Part IX

MODAL LOGICS AND SEMANTICS

26

Modal Logic

JOHAN VAN BENTHEM

1 Enriching Extensional Logic with Intensional Notions

When Frege wrote *Begriffsschrift*, he intentionally left out the key intensional notions of traditional logic before him. On one telling page he enumerates a list of things for which he sees no need – and readers of some erudition will recognize this anonymous enemy as Kant's famous "Table of Categories", including Modality. Nevertheless, in this century modal notions made their way back onto the logical agenda, leading to extensions of classical systems with operators of necessity, possibility, entailment, and other metaphysically inspired notions. These formalisms were influential as a tool for analyzing philosophical arguments. I still recall the shudder when reading my first sequences of symbols claiming to be a proof of God's existence 'out of a box.' But also, the semantics of modal logics in terms of possible worlds has formed a powerful philosophical union with the ontologies of Kripke and Lewis. These motivations also provided a watershed from mathematical logic, whose practitioners disliked modal logic instinctively, even though they are willing to countenance such deviations as intuitionistic or quantum logic. But worse than that, by the early 1980s, modal logic had also acquired powerful enemies within philosophy, preaching its imminent demise. I remember sneaking through corridors in those days, avoiding encounters with energetic colleagues who might be tempted to lend a helping hand to Historical Necessity. But modal logic did not die, its enemies never managed to invent an equally powerful substitute, its content and uses rather multiplied, and Handbooks wisely still include the subject.

2 Changing Views of Modal Logic

In what follows, I present a modern account of modal logic – not as a metaphysical system of any sort, but as a logical 'fine-structure formalism' for talking about actions, knowledge, and many other concrete things all around us. This view is very different from the original motivation in 'philosophical logic,' and I do not claim that it is uncontroversial in that field, especially among the *ancien regime*. But it is about time that a broader community learns what is really going on.

We have come a long way since the 1960s, because of two separate developments. First, what happened is a familiar phenomenon in science: originally non-intended applications of a theory take over. In the case of modal logic, these started with temporal and epistemic logic, then we had spatial logics, dynamic logics of action, and by now also modal logics of grammatical derivation, generalized quantifiers, games, or concept descriptions in AI. And this expansion is going on all the time. These applications provide new impetus to modal logic, at a time when it seems fair to say that ontology is no longer a live source of inspiration. A second influence came from inside modal logic. The mathematical theory of the subject that began to flourish in the 1970s yielded (as abstract mathematics should) surprising new viewpoints on what makes modal languages tick, which generated different perspectives – and in the end, a startling *inversion*. Viewed in one way, modal logics are typically extensions of classical logic with new operators. Viewed in another, and perhaps ultimately more insightful way, modal logics are *fragments* of classical logical languages, that serve as milestones in a natural 'fine-structure hierarchy' of expressiveness and reasoning.

Out of this historical panorama, we choose three notions as our major themes, *viz*. *fine-structure*, *information*, and *dynamics*. These will be introduced by looking at the basic modal language: propositional logic with box □ and diamond ◊. But first, let us mention another characteristic of much modal research: its exotic landscapes of different logics, such as K, T, S_4, S_5, or more bizarre code names. This seems a huge difference with a monotheistic religion like classical logic, which has only one set of validities – and hence many people associate modal logic with heathen botany. Now, the mathematical theory of the 1970s did create more unity in what has been called the 'jungle of modal logics.' Powerful meta-theorems appeared establishing properties like decidability, interpolation, frame-correspondence or completeness for whole families of 'modal logics' at once, or locating systematic failures – using methods from universal algebra and model theory. Nevertheless, and more controversially than our stance so far, we think this diversity is not a *fundamental* characteristic of the modal way of life – even though it is certainly one of its useful conveniences. Such 'logics' are different modal theories of special types of accessibility relation, comparable to special theories formulated on top of classical predicate logic. Our exposition will therefore concentrate on *modal base languages* and their properties, with an occasional excursion into the special frame classes of this other dimension of research.

3 A Précis of Basic Modal Logic

Language and interpretation

The basic modal language is very simple, and yet it has been a useful laboratory for new basic techniques. We interpret formulas in so-called possible worlds models – the grand name is still popular for its nostalgic mood – **M** = (W, R, V), according to the well-known truth definition:

$\mathbf{M}, s \models \Diamond A$ iff for some t with Rst, $\mathbf{M}, t \models A$
$\mathbf{M}, s \models \Box A$ iff for all t with Rst, $\mathbf{M}, t \models A$

It helps to think of the worlds as 'states' of some kind, while accessibility encodes possible moves that can be made to get from one state to another. But there are many other useful concrete views of these, in essence, 'decorated graphs' (figure 26.1).

Invariance for Bisimulation

The expressive power of this language is measured by a suitable notion of similarity between different models.

DEFINITION A *bisimulation* between two models **M**, **N** is a binary relation E between their states m, n s.t. whenever m E n, then (a) m, n satisfy the same proposition letters, (b1) if m R m′ , then there exists a world n′ with n R n′ and m′ E n′, (b2) the same 'zigzag clause' holds in the opposite direction.

