In Hedenius terminology, prescriptive sentences are called *genuine* and descriptive ones *spurious*. Wedberg calls the former ones *internal* and the latter ones *external*, Stenius uses the distinction between *modal* and *factual* sentences, Hansson [92] between *imperative* and *descriptive* ones and von Wright [207] *norms* and *norm-propositions* in the same spirit. Åqvist recommends studying previous attempts for logics of commanding [60, 91, 16, 17] to find systems, where addressees and authorities are made explicit.

But what are norms? Almost tautologically, norms are sentences with *normative impact* on human or social behavior. There are at least three different approaches to norms, namely the *linguistic approach*, according to which norms are certain linguistic entities, the *naturalistic approach*, according to which norms are behavioral regularities, and the *platonist approach*, according to which norms are ideal duties.

Weinberger [203] gives four properties of norms: Norms are no descriptions of reality, they cannot be verified, they cannot be reformulated as declarative sentences, and there is no deductive relation between norms and declarative sentences. But nevertheless it is common sense that one can *deduce* new norms old ones. This deduction of course must not be meant in the logical sense. It is also obvious, that systems of normative sentences or *normative codices* have properties like redundancy and consistency, which should be open to (logical) investigation. In von Wright's first system, for instance, deontic logic was conceived as a logic of norms, there was no mechanism to iterate deontic *predicates*, which on his account are not applied to norm-sentences, but to *act-types*, in order to achieve generality. Readers interested in the philosophical background of these problems are recommended to read the writings of Castañeda, for example [34], where moral and legal norms are investigated.

Instead of this approach, most logicians conceive deontic logic as the logic of norm-sentences or norm-propositions, such that one—as in the alethic case—has sentence-forming *deontic operators*. In this case the whole deductive apparatus of alethic modal logics applies to the deontic case and a semantics can be given. But after this decision in favor of a formalization in the shadows of alethic logics, there still remain two approaches, the *monadic* approach and the *dyadic* one.

The older monadic approach was soon troubled by paradoxes. I present it first to show that for deontic logics a simple transfer from alethic ones is not appropriate.

According to deontic tradition, there are three deontic operators, namely '**P**', '**O**' and '**F**', for *permission*, *obligation* and *prohibition*. For reasons, that will become clear later, I follow this tradition and do not use the usual modal operators '$\Box$' and '$\Diamond$'. These operators can be inter-defined as follows:

$$\begin{aligned} \mathbf{F}\phi &:= \neg\mathbf{P}\phi, \\ \mathbf{P}\phi &:= \neg\mathbf{O}\neg\phi. \end{aligned}$$

With these operators, one can build the deontic systems like the alethic ones, translating '**O**' for '$\Box$' etc. and all further concepts, for instance of a *deontic alphabet*, of a *deontic formula*, of *normality*, *regularity* and *congruence* are adapted from alethic logics. But already this basic definition is disputed in the literature [10]. Some logicians believe that the necessitation rule, which states, that all logical truths are obligatory, is not acceptable: von Wright, for example in his first system [204] explicitly denies that tautologies are obligatory. But instead of discussing this matter, let us consider the additional axioms of alethic modal logics.

- **T**, or the equivalent *ab-esse-ad-posse* principle $\phi \to \Diamond\phi$ is not acceptable for deontic logics. Its translation would be $\phi \rightarrow \mathbf{P}\phi$—whatever is actual is permitted. The opposite is the case: whoever violates a duty realizes a fact that is *not* permitted. Therefore one must weaken the axiom to the analog of the alethic axiom **D** in order to obtain $\mathbf{O}\phi \rightarrow \mathbf{P}\phi$ —whatever is obligatory is allowed. Now one can understand the name of the axiom **D**. It plays exactly the rôle that **T** plays in alethic logics.

- One could use a further axiom $\mathbf{T}_D$ ($\mathbf{O}(\mathbf{O}\phi \rightarrow \phi)$ instead of **T**, which can be deduced from **T** by simple application of the necessitation rule. With this axioms the formula $\mathbf{OO}\phi \rightarrow \phi$ can be deduced, if the system contains **K**. This formula is essential for the elimination of potences of modal operators in stronger systems. As noted above, this does not work only with axiom **D**. $\mathbf{T}_D$ states, sloppily formulated, that you are obliged to realize obligations. This seems reasonable.

- The translation of the alethic formula **E**—$\mathbf{P}\phi \rightarrow \mathbf{OP}\phi$—states that whatever is allowed ought to be allowed. It is open to doubt, that this principle should be accepted, but is is not counterintuitive.

- The Brouwer formula **B** ($\phi \rightarrow \Box\Diamond\phi$) is similar to the *ab-esse-ad-posse* principle considered above, it is unacceptable under the deontic interpretation. One should weaken it like **T**.

These considerations are already sufficient to construct the standard monadic deontic systems, corresponding to the alethic systems **K**, **T**, **S4**, **S5**, but note that the systems that presented here following [10], deviate from the systems that Bull and Segerberg [30] or Chellas [35] call 'deontic'. In our systems—Åqvist calls them *Smiley-Hanson* systems—the formula $\mathbf{T}_D$ is assumed as a replacement of the alethic formula **T**. Addition of formula **D** then gives stronger versions of the systems obtained so far. Bull and Segerberg, Chellas and others *start* with axiom **D**. In Chellas' calculus, $\mathbf{T}_D$ is added as a weakening of **E** [35]; in **S5** it is redundant, because it can be deduced from **E**, **K**, **nec** and propositional logic.

Summarizing, here is a list of the formulas considered so far:

$$\begin{array}{rc}
\mathbf{K} & \mathbf{O}(\phi \rightarrow \psi) \rightarrow (\mathbf{O}\phi \rightarrow \mathbf{O}\psi), \\
\mathbf{D} & \mathbf{O}\phi \rightarrow \mathbf{P}\phi, \\
\mathbf{T}_D & \mathbf{O}(\mathbf{O}\phi \rightarrow \phi), \\
\mathbf{4} & \mathbf{O}\phi \rightarrow \mathbf{OO}\phi, \\
\mathbf{E} & \mathbf{PO}\phi \rightarrow \mathbf{O}\phi, \\
\mathbf{B}_D & \mathbf{O}(\mathbf{PO}\phi \rightarrow \phi).
\end{array}$$

These formulas define the *Smiley-Hanson systems* in Lemmon code. Again, as in the alethic case, we want the systems given below to be the smallest normal systems of modal logic containing the formulas on the right side of the equations as axioms:

$$\mathbf{OK} \;:=\; \mathbf{K},$$

$$
\begin{aligned}
\mathbf{OT} &:= \mathbf{KT}_D, \\
\mathbf{OS4} &:= \mathbf{KT}_D\mathbf{4}, \\
\mathbf{OB} &:= \mathbf{KT}_D\mathbf{B}_D, \\
\mathbf{OS5} &:= \mathbf{K4E},
\end{aligned}
$$

and further

$$
\begin{aligned}
\mathbf{OK}^+ &:= \mathbf{OK} \cup \{\mathbf{D}\}, \\
\mathbf{OT}^+ &:= \mathbf{OT} \cup \{\mathbf{D}\}, \\
\mathbf{OS4}^+ &:= \mathbf{OS4} \cup \{\mathbf{D}\}, \\
\mathbf{OB}^+ &:= \mathbf{OB} \cup \{\mathbf{D}\}, \\
\mathbf{OS5}^+ &:= \mathbf{OS5} \cup \{\mathbf{D}\}.
\end{aligned}
$$

Deontic logics according to Bull and Segerberg and Chellas are constructed similar to the systems given by Åqvist by replacing $\mathbf{T}$ with $\mathbf{D}$.

The *semantics* also carries over from alethic logic, the accessibility relation of the alethic systems is now called *the relation of deontic alternative* or *co-permissibility*. Now what is the idea behind deontic possible worlds? According to relational semantics, a formula $\mathbf{O}\phi$ holds in a possible world iff ϕ holds in all its deontic alternatives. So the deontic alternative relation gives a notion of *deontic perfection*—in deontic alternatives all obligations are actual. This really captures Leibniz' idea of the *best of all possible worlds* and the name 'idealization' for the frame constraint of $\mathbf{D}$ is evident, too: For every possible world, there is a better one. $\mathbf{T}$ would impose reflexivity, but this clearly would destroy seriality between the possible worlds. Transitivity seems to be a natural requirement for the series of possible worlds, too. This calls for $\mathbf{4}$. The Brouwer formula, imposing symmetry, like $\mathbf{T}$ would destroy the ordering of possible worlds. Its weaker form $\mathbf{B}_D$ only imposes symmetry on deontic alternatives. $\mathbf{E}$ has the frame constraint of euclidity—alternatives of a possible world are alternatives of each other. Adding this axiom to $\mathbf{S4}$ leads to $\mathbf{S5}$, such that the frame constraint is an equivalence relation. This seems to be too strong for deontic logics. Now take a closer look at the frame constraint of $\mathbf{T}_D$: it imposes the condition, that deontic alternatives of a possible world are reflexive.

Suppose that $\mathbf{T}_D$ holds and that in our world an obligation is violated—let $\mathbf{O}\phi$ and $\neg\phi$ be true. Then our world must be a worst of all possible worlds [35]; our world cannot be better than any other world, because in all deontic alternatives of some world $\mathbf{T}_D$ must hold and therefore no obligation can be violated. This requirement is indeed very strict—maybe too strict. Consider the directed graph corresponding to the frame constraint of $\mathbf{T}_D$. It has two vertices, a line between the two and a loop at the second vertex. A weaker graph would be to open this loop, i.e. to modify it to a graph with three vertices and two lines, not between the same vertices. One possible frame constraint would be density, which corresponds to the formula $\mathbf{OO}\phi \rightarrow \phi$. On the other hand, this formula is an easy consequence of $\mathbf{T}_D$ with the help of $\mathbf{K}$ and *modus ponens*. In general it might be worth while investigating the correspondences between the *Gestalt*-change of the directed graphs and applications of rules and axioms. Chellas [35] believes, that **S4** supplemented with the last formula might be appropriate for applications. Of course, it would again be possible to turn from normal to regular or even to congruent systems, but this is obviously not very popular and I do not want to go further into details.

Kanger [110] gives an interesting translation from deontic to alethic modal logics, employing a new constant 'Q' for an ideal commitment. He defines

$$\begin{aligned}\mathbf{O}\phi &:= \Box(Q \to \phi),\\ \mathbf{P}\phi &:= \Diamond(Q \wedge \phi),\\ \mathbf{F}\phi &:= \Box(Q \to \neg\phi).\end{aligned}$$

Then, by addition of the axiom

$$\mathbf{Q} \qquad \Diamond Q$$

to the systems of alethic modal logics, containing 'Q' as a new member of the language one obtains, following the terminology of [10], the systems $\mathbf{K}^+_Q$, $\mathbf{T}^+_Q$, $\mathbf{S4}^+_Q$ $\mathbf{S5}^+_Q$ etc. For these systems, the alethic frames $\langle \mathcal{W}, \mathcal{R}\rangle$ must be extended to triples—augmented frames—$\langle \mathcal{W}, \mathcal{R}, \mathcal{O}\rangle$, where $\mathcal{O} \subseteq \mathcal{W}$ is the *set of best or optimal worlds*. Now we have the semantic requirement

$$\mathcal{M}, w \models Q \Leftrightarrow w \in \mathcal{O}.$$

The frame constraint on $\mathbf{Q}$ is

$$\forall w_0 \exists w_1 (\mathcal{R} w_0 w_1 \wedge w_1 \in \mathcal{O})$$

in addition to the usual alethic semantics. All of these alethic systems can be proved sound and complete [10]. One can further show [192], that there is a translation from these alethic systems to the deontic systems with the corresponding name.

In the *dyadic* approach, monadic operators $\mathbf{O}$ and $\mathbf{P}$ are now replaced by the dyadic operators $\mathbf{O}_\phi$ and $\mathbf{P}_\phi$, that formalize *relative modalities*, with the monadic operators as special cases:

$$\begin{aligned}\mathbf{O} &:= \mathbf{O}_\top,\\ \mathbf{P} &:= \mathbf{P}_\top,\\ \mathbf{F} &:= \neg\mathbf{P}_\top.\end{aligned}$$

Introducing a new *predicate*—in the monadic case it was a constant—*alethic definitions* of the deontic operators can be obtained:

$$\begin{aligned}Q &:= Q\top,\\ \mathbf{O}_\phi\psi &:= \Box(Q\phi \to \psi),\\ \mathbf{P}_\phi\psi &:= \Diamond(Q\phi \wedge \psi),\\ \mathbf{F}_\phi\psi &:= \Box(Q\phi \to \neg\psi).\end{aligned}$$

Dyadic axioms correspond to the monadic ones with slight modifications:

$$\begin{array}{rl}
\mathbf{K}^2 & \mathbf{O}_\phi(\psi \to \chi) \to (\mathbf{O}_\phi\psi \to \mathbf{O}_\phi\chi), \\
\mathbf{D}^2 & \mathbf{O}_\phi\psi \to \mathbf{P}_\phi\psi, \\
\mathbf{T}^2_D & \mathbf{O}_\phi(\mathbf{O}_\phi\psi \to \psi), \\
\mathbf{4}^2 & \mathbf{O}_\phi\chi \to \mathbf{O}_\psi\mathbf{O}_\phi\chi, \\
\mathbf{E}^2 & \mathbf{P}_\phi\mathbf{O}_\psi\chi \to \mathbf{O}_\psi\chi.
\end{array}$$

These axioms lead to the deontic systems $\mathbf{O^2K}$, $\mathbf{O^2T}$, $\mathbf{O^2S4}$, $\mathbf{O^2S5}$ etc. The alethic systems $\mathbf{K}^2_Q$, $\mathbf{T}^2_Q$, $\mathbf{S4}^2_Q$, $\mathbf{S5}^2_Q$ etc. are the usual alethic systems with augmented vocabulary, as in the monadic case.

Again, we have two different frames, the deontic ones and the augmented alethic ones. In the deontic case, the frames are similar to those of the usual alethic systems, except for the relation of deontic alternativeness, $\mathcal{R}^2_O$,which now is defined as a function from the set of sentences into the set of binary relations on the set of possible worlds. The truth conditions for the deontic operator $\mathbf{O}_\phi$ have to be relativized to the condition

$$\mathcal{M}, w_0 \models \mathbf{O}_\phi\psi \Leftrightarrow \text{ f.a. } w_1 \in \mathcal{W}(\mathcal{R}_\phi w_0 w_1 \Rightarrow \mathcal{M}, w_1 \models \psi),$$

and a similar condition holds for $\mathbf{P}_\phi$.

The set constraints have to be relativized, too. For example, the set constraint for the axiom $\mathbf{4}^2$ is

$$\forall \phi, \psi, w_0, w_1, w_2(\mathcal{R}_\phi w_0 w_1 \wedge \mathcal{R}_\psi w_1 w_2 \to \mathcal{R}_\psi w_0 w_2).$$

The augmented alethic models are again triples $\langle \mathcal{W}, \mathcal{R}, \mathcal{O} \rangle$. The alethic accessibility relation can be taken over from the monadic case, but now $\mathcal{O}$ is not a set anymore, but a function from the set of sentences of the alethic language enriched by the predicate 'Q', into the powerset of possible worlds. So, $\mathcal{O}$ is relativized to the conditions. But then, the truth condition for Q must also be modified:

$$\mathcal{M}, w \models Q\phi \Leftrightarrow w \in \mathcal{O}(\phi).$$

Soundness and completeness proofs for these systems as well as a translation theorem corresponding to that of the monadic case can be obtained [11]. Today, there are some considerable attempts to use dyadic systems that go much further than the systems given in this text, consider for example [105].

But now let us come back to informal questions. Why are there different attempts to formalize deontic concepts? Here is a brief consideration of the monadic systems, a detailed discussion may be found in [10].

Basically, the discussion of and the search for deontic systems was always guided by paradoxes, the most famous ones are [10, 54]

- Ross's Paradox,
- Prior's Paradox of derived obligation,
- Chisholm's contrary-to-duty imperative paradox,

- the dilemma of commitment and detachment,
- *prima facie* versus actual obligation, the *ceteris paribus* proviso,
- the ought-implies-can problem,
- CastañVeda's Good Samaritan paradox and the Jephta dilemma,
- the problem of act-utilitarism in relation to deontic logic.

For a discussion of these paradoxa and problems consider the references given above. We restrict ourselves to a brief sketch.

Ross's paradox rests on fact, that one can prove the formulas

$$\mathbf{O}\phi \to \mathbf{O}(\phi \vee \psi), \qquad \mathbf{P}\phi \to \mathbf{P}(\phi \vee \psi).$$

in normal monadic systems— $\phi \to (\phi \vee \psi)$ is a tautology and then the result is easily obtained with **nec**, **K** and *modus ponens*. Now consider the natural language sentences

(i) You post the letter, (ii) You burn the letter.

The translation of the 'ought'-formula is 'If I ought to post the letter, I ought to post or to burn it.' This looks paradoxical, because I can realize this obligation by burning the letter!

For the second formula consider the sentences

(i) You smoke, (ii) You kill somebody.

In this case, the translation is 'If I am allowed to smoke, then I am allowed to smoke or to kill somebody.' These paradoxes are much under discussion among philosophers. Some think that there is no paradox, but only confusion of material implication, but the majority finds it is worthwhile debating. Von Wright [209] gives a detailed analysis of the paradox concluding that it rests on subtle philosophical questions like free choice. Without going into details we refer the reader to [209, 10] and the references given there and in [54].

Prior's paradox rests on the fact, that sentences of either the form

$$\begin{gathered}\neg\phi \to (\phi \to \mathbf{O}\psi, \quad \mathbf{O}\psi \to (\phi \to \mathbf{O}\psi),\\ \mathbf{O}\neg\phi \to \mathbf{O}(\phi \to \psi), \quad \mathbf{O}\psi \to \mathbf{O}(\phi \to \psi)\end{gathered}$$

can be derived in all of the Smiley-Hanson systems of monadic deontic logic although they do not appear valid—the reader may check this by inserting appropriate sentences. According to Åqvist, there are two ways of treating these paradoxical sentences, namely to 'formulate them away', and to reconsider the applicability of monadic deontic logic to natural language discourse.

The first approach shows that simple manipulations reduce the formulas of the first line to harmless propositional tautologies and those of the second line to Ross's formulas. The second approach takes the problem seriously and claims, that monadic deontic logic is not able to formalise natural language in a satisfying way. For instance, if we attempt to read the formula $\mathbf{O}(\phi \rightarrow \psi)$ as 'ϕ commits us to do ψ', then we must face the fact that a forbidden act commits us anything. So von Wright [206] introduces *dyadic operators* similar to those introduced above and opened the way to a new line of research. In the dyadic systems given above, the formulas corresponding to those of Prior's paradox are in fact no longer derivable. The introduction of dyadic deontic operators therefore seems to be a more promising approach to concepts of commitment, obligation and permission and one may draw the general conclusion that deontic logic cannot be simply alethic modal logic with reformulated axioms.