Together, this 'atomic harmony' and the two zigzag clauses make bisimulation a natural notion of 'process equivalence' – and indeed it was independently discovered in computer science. Example (disregarding proposition letters): the two black worlds in **M**, **N** are connected by the drawn bisimulation, consisting of all the matches indicated by dotted lines – but there is no bisimulation which includes a match between the black worlds in **N** and **K** (figure 26.2).

INVARIANCE LEMMA If E is a bisimulation between **M** and **N**, and m E n , then m, n satisfy the same modal formulas.

That is, modal formulas are *invariant for bisimulation*. Thus, we can see the above failure of bisimulation by noting that the model in the middle satisfies the formula

$$\Diamond\Diamond\Box\bot$$

in its root, whereas the one to the right does not. The converse to the Lemma only holds for a modal language with *arbitrary infinite* conjunctions and disjunction – or for the plain modal language over special models. For instance:

PROPOSITION If m, n satisfy the same modal formulas in *finite* models **M**, **N**, then there is a bisimulation E between these with m E n.

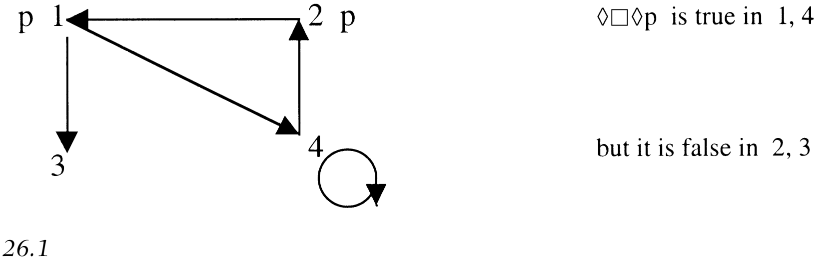

Figure 26.1

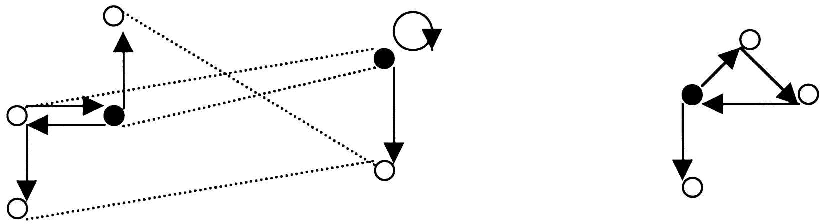

Figure 26.2

But there are still stronger definability results. For example, for any model **M**, s with a designated world s, there is an infinitary modal formula $\phi^{\mathbf{M},s}$ true in only those models **N**, t which are 'bisimilar' to **M**, s (i.e. some bisimulation links t to s).

Validity and proof systems

Universal validity is axiomatized in Hilbert-style by the so-called *minimal modal logic*:

1. all laws of propositional logic
2. a definition of $\Diamond\phi$ as $\neg\Box\neg\phi$
3. the modal distribution axiom $\Box(\phi \to \Psi) \to (\Box\phi \to \Box\Psi)$
4. the necessitation rule 'if $\vdash\phi$, then $\vdash\Box\phi$'

This looks like a standard axiomatization of first-order logic (with $\Box$ as $\forall$, and $\Diamond$ as $\exists$), but leaving out axioms with tricky side conditions on freedom and bondage: $\forall x\phi \to [t/x]\ \phi$ and $\phi \to \forall x\phi$. Modal deduction, either axiomatic or in other proof formats (sequents, natural deduction), is simple reasoning in perspicuous notation.

Modal logic games

Not intrinsic to modal logic, but a pleasant dynamic trend is this. All our notions have fine-structure as *games*. In an *evaluation game*, players Verifier (V) and Falsifier (F) disagree about a formula. Disjunction is a choice for V, conjunction for F, negation is role switch, $\Diamond$ makes V pick a successor of the current world, $\Box$ does the same for F. A game p is won by Verifier if the atom p holds in the current state, otherwise by Falsifier. A player also wins the game if the other player must move, for a modality, but cannot.

FACT **M**, s $\models \phi$ iff Verifier has a *winning strategy*
for the ϕ-game in **M** starting from s.

For example, our first model picture induces the following game tree for formula $\Diamond\Box\Diamond p$ starting from state 1, with bold-face indicating the winning positions for Verifier (figure 26.3).

In this game, V has two winning strategies: 'left,' and 'right,' ⟨'right,' 'down'⟩. These are indeed the two possible successful ways of verifying formula $\Diamond\Box\Diamond p$ in the given

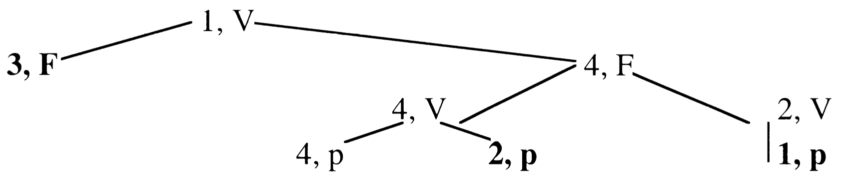

Figure 26.3

model **M** at world 1. Likewise, there are *model comparison games* between Duplicator (maintaining an analogy) and Spoiler (claiming a difference), playing over pairs (m, n) in two models **M**, **N**. These provide a fine-structured way of checking for the earlier bisimulation. In each round Spoiler chooses a state x in one model which is a successor of m or n, Duplicator responds with a corresponding successor y in the other model. If x, y are different in their atomic properties, Spoiler wins – if Duplicator cannot find a matching successor: likewise.

For example, in the non-bisimulation example **N**, **K** in figure 26.2, starting from a match between the two black worlds, Spoiler needs 3 rounds to win: forcing Duplicator in 2 rounds into a match where one world has no successor, while the other does.

FACT (a) Spoiler's winning strategies in a k-round game between **M**, s, **N**, t match the modal formulas of *operator depth* k on which s, t disagree. (b) Duplicator's winning strategies over an *infinite* round game between **M**, s , **N**, t match the bisimulations between them linking s to t.

One winning strategy for Spoiler in the preceding example exploits the earlier 'difference formula' $\Diamond\Diamond\Box\bot$. Many other logical notions can be 'gamified'. In particular, there are *construction games* determining if a given formula has a model, or *proof games* finding a derivation of it through a dialogue between two players.

Decidability and complexity

The basic modal language is a *decidable* 'miniature' of first-order logic. There are many decision methods for validity or satisfiability, exploiting special features of modal formulas – each with their virtues in generalization. Well-known methods are 'selection,' 'filtration,' and 'reduction.'

The deeper underlying issue is the precise computational *complexity* of various key tasks for a logic. These include not just satisfiability or validity testing, but also model checking and model comparison. Here are some landmark observations.

MODEL CHECKING Given a finite model **M**, s and a modal formula ϕ, checking if **M**, s $\models \phi$ takes polynomial time in length(ϕ) + size(**M**) .

This is better than for first-order logic, where the same task takes polynomial space.

SATISFIABILITY Checking if a given modal formula has a model takes *polynomial* space in the size of the given formula.

For propositional logic the same task takes just non-deterministic polynomial time. For the full first-order language, of course, it is undecidable.

MODEL COMPARISON Checking if there is a bisimulation between given finite **M**, s, **N**, t takes polynomial time in the size of these models.

This may look surprising, but simple algorithms exist throwing out successive pairs of worlds that cannot make any bisimulation. These benchmark complexity outcomes may differ as modal languages are varied, allowing us to detect 'jumps.' Complexity awareness is a new feature of increasing importance in logic.

Model theory

The model theory of basic modal logic is much like that of first-order logic: with classical highlights such as Craig interpolation, Los–Tarski preservation theorem for universal modal formulas, etc. The analogy gets lost for many special modal logics, where for example interpolation is much scarcer.

Translation

Modal operators behave much like first-order quantifiers. The following translation T takes all modal formulas ϕ to first-order formulas $T(\phi)$ with one free variable x having the same truth conditions on models **M**, s:

(a) $T(p) = Px$,
(b) T commutes with all the Boolean operators,
(c) $T(\Diamond\phi) = \exists y\,(Rxy \,\&\, [y/x]T(\phi))$, $T(\Box\phi) = \forall y\,(Rxy \rightarrow [y/x]T(\phi))$

With some care, only two variables x, y are needed in all these first-order translations (free or bound). E.g. $\Box\Diamond\Box p$ translates faithfully into the formula

$$\forall y\,(Rxy \rightarrow \exists x(Ryx \,\&\, \forall y(Rxy \rightarrow Py))).$$

Here is the essential semantic feature that makes these translated modal formulas special inside the full first-order language over R^2, P^1, Q^1, . . .

MODAL INVARIANCE THEOREM The following assertions are equivalent for all first-order formulas $\phi = \phi\,(x)$: (a) ϕ is equivalent to a translated modal formula, (b) ϕ is invariant for bisimulations.

The 'modal fragment' is a small fragment of FOL, sharing its 'nice' properties, but remaining decidable. What you get for free on this view are 'universal' properties of first-order formulas, such as the Löwenheim–Skolem Theorem. Not for free is, for example, the Interpolation Theorem: modal consequences might have non-modal first-order interpolants: honest work is required to show that indeed *modal* interpolants exist. The fragmentist perspective is general: many other modal languages live inside first-order logic or other standard logics, under some translation transcribing their standard semantics. We will see later what makes these fragments so well-behaved.

Landscapism

On top of the minimal logic, there are uncountably many different 'modal logics.' This landscape has two major highways: because of this

THEOREM Every normal modal logic is either a subset of the logic *Id* (with characteristic axiom $\phi \leftrightarrow \Box\phi$) or of *Un* (axiom $\Box\bot$).