The *Chisholm contrary-to-duty imperative paradox* is the following set of sentences:

(i) It ought to be that a certain man go to the assistance of his neighbors.
(ii) It ought to be that if he go, he tell them he is coming.
(iii) If he does not go, he ought not tell them he is coming.
(iv) He does not go.

It is commonly accepted on an intuitive basis, that the above sentences are *non-redundant* and *consistent*, but monadic logics seem unable to formalize the sentences in a way that pays tribute to our intuitions. So again, the right line of attack seems to be the use of dyadic deontic logic and it can indeed be shown [10, 105], that dyadic logics can handle the problem in a satisfactory way.

In brief, Ross's, Prior's and Chisholm's paradox can be overcome by the use of dyadic deontic logics, but—as a careful reader will have noticed—there are still five problems open. These problems are more difficult to state so that we cannot go into the details and refer the reader to [54, 10].

The *dilemma of commitment and detachment* arises from the fact, that one wants to be able to derive unconditional from conditional obligations under certain circumstances. But a straightforward solution leads again to a violation of Chisholm's paradox. This is the dilemma, but there are also attempts of solutions to it [105].

The *prima facie vs. actual obligation* is the problem, that obligations can be overruled by a change of situation.

The *Good Samaritan paradox* [33] is the following argument:

(i) If Bob pays \$500 to the man he will murder one week hence, then Bob will murder a man one week hence.
(ii) It ought to be, that Bob pays \$500 to the man he will murder one week hence.
(iii) It ought to be, that Bob will murder a man next week.

The argument deducing (iii) from (i) and (ii) is obviously not sound, but nevertheless, it can be obtained in all monadic systems given above.

The other problems are not so easy to introduce, but the general claim is [54], that all the systems considered so far are not sufficient. Van Eck therefore introduces a combination of time and deontic operators together with quantification—another example of a multi-modal approach is given in [35]—which helps him to face *all* the problems stated above. This makes van Eck's approach very promising and we highly recommend it to further investigation. Another system with quantifiers is that of Castañeda [34], which up to now has perhaps not received the attention it deserves. From a philosophical point of view, Castañeda was probably for the seventies and eighties what von Wright was for the decades before—the philosopher who made the deepest investigation in the formal structure of morality and ethics.

4 Quantificational Modal Logics

4.1 General Remarks

It is a pedagogic tradition in non-modal logics to introduce the propositional calculus first and reserve quantificational theory for a later chapter. The reason for this is merely to teach the reader some fundamental methods in an easy setting. Although there are some differences between propositional and quantificational first order logic—for example decidability and the proof procedures depending on that—there is no real methodic difference concerning the metatheorems. In modal logic the situation is completely different. So the reason, that I have not treated quantificational modal logic up to now is not a pedagogic one, but has to do with the problems connected with the incorporation of quantifiers. The problems are not syntactical ones—a syntactical treatment of quantifiers in modal logics is straight forward—it is the semantics, both formal and informal, which make the enterprise difficult. A consideration of the epistemic systems popular among the AI community [145, 146] reveals the almost total lack of quantificational systems—the same holds for deontic logics. But I think nevertheless that one should try to face the problems connected with quantification. Maybe the restriction to propositional calculi can help to solve some interesting small (engineering) problems, but everyone who subscribes to cognitive adequate systems must face the fact, that contemporary epistemic AI systems, under a wider perspective, are far from satisfactory, and that in general, any serious logical theory must be able to handle quantification. Therefore we give a perspective on quantified alethic modal logics, following Fitting [64] and Kreiser et al. [120] for the formal, Garson [81], Loux [140] and Kripke [124] for the philosophical aspects. Lenzen [132], who gives a sketch of a quantificational modal logic of knowledge and belief, may serve as a reference for the more specific epistemical problems, van Eck [54] for the deontic ones. In general, there seem to be few serious attempts for quantificational epistemic deontic logics, but we hope that, with the new strong interest in these domains, this situation may change.

The novice may wonder why quantified modal logic is considered difficult. This rhetorical remark stands at the beginning of Garson's chapter on quantification in modal logic in the *Handbook of Philosophical Logic* [81]. But one page later, the reader is confronted with a huge 'Quantified Modal Logic Road map' reflecting the spectrum of approaches to modal quantification, and an orientation in the philosophical literature leaves one with the impression that there are at least as many approaches to the field as there are philosophers or logicians working in it.

The *syntactic specifications* are nevertheless straight forward: simply build the modal operator(s) into the set of formulas. Again, as in the case of propositional logic, the *language L_{QM} of quantificational modal logic* is a language L of first order logic provided with the additional symbol '$\Box$': $L_M = L \cup \{\Box\}$ and all further syntactic concepts for formation of terms and formulas can be imported from the non-modal case, except the fact, that there is a formula-formation rule for the modal operator. The *calculus* is specified exactly as in non-modal first order logic, by adding the axioms of propositional modal calculus to these non-modal first order axioms and supplying the system with any of the modal rules (necessitation, monotonicity or congruence), one obtains a hierarchy of systems similar to the propositional case, with the only difference that they now may include equality.

The primary semantics for these systems is still straightforward. Instead of mapping formulas to truth values relative to possible worlds, one now attaches first order quantification domains or universes to any of the possible worlds. So a *first order frame* is a triple $\langle \mathcal{W}, \mathcal{R}, \mathcal{D} \rangle$, where $\mathcal{W}$ is a non-empty *set of possible worlds*, $\mathcal{R}$ is a binary relation on $\mathcal{W}$, the *accessibility relation* and $\mathcal{D}$ is a mapping from elements w of $\mathcal{W}$ to non-empty domains $\mathcal{D}(w)$ for the possible world w. This image is even more intuitive than its propositional counterpart. The worlds are not mere points anymore but carriers of objects, functions and relations as in a non-modal universe, but in general, one must impose restrictions on the universes. For example, if you live in a world in which certain facts hold, then the substrate that generates these facts should exist in the accessible worlds, too. That makes life much easier. One possibility is, to simply *fix the domains*. Another is to require *monotonicity*, i.e., that no world accessible from another world be smaller than the former.

4.2 The Maze of Metaphysics

Before entering details, let us consider a general problem connected with quantification in general with some impact on modal logics. There are basically two different theories of quantification, namely *objectual quantification* and *substitutional quantification.*

A mathematically educated logician might make a career without realizing this differentiation but philosophers might fail their BA, when not aware of it. The objectual interpretation is perhaps best highlighted by Quine's famous slogans *to be is to be the value of a variable* and *no entity without identity*, because quantification under this interpretation introduces *ontological commitments* and *standards of ontological admissibility.* According to the theory of eliminability of proper names, it is the bound variables and not the names that carry ontological information.

Under the objectual interpretation, a formula of the form $\forall x \phi$ is true iff for all objects in the domain, ϕ holds; a formula of the form $\exists x \phi$ is true iff for at least one object in the domain, ϕ holds.

Under the substitutional interpretation, a formula of the form $\forall x \phi$ is true iff all substitution instances for x in ϕ are true, a formula of the form $\exists x \phi$ is true iff at least one of the substitution instances for x in ϕ is true.

The objectual interpretation is the standard approach. Now what is the impact of the two theories on quantified modal logic? Let me give a rough description. Quine gives the following argument:

$$\begin{array}{ll} (1) & \Box(9 > 7), \\ (2) & 9 = \text{ the number of planets}, \\ (3) & \exists x \Box(x > 7), \text{ by existential generalization.} \end{array}$$

Another chain of inference would be

$$\begin{array}{ll} (1') & (9 > 7), \\ (2') & \exists x(x > 7), \text{ by existential generalization}, \\ (3') & \Box \exists x(x > 7), \text{ by necessitation.} \end{array}$$

The first derivation leeds to a *de re*, the second to a *de dicto* formula. The first derivation is fine. Indeed there *is* necessarily an object greater than seven—namely the object denoted by the definite description 'the number of planets'. But on the other hand, is it necessary that there is an *object* greater than seven? If you think that it is the number nine, then you are wrong, argues Quine, because nine is the number of planets and this number is only contingently greater than seven. What happens basically is the quantification into an intensional context. This problem of *quantifying in* is the main problem connected with quantificational intensional logics and most of the difficulties with those logics are consequences of it. We will return to this problem, later. Quine sees only one way out of this dilemma—the assumption of *essentialism*, the thesis, that objects have some of their properties necessarily and others contingently. So the number of planets is only contingently nine, but nine is necessarily seven plus two. The underlying problem is a clash between the quantification in intensional contexts and the congruence property of equality—the Leibniz principle of substitutability of equals. Quine's argument is that modal logic is not the appropriate mechanism to account for this kind of quantification and so the modal logics discussed in this text are inconvenient for a formalization of intensionality. There are many replies to Quine, some defending *de re* modalities and essentialism, others trying to detect failures in Quine's argument. For a detailed discussion consider [59]. One objection to Quine is, that his argument rests on the assumption of objectual quantification. If one assumes substitutional quantification, then these problems, including the commitment to *de re* modalities disappear [86].

Most of the quantificational modal approaches offered by the different authors react to and try to handle the problem proposed by Quine and another problem that comes with identity, namely the fact, that all identities are necessary, which can be derived from the obvious logical truth $\Box(x = x)$. My general impression is that none of the existing approaches is able to handle the problems in a completely satisfying way. So we will not investigate the best of all possible theories, but briefly sketch a variety of these.