On the former road lie the usual systems like T, S_4, S_5, on the latter, for example, 'Gödel–Löb logic' of arithmetical provability axiomatized by $\Box(\Box\phi \rightarrow \phi) \rightarrow \Box\phi$. Modal logics in this landscape can be studied by proof-theoretic or semantic methods, with a flourishing industry of completeness theorems providing bridges between the two.

Completeness

A typical modal completeness theorem runs like this.

THEOREM A formula is provable in K4 (K plus the axiom $\Box\phi \rightarrow \Box\Box\phi$) iff it is true in all models whose accessibility relation is *transitive*.

There are many techniques for proving such results, ranging from simple inspection of the Henkin model of all complete theories in the logic to drastic 'model surgery.' The motivation for completeness theorems can come from two directions. Either one has a pre-existing logic given by axioms and rules (such as the above cases), and seeks a useful corresponding model class – or one has a natural model class (say, some interesting space–time structure), and wishes to axiomatize its modal validities for 'simple reasoning.' The literature is replete with both. In this survey, we do not pursue this completeness line, since it gets so much exposure anyway.

Correspondence

The preceding correspondence between modal axioms and properties of the accessibility relation is a major attraction of modal logic. It can also be studied directly in the semantics, calling a modal formula true in a *frame* (a model stripped of its valuation) if it holds under all valuations. This line of research has produced two key results of a model-theoretic nature:

THEOREM A modal formula defines a first-order frame-property iff it is preserved under taking 'ultrapowers' of frames.

THEOREM A first-order frame-property is modally definable iff it is preserved under taking (a) 'generated subframes,' (b) 'p-morphic frame images,' (c) 'disjoint unions,' and (d) 'inverse ultrafilter extensions.'

Known non-first-order modal principles are the McKinsey Axiom $\Box\Diamond p \to \Diamond\Box p$, and the earlier-mentioned Gödel–Löb Axiom. Useful in practice is the *Sahlqvist Theorem*, describing an effective method for constructing first-order equivalents for most widely used modal axioms, which has by now reached the world of automated theorem proving. It proceeds by substituting first-order descriptions of 'minimal valuations' into a modal axiom to get a natural first-order equivalent (if available).

EXAMPLE The above K4 axiom $\Box p \to \Box\Box p$ has a first-order translation $\forall y\,(Rxy \to Py) \to \forall y\,(Rxy \to \forall z\,(Ryz \to Pz))$. A minimal valuation for p making the antecedent true is $Pu := Rxu$. Substituting this into our formula, and dropping the then tautologically true antecedent, we are left with a consequent of the syntactic form $\forall y\,(Rxy \to \forall z\,(Ryz \to Rxz)))$, which is precisely frame *transitivity*.

In a sense this whole mathematical theory is a study of simple modal fragments of the complex realm of *second-order logic*, a perspective we will not pursue here.

The basic modal language has limited expressive power. But it has been the main mathematical laboratory for notions, techniques, and results. In what follows here, we look at some modern extensions, and new basic issues to which these give rise.

4 The Major Applications

Contemporary 'applications' of modal logic are not routine uses of existing notions and techniques: they add things not dreamed of in the original framework. This short article cannot really do justice to the variety of developments here. Here are some major directions that are arguably most influential in the 'drift' of the field.

Epistemic logic

Propositional attitudes like knowledge show logical behavior like that of ontological modalities. In particular, the epistemic operator

$K_i\phi$ 'agent **i** knows that ϕ' is like a universal modality stating that ϕ is true *in all of i's epistemically indistinguishable situations*.

And the same is true to some extent for other epistemic propositional attitudes, such as 'belief.' On this view, accessibilities are often equivalence relations for each agent – though alternatives exist. Languages like this express basic epistemic statement patterns that we often use in natural discourse, such as

$K_i\phi \vee K_i\neg\phi$ 'agent i knows whether ϕ is the case'

and modal axioms acquire a new flavor:

$K_i\phi \to K_i K_i\phi$ 'positive introspection'
$\neg K_i\phi \to K_i K_i\neg\phi$ 'negative introspection'

But the major new theme in this epistemic setting is a ‘social one.’ It is not the Lonely Thinker that is essential to understanding cognition, but *interaction* between different agents in a *group:* $K_iK_j\phi$ or $K_i\neg K_j\phi$. What I know about your knowledge or ignorance is crucial, both to my understanding and to my actions. (For example, I might empty your safe tonight if I think you don’t know that I know the combination.) Some forms of ‘group knowledge’ even transcend simple iterations of individual knowledge assertions. The central example here is *common knowledge*: if everyone knows that your partner is unfaithful, but no more, you have private embarrassment – if it is common knowledge, you have public shame and perhaps a need for violent action. Technically, common knowledge works as follows:

> $C_G\phi$ ϕ holds at every world reachable via any finite chain of uncertainty relations for actors in G.

For example, in the picture in figure 26.4, where p holds in the current world (the black dot), in the black world, (a) agent Q does not know if p is the case: $\neg K_Q p$ & $\neg K_Q\neg p$; (b) agent A does know that p is the case: $K_A p$; while (c) it is common knowledge in the group {Q, A} that A knows *whether* p is the case: $C_{\{Q, A\}}$ $(K_A p \vee K_A\neg p)$. Incidentally, this might be a good situation for Q to ask A a *question* about p: but more on epistemic *actions* below.

Dynamic logic

Accessibilities can also be viewed as *transitions* for actions that change states. In ‘dynamic logic’ – originally designed to describe the execution of computer programs, but now used as a general logic of action, we have

> $[\pi]\phi$ says that *after every successful execution of action* π*,* ϕ *holds.*

Thus, modal statements relate actions to ‘postconditions’ describing their effects (and also to ‘preconditions’ for their successful execution). A concrete model of this are *games*, where actions are moves available to players. For example in the tree shown in figure 26.5, player E has a *strategy* for achieving an outcome satisfying p.

This strategic assertion is captured by the dynamic modal formula $[a\cup b]\langle c\cup d\rangle p$. Again we get a minimal modal logic for universal validity here, this time set up as a *two-level system* treating propositions and actions (transition relations) on a par. This joint set-up allows for a logical analysis of important action constructions, encoded in valid principles for general operations as well as specific actions:

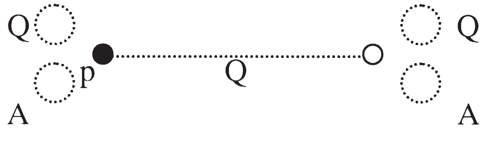

Figure 26.4

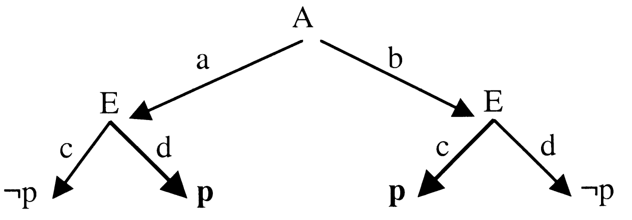

Figure 26.5

$[\pi;\pi']\phi \leftrightarrow [\pi][\pi']\phi$	sequential composition
$[\pi\cup\pi']\phi \leftrightarrow [\pi]\phi \;\&\; [\pi']\phi$	choice
$[(\phi)?]\psi \leftrightarrow (\phi\rightarrow\psi)$	test for proposition ϕ

A major new feature here is unbounded finite *repetition* of actions: π^*. This is typical for computation, and it is not first-order definable. This shows in axioms

$[\pi^*]\phi \leftrightarrow (\phi \;\&\; [\pi][\pi^*]\phi)$	fixed-point axiom
$(\phi \;\&\; [\pi^*](\phi \rightarrow [\pi]\phi)) \rightarrow [\pi^*]\phi$	induction axiom

Thus, dynamic logics resemble *infinitary* fixed-point extensions of classical logic, but they do retain the 'modal stamp': being bisimulation-invariant, and decidable. Fixed-point definitions are ubiquitous in computer science, but also in mathematics or linguistics, because many natural notions involve a kind of 'implicit' recursion. An elegant current system of this kind for actions is a generalization of dynamic logic allowing arbitrary fixed-point definitions: the so-called 'μ-calculus.'

Temporal and spatial logic

A more traditional, but very lively application area of modal logic concerns 'physical' rather than 'human' nature. We mention this as a counterpoint to our cognitive slant. One concrete interpretation of modal models is as *flows of time*, accessibility being 'earlier than.' The universal modality will then say 'everywhere in the future,' which comes with an obvious dual 'everywhere in the past.' Temporal logics are prominent in computer science and AI, where they show a great diversity beyond this basic modal point of departure. In particular, they can live over different primitive entities: duration-less points, or extended 'periods.' Usually, the vocabulary of temporal languages is much richer than the basic modal language. A typical example are operators allowing us a view of what goes on *during* the successful execution of a program or plan:

UNTIL $\phi\psi$ *at some point later than now* ϕ *holds,*
while at all intermediate points ψ *is true*

In this same physical arena, modal logics of *space* are also gaining importance, for example in knowledge representation. One of these revives an old mathematical idea. Let our models be topological spaces endowed with a valuation. Then the modality

$\Box\phi$ may be read as saying that the current point lies in the *topological interior* of the set $[[\phi]]$ of all points where ϕ holds.

Then, modal laws come to encode various topological facts about space, for example:

$\Box(\phi \& \psi) \leftrightarrow \Box\phi \ \& \ \Box\psi$ says that open sets are closed under intersections.

This style of analysis may be extended to modal fragments of geometry. It provides an alternative to our standard semantics quantifying over successors in some binary world-to-world relation. (Technically, it is a 'neighborhood semantics,' of a sort developed in the 1960s to explore landscapes below the minimal modal logic K.) Thus our spatial excursion also shows that the 'standard approach' is not sacrosanct.