The various approaches to quantificational modal logic can be classified according to several aspects, their *informal account of possible worlds*, their handling of the *transworld identity problem*, and their *formal semantic structure*.

The first and the last aspect do not need any further discussion. The transworld identity problem is the problem of giving adequate criteria for tracing an object through possible worlds. Compared to the two other aspects given, it may seem as a somewhat limited and special problem, but this is not true: the problem of transworld identity is—according to literature—*the* central problem of quantified modal logic. The three aspects with regard to modal logics are of course not independent of another. It is clear, that formal semantical constraints have an impact on the philosophical positions—for example via ontological commitments, but in general, it is the philosophical position, which should be seen as the foundation for the formal.

Now there is a large amount of philosophical or semi-formal attitudes towards modal quantification: the *conceptualism approach* (Kripke), and the *realism approach* with a bifurcation in the *possibilism* (Lewis) and the *actualism approach* (Platinga, Stalnaker). Discussion of these positions follows below.

Transworld identity led to the following principal positions: the (trivial) *identity theory*, Lewis's *the counterpart theory*, Kripke's *theory of rigid designators*, the *predicate abstraction theory* by Stalnaker,Thomason, Konolige and Fitting and Platinga's *trivial solution*. These positions will be discussed in the sequel, too.

Let me now expose the formal details of the various systems. Garson [81] gives a long list of different approaches to modal quantification. We follow his terminology:

1. objectual domain
    - 1.1 rigid designators
        - 1.11 fixed domain
        - 1.12 world-relative domains
            - 1.121 free logic
            - 1.122 classical logic and term elimination
            - 1.123 classical logic, truth value gaps and nested domains
            - 1.124 classical logic, truth value gaps and no domain restrictions
    - 1.2 non-rigid designators
        - 1.21 local terms
        - 1.22 global terms
2. conceptual domain
    - 2.1 fixed domain
    - 2.2 world-relative domain
3. substantial domain

3.1 standard predicates

3.2 intensional predicates

Let us briefly describe the most important concepts of this list, further discussion may be found in [81]. One of the concepts at the highest level in our scheme is the concept of *objectual domain*. This concept denotes the standard approach, where only objects are admitted to the domains at the possible worlds. Unlike the two other approaches at the same level, this one needs no further motivation. If the decision for an objectual domain approach is made, one has another alternative, that between *rigid terms* and *non-rigid terms*.

4.3 Rigid Designators

The idea of rigid designators is due to Kripke and stems from his metaphysical convictions [122, 124]. From a semantical point of view, it is very easy to characterise this position. It is simply required, that *proper names are rigid denotators iff they denote the same object in every possible world*. According to our definition of intensions, that means, that *the intension of a rigid designator is a constant function*. Kripke, in order to cope with these problems, assumed, that proper names denote directly, so that they have only ex- but no intensions. If we subscribe to Kripke's deviation from Frege sematics, it is straight forward to define *rigid interpretations* and consequently *rigid models* [64], where the denotations of all constant symbols—all proper names—are determined by their valuation.

For Kripke, the problem of transworld identity does not exist—objects in different possible worlds are identical *by stipulation*. This must be seen in the general framework of *conceptualism*. Possible worlds—according to Kripke—are not entities existing in reality, but cognitive alternatives to the *one* real world—ideas of how the world *could be*. This again is reminiscent to Wittgenstein's *logical space*, where the general structure is prescribed by logic (all the laws of nature, causality etc. are in their structure reducible to logic) and a substance, that, when filled into the logical frame, determines the real world. So, on Kripke's conceptualist account, the general structure of the world is given and the possible worlds are constructed by the human mind by filling in different configurations of substance. With this rigid ontology, of course, everything equal is necessarily equal and Quine's argument concerning the number nine and the number of planets is trivial—Kripke does not try to avoid *de re* modalities, he shoes, how to handle them. 'The number of planets', on Kripke's account, is a *non-rigid* designator for the number nine holding in some worlds and in some others not—by definition. Another paradigm for Kripke's account for possible worlds would again be the mathematical model of a sample space.

The constraint of *domain fixing* reduces the fill-in of substance to a pure combinatorics. It can be seen as the frame constraint of the *Barcan formula* $\forall x \Box \phi \rightarrow \Box \forall x \phi$, which can be deduced in quantificational **B**, but not in the weaker systems. It is obvious, that symmetry—the frame constraint of the formula **B**—imposes the ontological commitment of fixed domains. In a fixed domain universe, even the formula $\forall x \Box \exists y(x = y)$ is valid—on the objectual interpretation of quantification, everything exists necessarily, but it is questionable whether one should subscribe to this and not assume different objects in different possible worlds.

In all logics with non-fixed domains, if one has the necessitation rule, one can deduce the converse of the Barcan formula, namely the formula $\Box\forall x\phi \rightarrow \forall x\Box\phi$. But this imposes the frame constraint of monotonicity or *nested domains*—a domain of a possible world must be contained in all possible worlds accessible from it. Another problem of the approaches given so far is the treatment of non-existing objects, like 'Pegasus' or 'God'. This calls for *free logic* or at least *truth value gaps* [81].

4.4 Non-Rigid Designators

Some philosophers find rigid terms problematic for several reasons.

Kripke's theory of proper names deviates from the standard Frege-Russell theory of proper names as definite descriptions. For Kripke, proper names are *rigid* designators and definite descriptions are not [15].

With dynamic logic as a paradigm, Pascal allows rigid program constants only but not non-rigid program variables, whereas LISP needs non-rigid function symbols, as programs can redefine functions [64].

Non-rigid designators are necessary for Skolemization. If we Skolemize a formula $\Box\exists x\phi x$, we are not allowed to take one single Skolem-constant, say c, and insert it for x, because in different possible worlds, there might be different objects satisfying ϕ. So, if we want to introduce a convenient Skolem-constant, it must be non-rigid.

Among the philosophers devising non-rigid approaches Lewis and Platinga, who can both be counted among the *realists*, i.e. both of them believe, that possible worlds *exist* independent of our perception, are most influential.

Lewis tries to depict a way mediating between trans-world identity and essentialism à la Kripke. He proposes a *counterpart theory*, according to which individuals are *worldbound*, exist in only one world, but may have *counterparts* in other worlds. Counterparts are only similar to, but not identical with each others. Counterpart theory has some advantages over the rigid terms approach, but nevertheless it was also severely critisised by several authors [164, 123, 124, 139]. As Kripke points out, it cannot handle the Leibniz' Principle of the identity of indiscernibles in a correct way. Another argument goes as follows: Submit an object to slight changes in its properties, when passing from one world to another. In this series, every 'property list' may be similar to that of its neighbors, but after a sufficient number of steps, no one would ascribe similarity to the first and the last list of these series [162]. Further discussion of this position is included in the texts mentioned above and [135]. Another facet of Lewis's theory is the way he handles non-existing object. He is a *possibilist* assuming a subdomain of non-existent *possibilia*, so that under the objectual interpretation, there exist things which do not exist. This formula must not be formalized as $\exists x\neg\exists y(x = y)$—which is absurd, Lewis introduces additional

[15] Nevertheless, Kripke clearly offers a consistent alternative to the Frege-Russell doctrine.

quantifiers restricted to existing things (this is equivalent to the introduction of sorts, of course) [135]. Another way to handle this kind of problems is the introduction of *free logics*.

Plantinga, on the other hand, is an *actualist*, rejecting quantification over non-existent objects. Instead, he uses the old platonic idea of differentiation between existence and instantiation to distinguish the actual world from the possible others [163]. For Plantinga, a possible world is a *possible state of affairs, [but] not just any possible state of affairs is a possible world.* Only *complete* or *maximal* states of affairs are, this concept being similar to *maximal consistency* in logics. Possible worlds are constructed like maximal consistent sets or Hintikka sets in logics. Identity must not be traced through possible worlds, it exists *per constructionem* and is completely trivial. On the other hand, Platinga's approach lacks the intuitivity of Lewis's or Kripke's account.

There are still some severe problems with non-rigid designators, namely the problem of *syntax ambiguity*, the problem of *de re* and *de dicto modalities* and the problem of *quantifying in* (again).

In fact, these problems are one single problem, namely the problem of *quantifying in de dicto contexts*—a scope problem.

Fitting [63, 64] develops a formalism for non-rigid designators in quantified modal logics, including a new symbol, the *predicate abstraction* λ. He defines

$$(\lambda x.\phi x)c := \exists x(\phi x \wedge x = c)$$

From the right side of the definition, it is straight forward to derive the semantics of the predicate abstraction operator:

$$
\begin{aligned}
&\mathcal{M}, w_0 \models \exists x(\phi x \wedge x = c)[v] \\
&\qquad \Leftrightarrow \quad \text{ex.}\kappa \in Dom(w_0).\mathcal{M}, w_0 \models (\phi x \wedge x = c)[v[{}^{\kappa}_{x}]] \\
&\qquad \Leftrightarrow \quad \text{ex.}\kappa \in Dom(w_0)(\mathcal{M}, w_0 \models \phi x[v[{}^{\kappa}_{x}]] \text{ and} \\
&\qquad\qquad \mathcal{M}, w_0 \models (x = c)[v[{}^{\kappa}_{x}]]) \\
&\qquad \Leftrightarrow \quad \text{ex.}\kappa \in Dom(w_0)(\mathcal{M}, w_0 \models \phi x[v[{}^{\kappa}_{x}]] \text{ and} \\
&\qquad\qquad v[{}^{\kappa}_{x}](x, w_0) = v[{}^{\kappa}_{x}](c, w_0) = v(c, w_0)) \\
&\qquad \Leftrightarrow \quad \mathcal{M}, w_0 \models \phi x[v[{}^{v(c,w_0)}_{x}]]
\end{aligned}
$$

Now Fitting defines *de re* and *de dicto* formulas as having the following form:

$$
\begin{array}{ll}
(\lambda x.\Box\phi)c & \textit{de re}, \\
\Box(\lambda x.\phi)c & \textit{de dicto}.
\end{array}
$$

The reason for Fitting's introduction of predicate abstractors are the *scope* problems connected with non-rigid designators. Consider for instance the formula $\Diamond\phi c$, where c is a non-rigid designator. On a *de re* reading, it would mean, that one picks out the denotation of c at 'our' possible world and looks for an accessible possible world, where this same object has the property that in this world corresponds to ϕ. On a *de dicto* reading, we look for a world accessible from our world such that the object denoted by c in the accessible world has the property that in this world corresponds to ϕ. Of course, with rigid designators, the *de re* and *de dicto* case coincide, because c denotes the same object in all possible worlds.

To give an example let 'w_0' denote our actual world, let in our world c denote Napoleon and $\phi =$ the concept 'won the battle of Waterloo'. So ϕc is false in w_0. On a *de re* reading, $\Diamond\phi c$ is interpreted as 'Napoleon (from our world) won the battle of Waterloo in some accessible possible world.' and on a *de dicto* reading 'Some Napoleon (who- or whatever this may represent) from some accessible possible world won the battle of Waterloo in that world'. Now assume the accessibility relation to be irreflexive and further that in all possible worlds different from our world, 'Napoleon' denotes a certain brand of wine. Let in some possible world *our* Napoleon win the battle of Waterloo under the name of Wellington. Then the *de re* reading becomes true and the *de dicto* reading false. In a rigid designator model, our Napoleon is the same in all possible worlds, such that the *de re* reading turns out true if and only if the *de dicto* reading does.

But is this really—as Fitting believes—a matter of ambiguous syntax? At least a semantical analysis of the formula $\Diamond\phi c$ leads directly to the *de dicto* interpretation and it seems, as if the ambiguity was merely due to insufficient care of translation between natural and formal language. The *de re* and *de dicto* readings should be explicitly formalized as:

$$\begin{array}{lll} (\lambda x.\Box\phi x)c: & \exists x(\Box\phi x \wedge x = c) & \textit{de re}, \\ \Box(\lambda x.\phi x)c: & \Box\exists x(\phi x \wedge x = c) \quad \text{or} \quad \Box\phi c & \textit{de dicto}. \end{array}$$

Up to now, there is no ambiguity. But in fact, there is a problem with quantification, namely the problem of quantifying in modal contexts. According to the usual quantificational axioms, it is allowed to infer $\exists x\phi x$ from ϕc by *modus ponens*. This clearly leads to the paradox with the number nine and the number of planets, because this can be done for all formulas, but in order to avoid the paradox, one simply has to restrict this quantification rule by taking care of the modal operators: quantifying in modal contexts has to be forbidden. Fitting [64] gives a similar argument, involving the morning star/evening star paradox. This example is a bit obscure in the context of equality, as Fitting exposes it, it is really again the quantifying-in problem. Unfortunately, his argument is even further obscured by the fact that he talks about identity of objects and writes *[...] equality [*in fact, Fitting means identity*] should be thought of [...] as a relation on objects, not on names for them.* How can two (sic!) objects be identical! Frege and Wittgenstein would spin in their graves.

Fitting's argument is interesting for another reason. In modal calculus with equality, one can deduce the formula

$$(a = b) \rightarrow \Box(a = b).$$

Now, on our reading, this is the *de dicto* formula

$$(a = b) \rightarrow \Box(\exists x, y(x = y \wedge x = a \wedge y = b).$$

On a non-rigid interpretation, one would reject this formula, because it simply is not valid. It is not the case that the object denoted by a is identical with the object denoted with b in all possible worlds. Instead of this, one would accept the *de dicto* formula

$$(a = b) \rightarrow \exists x, y(\Box(x = y) \wedge x = a \wedge y = b),$$

the interpretation of which is trivial. Now one has to look for axioms, which take these facts into account [64]...

4.5 Further Alternatives—The Call for Free Logics

Garson [81] shows, that the realism approach together with free logic instead of classical logic, in order to handle non-existing objects, makes a change of the pre-modal axioms or rules necessary, but this increases the difficulty of completeness proofs. For some variants, completeness proofs even do not seem to exist. This all makes the non-rigid approach *not especially satisfying* [81]. In the *conceptual interpretation*, one tries to avoid these inconveniences by quantifying also over individual concepts. This modification introduces an essential weakness to the system: *no (consistent) system is complete for this [the usual] semantics* [81]. The problem is of course, that here one enters the domain of second order logic, where completeness can only be obtained with respect to *general (Henkin) models.* Finally, there is the *substantial interpretation*, which *restrict[s] the domain of quantification to the term intensions that reflect 'the way things are' across possible worlds* [81]. For a discussion, see the same reference. The important thing about the substantial interpretation is the fact that a set of substances must be specified for every possible world within the frame. The most general modal logics with a substantial semantics even cannot be axiomatized [28, 81]. But their first-order fragments, according to Garson, show a way of how to proceed in the future.

A general formal problem with quantificational modal logics is to prove completeness. For some of the systems, completeness proofs do not exist, and in general, there is no canonical way of doing it. This problem is also discussed at length in Garson [81]. For a general introduction to the philosophical problems connected with modal logics, read also Cocchiarella's Handbook article [44].

Epistemic interpretations introduce some new variants of our problems to quantificational theory [132], but in general, there is a straight transfer from the alethic problems. Lenzen, in order to avoid some of them, considers only rigid models with the philosophical theory of Kripke in the background. But this is still difficult enough—Lenzen is not in a position to give completeness proofs for all of his systems, especially not for the mixed-modality systems he proposes.

As already stated, there are quantificational systems for deontic modal logics [54, 34], but I will leave the discovery and investigation of those systems to the reader. The system of van Eck requires temporal logics, which lies beyond the scope of the present paper and Castañeda's system can, I think, be appreciated only by those familiar with his general philosophical ideas.

5 Conclusion

In this paper, I tried to sketch some philosophical logics, to provide a philosophical background for some current AI research, as a search through philosophical and computer science literature revealed that computer scientists are often not aware of the philosophical problems connected with the logics they consider. Unless general philosophy of science, where philosophy first of all deals with foundational or general methodical questions with almost no impact to the scientist's every day's work, in the case of philosophical logics, philosophical problems are not marginal, but the reason for the many insufficiencies with alethic, epistemic and deontic logics. The principal problem behind that is the fact that no one really knows how to handle intensionality in general. One can easily verify this argument by considering AI systems of belief. Usually, propositional **S5** or **S4** is used, but one hardly finds a discussion of this choice going beyond statements of the form: '**S4** means positive introspection, **S5** negative introspection.' Instead, knowledge and belief are difficult stratified concepts and really worth a thorough investigation. On the other hand, the philosophical enterprise towards general intensional logics, connected with names like Anderson, Church, Montague, Gallin, Kaplan and Castañeda, in order to mention only a few, is far from complete.

If one divides research on modal logics into three main objectives, namely research into formal mathematical properties, into structural philosophical properties and into the construction of cognitive adequate models of epistemic or deontic agents, only the first group of scientists might neglect philosophical problems, AI researchers should not. Cognitive adequacy requires acquaintance with cognitive theories, with psychology, linguistics and philosophy, but the proper problems connected with modal expressions are genuinely philosophical.

My sketch of the relevant features of philosophy of mind and language was neither intended to be deep nor exhausting, it should simply make computer scientists sensitive and allow for an orientation in the philosophical literature. Therefore I gave some introductory remarks on Frege semantics and some resulting features of intensionality. Then the connection to modal metaphysics and modal logics was pointed out and a discussion of the various alethic modal systems was given. It was shown, to what extent a simple parallelism between alethic, epistemic and deontic logics raises problems and what methods are employed to solve them. Epistemic modalities strictly behave non-intensionally and therefore some debatable assumptions must be introduced to assimilate epistemic logics to the alethic ones. Deontic logics are even more debated. Many philosophers doubt whether they should be given a semantics at all. To all these points, detailed references are given, such that the reader should in principle be able to undertake a closer examination.

One major problem with modal logics is the underlying difficulty of quantification into intensional contexts. I presented a method proposed by Fitting for fixing the scope of variables and explicitly depicting *de re* and *de dicto* formulas, which attempts to be able to handle this problem. Another central problem of quantificational modal logics, the problem of transworld identity was discussed.

Various ontological and semantical approaches to this problem, including the debate of rigid and non-rigid designators and Kripke's and Kaplan's deviations from Fregean semantics were sketched.

Besides these considerations within possible worlds metaphysics and semantics, I also mentioned possible-world-free alternatives, like algebraic semantics and general approaches to intensionality like those of Church and Bealer. Unfortunately, I could not present convincing solutions for most of the problems under consideration, but only point out where future research might be important—for example in free logics.

As a main conclusion, I want to note that the field under investigation is in flux on both sides, the philosophical and the artificial intelligence one. One should appreciate the AI approaches from the phenomenological side, because they brought a new dynamics to the field and one should never wait for general solutions, but nevertheless it is worth taking serious the philosophical questions, in favor of synergetic effects. Philosophical logics are maybe the only domain where contemporary philosophy has relevant ideas to contribute.