AI, linguistics, mathematics

Modal logic has either been applied, or rediscovered, in such areas as artificial intelligence ('description languages,' 'context logics'), linguistics ('categorial grammar,' 'feature logics'), and indeed mathematics, with flourishing areas such as 'provability logic,' and in recent years also modal versions of set theory. This list is not complete (intuitionistic logic or relevant logic or linear logic are also similar in some of their key features), but it does show that modal structures occur naturally across a wide range of disciplines.

5 Fine-Structure of Expressive Power

Modal logic today shows several new general themes that cut across these various applications. We mention a few, though there is certainly no consensus on a simple synthesis out of the current research scene. One is *extension of expressive power*.

Logical extensions

Modal languages can be enriched over their original models. A popular 'logical extension' of this sort adds a *universal modality*

$U\phi$ saying that *ϕ is true at all worlds*, accessible or not.

This gives more expressive power, which one can use to state 'global facts,' such as the inclusion of one region of the model in another. But our standard techniques generalize, for example, the language of $\{\Box, U\}$ matches up with 'total bisimulations,' whose domains and ranges are the whole models being compared. And also: its minimal logic remains decidable – though the complexity of validity goes up to exponential time. (When added to more complex languages, indeed, U may push a decidable logic over the brink into undecidability.) In earlier years, extending the basic modal language was 'not done,' because it would change the rules of the game, and make life too easy. Here is another example. Having *names for specific worlds* would be a great convenience, both in practice and in the modal metatheory, but the basic language does not allow it. For example, much has been made of the latter's inability to express the frame property of *irreflexivity* ($\forall x\ \neg Rxx$). But this is expressed quite simply by the following axiom in an extended modal language:

$i \rightarrow \neg \Diamond i$ where the 'nominal' i is a special proposition letter ranging over only *singleton sets* of worlds.

Nowadays the tendency is to add such devices freely, only subject to striking a good balance between increased expressive power and manageable complexity. Another example is the above operator 'Until' of temporal logic, where inhibitions as to enrichment have always been weaker. What keeps these extensions 'modal' is that they allow for bisimulation analysis, while staying decidable. Much is known by now about which added operator leads to which jump in decidable complexity for our benchmark tasks of satisfiability, model checking, and model comparison.

'Geometrical' extensions

By contrast to the preceding move, 'geometric extensions' enrich the similarity type of our models, adding modalities with *new accessibilities*, as in epistemic or dynamic logic, or in *polyadic* modal languages with n-ary alternative relations. For example an existential 'dyadic modality'

$\Diamond\phi\psi$ holds at s iff $\exists t, u$ *s.t.* $R^3 s, tu$, ϕ *holds at t*, ψ *holds at u*

Concrete interpretations for such ternary accessibility relations R include:

s is the concatenation of two expressions t, u,
s is the merge of the two resources t, u.

Guarded fragment

One limit to which many extensions of both types tend is the so-called *Guarded Fragment* of first-order logic. This is defined inside the full first-order syntax by allowing only quantifications of the 'guarded' form

$\exists \mathbf{y}(G(\mathbf{x}, \mathbf{y}) \ \&\ \phi(\mathbf{x}, \mathbf{y}))$

where **x**, **y** are tuples of variables, G(**x**, **y**) is an atomic formula whose variables occur in any order and multiplicity, and ϕ is a guarded formula having only variables from **x**, **y** free. Many modalities are guarded in this syntactic sense:

$\Diamond p$ $\exists y(Rxy \ \&\ Py)$
$\Diamond pq$ $\exists yz(Rxyz \ \&\ Py \ \&\ Qz)$

This sublanguage of first-order logic, where groups of objects are only introduced 'under guards' still yields to modal analysis supporting a 'nice' meta-theory.

THEOREM The Guarded Fragment has a characteristic bisimulation.
THEOREM The Guarded Fragment is decidable in doubly exponential time.

These properties even transfer to certain extensions. Another interesting property exemplified in this setting is *robust decidability*: small modal languages sometimes bear the weight of expressive extensions that otherwise explode reasoning complexity. An example are fixed-point operators for inductive definitions. On top of first-order logic, these make the resulting language non-axiomatizable – when added to the Guarded Fragment, however, they do not increase complexity at all.

Two dimensions

The earlier 'landscape' of modal logic was really one-dimensional: it kept the basic language constant in expressive power, varying deductive strength of special theories expressed in it. But now we have a second dimension: systematic variation of expressive power. This new two-dimensional landscape has many 'thresholds of complexity' which are currently being charted.

6 System Combination: Action and Information

Other main themes in general modal logic today are *many agents*, *dynamics*, and *system combination*. The former has already occurred in our survey. As to the latter, many applications are 'multi-modal,' putting together various modal logics in one system: say of action, knowledge, and time. There are several ways of doing this, ranging from mere 'juxtaposition' to more intricate forms of interaction between the component logics. One then wants to predict expressiveness and complexity of the combination from those of its parts – plus the mode of combination used. There is an incipient general theory of relevant modes of combination, including new constructions of 'product' and 'fibering.' This style of thinking even shows in modern technical views of *modal predicate logic*. One can 'deconstruct' this famous system into a combination of two modal logics: a static one of world accessibility, and a dynamic one of object assignment to variables. The main challenge arising then are the unpredictable effects of various combinations. Disregarding further generalities concerning composition of logics, we describe two rather exciting recent special combinations of long-standing modal ideas.