References

[1] E. Agazzi, editor. *Modern Logic—A Survey.* Reidel, 1981.

[2] P. Agre and D. Chapman. An implementation of a theory of activity. In *AAAI-87*, pages 268–272, 1987.

[3] A. Al-Hibri. *Deontic Logic: A Comprehensive Appraisal and a New Proposal.* University Press of America, 1978.

[4] L. Alanen and S. Knuuttila. The foundations of modality and conceivability in descartes and his predecessors. In S. Knuuttila, editor, *Modern Modalities.* Kluwer Academic Press, 1988.

[5] C.E. Alchourrón and E. Bulygin. *Normative Systems.* Springer-Verlag, 1971.

[6] J.F. Allen and C.R Perrault. Analyzing intention in utterances. *Artificial Intelligence*, 15:143–178, 1980.

[7] J. Almog et al., editors. *Themes from Kaplan.* Oxford University Press, 1989.

[8] A.R. Anderson and N.D. Belnap. *Entailment: The Logic of Relevance and Necessity*, volume 1. Princeton University Press, 1975.

[9] C.A. Anderson. General intensional logic. In D. Gabbay and Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 355–385. Reidel, 1984.

[10] L. Åqvist. Deontic logic. In D. Gabbay and Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 605–714. Reidel, 1984.

[11] L. Åqvist. *An Introduction to Deontic Logic and the Theory of Normative Systems.* Bibliopolis, 1985.

[12] N. Asher. Belief in discourse representation theory. *Journal of Philosophical Logic*, 15:127–189, 1986.

[13] N. Asher. Information, interpretation and attitudes. In P. Hanson, editor, *British Columbia Studies in Cognitive Science*, volume 1. University of British Columbia Press, 1990.

[14] J. Bacon. The logical form of perception statements. *Synthese*, 41:271–308, 1979.

[15] P. Bailhache. Sémantiques pour des systèmes déontiques intégrant permission faible et permission forte. *Logique et Analyse*, 20:286–316, 1977.

[16] P. Bailhache. Analytical deontic logic; authorities and addresses. *Logique et Analyse*, pages 65–80, 1981.

[17] P. Bailhache. *Essai de logique deontique.* Librairie Philosophique Vrin, Paris, 1991.

[18] J. Barwise and J.Perry. *Situations and Attitudes.* Bradford Books, MIT Press, 1983.

[19] G. Bealer. *Quality and Concept.* Clarendon Press, 1982.

[20] J. van Benthem. Syntactic aspects of modal incompleteness theorems. *Theoria*, 45(63–77), 1979.

[21] J. van Benthem. Correspondence theory. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 2: Extensions of classical Logic, pages 167–247. D. Reidel, 1984.

[22] J. van Benthem. Notes on modal definability. *Notre Dame Journal of Formal Logic*, 39:20–39, 1989.

[23] J. van Benthem. Beyond accessibility: Functional models for modal logic. In M. de Rijke, editor, *Diamonds and Defaults*, pages 175–214. Kluwer Academic Publishers, 1993.

[24] W. Bibel. First order reasoning about knowledge and belief. *Artificial Intelligence and Inf. -Control Syst. Robot.*, 3:9–16, 1984.

[25] U. Blau. *Glauben und Wissen. Eine Untersuchung zur Epistemischen Logik.* PhD thesis, Universität München, 1969.

[26] U. Blau. Zur Situation der deontischen Logik. *Papiere zur Linguistik*, 6:90–100, 1974.

[27] M.E. Bratman. *Intention, Plans and Practical Reason.* Harvard University Press, 1987.

[28] A. Bressan. *A General Interpreted Modal Calculus.* Yale University Press, 1973.

[29] C. Brink et al. Towards automating duality. Technical Report MPI-I-93-220, Max-Planck-Institut für Informatik, 1993.

[30] R. Bull and K. Segerberg. Basic modal logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 2: Extensions of classical Logic, pages 1–88. D. Reidel, 1984.

[31] R. Carnap. *Meaning and Necessity.* Chicago University Press, 1947.

[32] H.-N. Castañeda. Direct reference, the semantics of thinking, and guise theory (constructive reflexions on David Kaplan's theory of indexical reference). In J. Almog et al., editors, *Themes from Kaplan*, pages 105–144. Oxford University Press, 1989.

[33] H.N. Castañeda. *The Structure of Morality.* Charles C. Thomas, 1974.

[34] H.N. Castañeda. *Thinking and Doing: The Philosophical Foundations of Institutions.* Reidel, 1975.

[35] B.F. Chellas. *Modal Logic, an Introduction.* Cambridge University Press, 1980.

[36] R.M. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24:33–36, 1963.

[37] R.M. Chisholm. The logic of believing. *Pacific Philosophical Quarterly*, 61:31–49, 1980.

[38] R.M. Chisholm and R.G. Keim. A system of epistemic logic. *Ratio*, 14:99–115, 1972.

[39] T. Chrz. A system for knowledge representation based on intensional logic. *Comput. Artif. Int.*, 3:193–209, 305–317, 1984.

[40] A. Church. A formulation of the logic of sense and denotation. In P. Henle et al., editors, *Structure, Method and Meaning: Essays in Honour of Henry M. Sheffer*, pages 3–24. Liberal Arts Press, 1951.

[41] A. Church. Intensional isomorphism and identity of belief. *Philosophical Studies*, 5:65–73, 1951.

[42] A. Church. Outline of a revised formulation of the logic of sense and denotation (I). *Noûs*, 7:24–33, 1973.

[43] A. Church. Outline of a revised formulation of the logic of sense and denotation (II). *Noûs*, 8(135–156), 1974.

[44] N.B. Cocchiarella. Philosophical perspectives on quantification in tense and modal logic. In D. Gabbay and Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 309–353. Reidel, 1984.

[45] P.R. Cohen and H.J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 49:213–261, 1990.

[46] A.G. Conte et al., editors. *Deontische Logik und Semantik*. Athenaion, 1977.

[47] M.J. Cresswell. Intensional logics and logical truth. *Journal of Philosophical Logic*, 1:2–15, 1972.

[48] M.J. Cresswell. An incomplete decidable modal logic. *Journal of Symbolic Logic*, 49:520–527, 1984.

[49] M. de Rijke, editor. *Diamonds and Defaults*. Kluwer Academic Publishers, 1993.

[50] Farinas del Cerro. Un principe de résolution en logique modale. *RAIRO Inf. Theor.*, 18:161–170, 1984.

[51] D.C. Dennett. Recent work in philosophy of interest to AI. *Artificial Intelligence*, 19:3–5, 1982.

[52] D.C. Dennett. Recent work in philosophy II. *Artificial Intelligence*, 22:231–233, 1984.

[53] D.C. Dennett. *The Intantional Stance*. MIT Press, 1987.

[54] J.A. van Eck. A system of temporally relative modal and deontic predicate logic and its philosophical applications. *Logique et Analyse*, 25:249–290,339–381, 1982.

[55] D. Elgesem. *Action Theory and Modal Logic*. PhD thesis, Department of Philosophy, University of Oslo, 1992.

[56] R. Fagin and J. Y. Halpern. Belief, awarness and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

[57] K. Fine. An incomplete logic containing **s4**. *Theoria*, 40(23–29), 1974.

[58] K. Fine. Model theory for modal logic I, II. *Journal of Philosophical Logic*, 7:125–156, 277–306, 1978.

[59] K. Fine. The problem of de re modality. In J. Almog et al., editors, *Themes from Kaplan*, pages 197–272. Oxford University Press, 1989.

[60] M. Fisher. A logical theory of commanding. *Logique et Analyse*, 4:154–169, 1961.

[61] F. Fitch. Natural deductoin rules for obligation. *American Philosophical Quarterly*, 3:27–38, 1966.

[62] M. Fitting. *Proof Methods for Modal and Intuitionistic Logics*. Number 169 in Synthese Library. Reidel, 1983.

[63] M. Fitting. Modal logic should say more than it does. In J.-L. Lassez and G. Plotkin, editors, *Computational Logic, Essays in Honor of Alan Robinson*, pages 113–135. MIT Press, 1991.

[64] M. Fitting. Basic modal logic. In D. Gabbay et al., editors, *Handbook of Logic in Artificial Inelligence and Logic Programming*, volume 1 Logical Foundations, pages 368–448. Clarendon Press, 1993.

[65] D. Føllesdal and R. Hilpinen. Deontic logic: An introduction. In R. Hilpinen, editor, *Deontic Logic: Introductory and Systematic Reading*, pages 1–35. Reidel, 1971.

[66] G. Frege. Der Gedanke. In *Funktion, Begriff, Bedeutung*, pages 30–53. Vandenhoek & Ruprecht, 3rd edition, 1986.

[67] G. Frege. *Funktion, Begriff, Bedeutung.* Vandenhoek & Ruprecht, 3rd edition, 1986.

[68] G. Frege. Funktion und Begriff. In *Funktion, Begriff, Bedeutung*, pages 17–39. Vandenhoek & Ruprecht, 3rd edition, 1986.

[69] G. Frege. *Logische Untersuchungen.* Vandenhoek & Ruprecht, 1986.

[70] G. Frege. Über Begriff und Gegenstand. In *Funktion, Begriff, Bedeutung*, pages 66–80. Vandenhoek & Ruprecht, 3rd edition, 1986.

[71] G. Frege. Über Sinn und Bedeutung. In *Funktion, Begriff, Bedeutung*, pages 40–65. Vandenhoek & Ruprecht, 3rd edition, 1986.

[72] G. Gaba and N. di Nola. A doxastic logic. *Riv. Nature Pure e Appl.*, 31(1):19–33, 1981.

[73] D. Gabbay et al., editors. *Handbook of Logic in Artificial Inelligence and Logic Programming*, volume 1 Logical Foundations. Clarendon Press, 1993.

[74] D. Gabbay and H.J. Ohlbach. Quantifier elimination in second-order predicate logic. Technical Report MPI-I-92-231, Max-Planck-Institut für Informatik, 1992.

[75] D. Gallin. *Intensional and Higher-Order Modal Logic.* North-Holland, 1975.

[76] P. Gärdenfors. On the interpretation of deontic logic. *Logique et Analyse*, 21:371–398, 1978.

[77] P. Gärdenfors. Epistemic importance and the logic of theory change. Int. Congr. Log., Met. and Phil. Sci., Salzburg, 1985.

[78] P. Gärdenfors. *Knowledge in Flux: Modelling the Dynamics of Epistemic States.* MIT Press, 1988.

[79] P. Gärdenfors and N.-E. Sahlin. Unreliable probabilities, risk taking, and decision making. *Synthese*, 53:361–386, 1982.

[80] J.L. Gardies. Logique d'action (ou du changement) et logique déontique. *Logique et Analyse*, 26:71–89, 1983.

[81] J.W. Garson. Quantification in modal logic. In D. Gabbay and Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 249–307. Reidel, 1984.

[82] M.L. Ginsberg and D.E. Smith. Reasoning about action I, II. *Artificial Intelligence*, 35:165–195, 197–226, 1988.

[83] R.I. Goldblatt. A new extension of S4. *Notre Dame Journal of Formal Logic*, 14:567–574, 1973.

[84] R.I. Goldblatt. Metamathematics of modal logics. *Reports on Mathematical Logic*, 6,7:41–78,21–51, 1976.

[85] N. Grana. *Logica deontica paraconsistenta.* Liguori, Napoli, 1990.

[86] S. Haak. *Philosophy of Logics.* Cambridge University Press, 1978.

[87] A.R. Haas. A syntactic theory of belief and action. *Artificial Intelligence*, 28:245–292, 1986.

[88] J.Y. Halpern and R. Fagin. Two views on belief: belief as generalized probability and belief as evidence. *Artificial Intelligence*, 54:275–317, 1992.

[89] J.Y. Halpern and Y.O. Moses. A guide to the modal logics of knowledge and belief. In *Proceedings IJCAI-85*, pages 480–490, 1985.

[90] J.Y. Halpern and Y.O. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54:319–379, 1992.

[91] W.H. Hanson. A logic of commands. *Noôus*, 3:373–398, 1966.

[92] B. Hansson. An analysis of some deontic logics. In R. Hilpinen, editor, *Deontic Logics: Introductory and Systematic Reading*, pages 121–147. Reidel, 1971.

[93] A. Hauttamaeki. Points of view and their logical analysis. *Acta Philosophica Fennica*, 41, 1986.

[94] R. Hilpinen, editor. *Deontic Logic: Introductory and Systematic Readings.* Reidel, 1971.

[95] R. Hilpinen, editor. *New Studies in Deontic Logic.* Reidel, 1981.

[96] J. Hintikka. Quantifiers in deontic logic. Technical Report mentationes humanarum litterarum, 23:4, Societas Scientiarum Fennica, 1957.

[97] J. Hintikka. *Knowledge and Belief.* Cornell University Press, 1969.

[98] J. Hintikka, editor. *Models for Modalities: Selected Essays.* Synthese Historical Library, 23. Reidel, 1969.

[99] J. Hintikka. The modes of modality. In M.J. Loux, editor, *The Possible and the Actual, Readings in the Metaphysics of Modality*, pages 65–79. Cornell University Press, 1979.

[100] J. Hobbs and C.Moore R, editors. *Formal Theories of the Commonsense World.* Ablex, Norwood, NJ, 1985.

[101] W. van der Hoek. *Modalities for Reasoning about Knowledge and Quantities.* PhD thesis, Vrije Universiteit van Amsterdam, 1992.

[102] G.E. Hughes and M.J. Cresswell. *An Introduction to Modal Logic.* Methuen, 1968.

[103] G.E. Hughes and M.J. Cresswell. *A Companion to Modal Logic.* Methuen, 1984.

[104] R.E. Jennings and P.K. Schottch. Epistemic logic, skepticism and non-normal modal logic. *Philosophical Studies*, 40:47–67, 1981.

[105] A.J.I. Jones. On the logic of deontic conditionals. In *Milestone 2—Deliverables*, volume 1 of *Esprit 3125–Medlar.* Department of Computer Science, Imperial College, 1992.

[106] E. Jónsson and A. Tarski. Boolean algebras with operators I. *Am.J.Math*, 73:891–939, 1951.

[107] G. Kalinowski. *La Logique des Normes.* Presses Universitaires de France, 1972.

[108] G. Kalinowski and F. Selvaggi, editors. *Les Fondements Logiques de la Pensée Normative.* Acte du Colloque de Logique Déontique de Rome (1983. Ed. Pontificia Università Gregoriana, 1985.

[109] S. Kanger. *Provability in Logic.* PhD thesis, University of Stockholm, 1957.

[110] S. Kanger. New foundations for ethical theory. In R. Hilpinen, editor, *Deontic Logic: Introductory and Systematic readings*, pages 36–58. Reidel, 1971.

[111] I. Kant. *Kritik der reinen Vernunft.* Felix Meiner Verlag, 1956.

[112] D. Kaplan. Demonstratives. In J. Almog et al., editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, 1989.

[113] D. Kaplan. Transworld heir lines. In M.J. Loux, editor, *The Possible and the Actual, Readings in the Metaphysics of Modality*, pages 88–109. Cornell University Press, 1997.

[114] W. Kneale and M. Kneale. *The Development of Logic*. Clarendon Press, 1962.

[115] S. Knuuttila. The emergence of deontic logic in the fourteenth century. In R. Hilpinen, editor, *New Studies in Deontic Logic*, pages 225–248. D. Reidel, 1981.

[116] S. Knuuttila. *Modern Modalities*. Kluwer Academic Press, 1988.

[117] K. Konolige. *A Deduction Model for Belief*. Morgan Kaufmann Inc., 1986.

[118] L. Kott, editor. *Knowledge, Belief and Time*, Rennes, 1986. ICALP 13, Springer Verlag.

[119] M. Kracht. How completeness and correspondence theory got married. In M. de Rijke, editor, *Diamonds and Defaults*, pages 175–214. Kluwer Academic Publishers, 1993.

[120] L. Kreiser et al. *Nichtklassische Logik*. Akademie-Verlag, 1990.

[121] S. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14, 1959.

[122] S. Kripke. Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16:83–94, 1963.

[123] S. Kripke. Identity and necessity. In M. Munitz, editor, *Identitiy and Individuation*. New York University Press, 1971.

[124] S. Kripke. *Naming and Necessity*. Harvard University Press, 1980.

[125] F. von Kutschera. *Einführung in die intensionale Semantik*. de Gruyter, 1976.

[126] K. Lambert, editor. *Philosophical Problems in Logic*. D. Reidel, 1970.

[127] K.B. Laskey and Lehner P.E. Assumptions, beliefs and probabilities. *Artificial Intelligence*, 41:65–77, 1990.

[128] J.-L. Lassez and G. Plotkin. *Computational Logic, Essays in Honor of Alan Robinson*. MIT Press, 1991.

[129] J. Legris. *Eine epistemische Interpretation der intuitionistischen Logik*. Koenigshausen & Neumann, 1990.

[130] H. Lenk, editor. *Normenlogik*. Verlag Dokumentation, 1974.

[131] W. Lenzen. Recent work in epistemic logic. *Acta Philosophica Fennica*, 30, 1978.

[132] W. Lenzen. *Glauben, Wissen und Wahrscheinlichkeit*. Springer Verlag, 1980.

[133] W. Lenzen. Doxastic logic and the Burge-Buridan paradoxon. *Philosophical Studies*, 39:43–49, 1981.

[134] H. Levesque. A logic of implicit and explicit belief. In *AAAI-84*, 1984.

[135] D. Lewis. Counterpart theory and quantified modal logic. In M.J. Loux, editor, *The Possible and the Actual, Readings in the Metaphysics of Modality*, pages 110–128. Cornell University Press, 1979.

[136] L.M Lewis. *The Ontology, Syntax, and Computability of Deontic Logic*. PhD thesis, University of Georgia, 1986.

[137] F. Lin and Y. Shoam. A logic of knowledge and justified assumptions. *Artificial Intelligence*, 57:271–289, 1992.

[138] J. Łoś. Logiki wielowartosciowe a formaloszacja funkcij intensionalnych. *Kwartalnik Filozofisczny*, 17:59–78, 1948.

[139] M.J. Loux. Introduction: Modality and metaphysics. In M.J. Loux, editor, *The Possible and the Actual, Readings in the Metaphysics of Modality*, pages 15–64. Cornell University Press, 1979.

[140] M.J. Loux, editor. *The Possible and the Actual, Readings in the Metaphysics of Modality*. Cornell University Press, 1979.

[141] E. Mally. *Grundgesetze des Sollens. Elemente der Logik des Willens*. Leuschner & Lubensky, 1926.