Information update

Models of epistemic logic serve as information states for groups of agents. Epistemic formulas are then evaluated against worlds in such states, telling us what is true or not in them. But knowledge usually functions in *communication*: it is conveyed to others via speech acts, and influenced by theirs. To model such *cognitive actions*, we need to combine two earlier systems: epistemic logic and dynamic logic. In particular, a communicative action changes the current epistemic model! In the simplest case, this 'update' works as follows:

public announcement of a proposition ϕ *to a group of agents eliminates all worlds in the current model* ***M*** *that satisfy* ϕ

Suppose that in our earlier two-agent two-world picture Q asks A : "p?" and A then truthfully answers 'Yes.' Then the ¬p-world gets eliminated, and we are left with a one-world model where p has become common knowledge among {Q, A}. But more subtle cases are possible, even with very simple models of this sort. For example, a question itself may convey crucial information! Suppose that, by asking, Q conveys the information that she does not know the status of p. Even if A did not know the answer at the start, this may tell him enough to settle p, and thereby answer the question. Figure 26.6 shows one scenario where this happens.

But the modeling power of combined epistemic dynamics is still higher. For example, suppose neither Q nor A knew about p, but A asks expert R, who answers only to Q. Then A learns whether p, Q is no wiser about p, but it has become common knowledge that A knows whether p. This requires 'arrow elimination' (figure 26.7).

These simple pictures hide delightful subtleties. For example, one may check that, on this account, a public announcement that some formula ϕ is the case need not always result in our 'learning that ϕ', in the form of an updated model where ϕ holds! (For, truth value switches may happen when we process an announcement of *ignorance*.) Precise algorithms for performing updates associated with communicative acts, public or private, have been proposed in recent years – and these provide an entirely new use of our 'standard models.' Eventually, in this view of communication, one wants to describe not just information update, but also actions of 'withdrawal' or *revision*, triggered by propositions that contradict the content of our current information state. These cognitive actions require modal logics of *counterfactual conditionals* that feed into modern *belief revision theory*.

Game logics

There can be more than one attractive way of putting modal ideas together. Another interesting mix of 'epistemics' and 'dynamics' occurs in the analysis of *games*. As players move through a game tree, their information changes. Plain game trees are described in dynamic logic, as we saw in an earlier section, though realistic reasoning

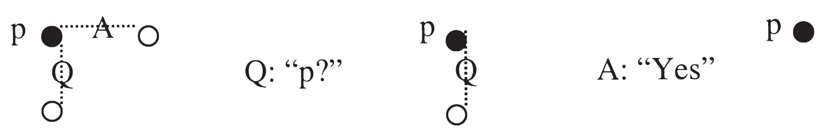

Figure 26.6

Figure 26.7

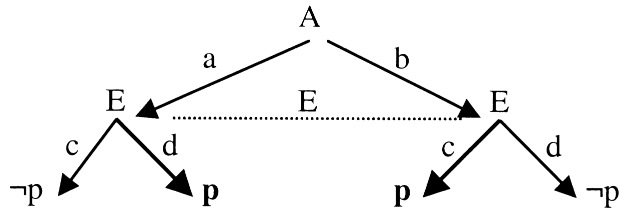

Figure 26.8

about future game actions also require a logic of players' *preferences*. Especially interesting, however, are *imperfect* information games, where players may not know the precise moves played by their opponents. Thus, in these games, the primary epistemic uncertainty is *between actions*, and only in a derived sense between the resulting game states. (Think of a card game where we cannot observe which initial hand Nature is dealing to our opponent, or where some mid-play moves by our opponents may be partially hidden.) An informative example is the earlier game tree, but now with an uncertainty link for player E at the second stage – she does not know the precise opening move played by A (figure 26.8).

We can view this as a model for an obvious combined dynamic-epistemic language, having both epistemic modalities K_i and dynamic ones [a], which may interact. In particular, half-way, player E knows 'de dicto' that she has a winning move

$$K_E(\langle c\rangle p \vee \langle d\rangle p)$$

but she does not know any particular winning move 'de re':

$$\neg K_E\langle c\rangle p \;\&\; \neg\, K_E\langle d\rangle p!$$

Indeed, this game is 'non-determined' in a natural sense: E cannot force an outcome p, but neither can A force outcome ¬p. The general logic of these game trees is the minimal propositional dynamic logic plus epistemic 'multi-S5.' But on top of that, the combined dynamic-epistemic language can also express *modes of playing games*. Take the game-theoretic notion of 'Perfect Recall.' This describes players whose *own* actions never introduce any uncertainties that they did not have before. Properly understood, this validates an interchange axiom

$$(turn_E \;\&\; K_E[a]\phi) \rightarrow [a]K_E\phi:$$

what we know about the result of our own game moves is still known to us after we perform them. (To understand this better, contrast the effects of non-'epistemically neutral' actions like drinking genever.) Thus, we can correlate modal logics in this epistemic-dynamic language with special styles of playing a game. Another mode is 'Bounded Memory' – whose treatment requires a universal modality. This simple example also illustrates a general point. Games are a nice target for logical analysis because they show cognition at work under well-defined 'laboratory circumstances.'