[142] A.A. Martino. *Deontic Logic, Computational Linguistics and Legal Information Systems*, volume 2. North Holland, 1981.

[143] J.P. Martins and S.C. Shapiro. A model for belief revision. *Artificial Intelligence*, 35:25–79, 1988.

[144] J.-J.C. Meyer and R.J.Wieringa, editors. *Proc. of 1st Int. Workshop on Deontic Logic in Computer Science*, Free University of Amsterdam, 1991.

[145] J.-J.Ch. Meyer et al. Epistemic logic for computer scientists: A tutorial. *EATCS bulletin*, 44:242–270, 1991. (I).

[146] J.-J.Ch. Meyer et al. Epistemic logic for computer scientists: A tutorial. *EATCS bulletin*, 45:256–287, 1991. (II).

[147] P. Milgrom. An axiomatic characterisation of common knowledge. *Ecnometrica*, 49(1):219–222, 1981.

[148] C.A. Mortari. *Issues in Epistemic and Modal Logics and Their Applications*. PhD thesis, Universität Tübingen, 1991.

[149] G.H. Müller and W. Lenski. *Omega-Bibliography of Mathematical Logic*, volume 2, Non-Classical Logics. Springer-Verlag, 1978.

[150] R.M. Neal. Connectionist learning of belief networks. *Artificial Intelligence*, 56:71–113, 1992.

[151] I. Niemelä. Towards automated autoepistemic reasoning. In J. van Eijk, editor, *Logics in AI, European Workshop JELIA'90*, pages 428–443. Springer-Verlag, 1991.

[152] I. Niiniluoto and E.Saarinen, editors. *Intensional Logic: Theory and Applications.* Acta Philodophica Fennica, vol 35, 1982.

[153] U. Nortmann. Deontische Logik: Die Methode der lokalen Äquivalenz. *Erkenntnis*, 25:275–318, 1986.

[154] D. Nute. Permission. *Journal of Philosophical Logic*, 14:169–190, 1985.

[155] H.J. Ohlbach. *A Resolution Calculus for Modal Logics.* PhD thesis, Universität Kaiserslautern, 1988.

[156] H.J. Ohlbach. Translation methods for non-classical logics—an overview. Technical Report MPI-I-93-225, Max-Planck-Institut für Informatik, 1993.

[157] E. Orlowska. Logic of vague concepts. *Bull. Sect. Log., Pol. Akad. Sci.*, 11:115–129, 1982.

[158] E. Orlowska and Z. Pawlak. Logical foundations of knowledge representation. *Prace, Inst. Podstazo Inf. Pol. Akad. Nauk.*, 551:108 pp, 1984.

[159] A. Pap. Belief and proposition. *Philosophy of Science*, 24:123–136, 1957.

[160] D. Perlis. Language with self-reference II: knowledge, belief and modality. *Artificial Intelligence*, 34:179–212, 1988.

[161] S. Peters and E. Saarinen, editors. *Processes, Beliefs and Questions.* D. Reidel, 1982.

[162] A. Plantinga. World and essence. *Philosophical Review*, 79:461–492, 1970.

[163] A. Plantinga. Actualism and possible worlds. In M.J. Loux, editor, *The Possible and the Actual, Readings in the Metaphysics of Modality*, pages 253–273. Cornell University Press, 1979.

[164] A. Plantinga. Transworld identity or worldbound individuals? In M.J. Loux, editor, *The Possible and the Actual, Readings in the Metaphysics of Modality*, pages 146–165. Cornell University Press, 1979.

[165] D. Prawitz. *Natural Deduction: A Proof-Theoretic Study*, volume 3 of *Stockholm Studies in Philosophy.* Almqvist and Wiksell, 1965.

[166] W.v.O. Quine. *From a Logical Point of View.* Harper Torchbooks, 1953.

[167] W.v.O. Quine. Two dogmas of empiricism. In *From a Logical Point of View.* Harper Torchbooks, 1953.

[168] W.v.O. Quine. *Word and Object.* Wiley, 1960.

[169] V. Rantala. Impossible worlds semantics and logical omniscience. In I. Niiniluoto and E.Saarinen, editors, *Intensional Logic: Theory and Applications.* Acta Philodophica Fennica, vol 35, 1982.

[170] V. Rantala. Quantified modal logic: Non-normal worlds and propositional attitudes. *Studia Logica*, 41:41–65, 1982.

[171] H. Rasiowa and R. Sikorsky. *The Mathematics of Metamathematics.* Państwowe Wydawnictwo Naukowe, 1963.

[172] N. Rescher. An axiom system for deontic logic. *Philosophical Studies*, 9:24–30, 1958.

[173] N. Rescher. *Hypothetical Reasoning.* North-Holland, 1964.

[174] N. Rescher. *The Logic of Commands.* Routledge & Kegan Paul, 1966.

[175] N. Rescher, editor. *The Logic of Decision and Action.* Pittsburgh University Press, 1967.

[176] N. Rescher. *Topics in Philosophical Logic.* Reidel, 1968.

[177] N. Rescher. *Studies in Modality.* Blackwell, 1974.

[178] N. Rescher. *Empirical Inquiry.* The Athlone Press, 1982.

[179] W. Roos. A logic for reasoning with inconsistent knowledge. *Artificial Intelligence*, 57:69–103, 1992.

[180] S.J. Rosenschein. Formal theories of knowledge in AI and robotics. *New Generation Computing*, 3(4), 1985.

[181] K. Sadegh-Zadeh. Recension of: 'W. Lenzen, Glauben, Wissen und Wahrscheinlichkeit'. *Erkenntnis*, 23:97–112, 1985.

[182] G.N. Schlesinger. *The Range of Epistemic Logic*, volume 6 of *Scots Philosophical Monograph Series.* Aberdeen University Press, 1985.

[183] J.R. Searle. *Intentionality: An Essay in the Philosophy of Mind.* Cambridge University Press, 1983.

[184] K. Segerberg. Action-games. *Acta Philosophica Fennica*, 32:220–231, 1981.

[185] K. Segerberg. A deontic logic of action. *Studia Logica*, 41:269–282, 1982.

[186] K. Segerberg. The logic of deliberate action. *Journal of Philosophical Logic*, 11:233–254, 1982.

[187] K. Segerberg. A topological logic of action. *Studia Logica*, 43:415–420, 1984.

[188] D. Siemens Jr. Fitch-style rules for many modal logics. *Notre Dame Journal of Formal Logic*, 18:631–636, 1977.

[189] M.P. Singh and M. Asher. Towards a formal theory of intensions. In J. van Eijk, editor, *Logics in AI, European Workshop JELIA'90*, pages 472–486. Springer-Verlag, 1991.

[190] M.P. Singh and M. Asher. A logic of intentions and beliefs. *Journal of Philosophical Logic*, 22:513–544, 1993.

[191] T.J. Smiley. The logical basis of ethics. *Acta Philosophica Fennica*, 16:237–246, 1963.

[192] T.J. Smiley. Relative necessity. *Journal of Symbolic Logic*, 28:113–134, 1963.

[193] R. Stalnaker and R. Thomason. Abstraction in first-order modal logic. *Theoria*, 34:203–207, 1968.

[194] R.C. Stalnaker. *Inquiry.* MIT Press, 1984.

[195] W.R. Stark. A logic of knowledge. *Zeitschrift für Logik und Grundlagen der Mathematik*, 27:371–374, 1981.

[196] W. Stelzner. *Epistemische Logik.* Akademie–Verlag, 1984.

[197] R. Thomason and R. Stalnaker. Modality and reference. *Nôus*, 2:203–207, 1968.

[198] S.K. Thomasson. Reduction of second order logic to modal logic. *Zeitschrift für Logik und Grundlagen der Mathematik*, 21:107–114, 1975.

[199] J. van Eijk, editor. *Logics in AI, European Workshop JELIA'90,*. Springer-Verlag, 1991.

[200] F. Voorbraak. Generalized Kripke models for epistemic logic. Logic Group: Logic Group preprint series 71, Rijksuniversiteit Utrecht, 1991.

[201] F. Voorbraak. The logic of objective knowledge and rational belief. In *Logics in AI, European Workshop JELIA'90*, pages 499–515. Springer-Verlag, 1991.

[202] L.A. Wallen. *Automated Deduction in Nonclassical Logics.* MIT Press, 1990.

[203] O. Weinberger. *Studien zur Normenlogik und Rechtsinformatik.* J. Schweitzer Verlag, 1974.

[204] G.H. von Wright. Deontic logic. *Mind*, 60:1–15, 1951.

[205] G.H. von Wright. *An Essay in Modal Logic.* North-Holland, 1951.

[206] G.H. von Wright. A note on deontic logic and derived obligation. *Mind*, 65:507–509, 1956.

[207] G.H. von Wright. *Norm and Action.* Routledge & Kegan Paul, 1963.

[208] G.H. von Wright. A new system of deontic logic. In R. Hilpinen, editor, *Deontic Logic*, pages 105–114. North–Holland, 1964.

[209] G.H. von Wright. An essay in deontic logic and the general theory of action. *Acta Philosophica Fennica*, 21, 1968.

[210] G.H. von Wright. *Handlung, Norm und Intention: Untersuchungen zur Deontischen Logik.* de Gruyter, 1977.

[211] G.H. von Wright. Problems and prospects of deontic logic: A survey. In E. Agazzi, editor, *Modern Logic—A Survey*, pages 299–423. Reidel, 1981.

[212] J.J. Zeman. *Modal Logic: The Lewis-Modal Systems.* Clarendon Press, 1973.