7 Back to the Heartland

Modal logic started as an epicycle on standard logic. And it is still viewed by most people as a 'nonstandard' topic beyond The Core. But latterly, it has started to influence the heartland itself. We conclude with two examples of this 1990s trend.

Modal foundations of predicate logic

Predicate logic *itself* is a form of modal or dynamic logic! The key truth condition for the existential quantifier reads

$\mathbf{M}, s \models \exists x\phi$ iff there exists d in D^M s.t. $\mathbf{M}, s[x{:=}d] \models \phi$

This has the modal pattern for evaluating an existential modality $\langle x\rangle$:

$\mathbf{M}, s \models \exists x\phi$ iff there exists t s.t. $R^x st$ with $\mathbf{M}, t \models \phi$

Viewed in this light, the usual set of 'valid laws' of first-order logic can be deconstructed into several layers: (1) Its decidable(!) core is the minimal modal logic, which contains such laws as Monotonicity: $\forall x(\phi \to \psi) \to (\forall x\phi \to \forall x\psi)$. This level makes no presuppositions whatsoever concerning the form of the models, which could have any kind of 'states' and 'variable shifts' R^x. (2) Next, there are laws recording universal effects of taking variable assignments for states, plus the special shift relation of 'agreeing up to the value for x.' For example $\forall x\phi \to \forall x\, \forall x\phi$ expresses the *transitivity* of R^x: indeed, all of S_5 holds. (3) Most 'specifically', some first-order laws express *existence* properties for states. Here is an example:

$\exists x\, \forall y\phi \to \forall y\, \exists x\phi$ expresses *confluence*: whenever s R^x t and s R^y u, then there also exists a state v s.t. t R^y v and u R^x v (figure 26.9).

Thus, modal analysis reveals unexpected 'fine-structure' in the class of what is usually lumped together as 'standard validities': they are valid for different reasons!

Moreover, on our general modal models, the predicate-logical language gets increased expressive power, because new distinctions come up. For example:

polyadic quantifiers $\exists xy\bullet$

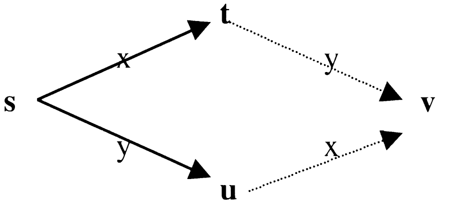

Figure 26.9

introducing two objects becomes different from iterations ∃x ∃y• or ∃y ∃x•.
Summing up, we get a highly unorthodox view. The 'modal core' of standard logic is decidable, *pace* Church and Tarski – but piling up special (existential) model conditions makes state sets behave so much like *full function spaces* D^{VAR} that their total logic becomes undecidable.

Dynamic predicate logic

Another dynamic view on first-order logic rather emphasizes the *state change* implicit in evaluating an existential quantifier. We move to a new state containing a suitable 'witness value' for x. More generally, one can let first-order formulas denote *actions of evaluation*:

(a) atomic formulas are *tests* if the current state satisfies the relevant fact,
(b) an existential quantifier picks an object and assigns it to x (*random assignment*),
(c) a substitution operator [t/x] is a *definite assignment* x:=t,
(d) a conjunction is sequential action *composition*,
(e) a negation ¬ϕ is a test for the *impossibility* of succesfully executing the action ϕ.

The resulting 'dynamified' version of first-order logic has applications in the semantics of natural language – as anaphoric pronouns 'he,' 'she,' 'it,' show this kind of dynamic behavior. One nice illustration occurs with sentences like

$\exists x\, Kx \rightarrow Hx$ 'if you get a kick, it hurts'

The standard logical folklore must 'improve' natural language here to arrive at the universal first-order form $\forall x\, (Kx \rightarrow Hx)$. But with dynamic semantics, this meaning arises automatically, as any value assigned by the existential move in the antecedent will be bound to x when the consequent is processed. This system has also inspired programming languages for dynamic execution of specifications.

'Dynamic predicate logic' exemplifies a general paradigm of bringing out the implicit cognitive dynamics which underlies existing logical systems. This allows one to view natural language meanings in terms of updates of propositional content, perspective, and other parameters that determine the transfer of information.

8 Conclusion

This survey is different in spirit from standard wisdom in philosophical logic. We have presented modal logic as a tool for fine-structure analysis of the expressiveness and complexity of logical languages, including effects of their combinations, and the major applications (information, action) that drive abstract theory today. There is no uniform conclusion, or even a new definition of modal logic in the end: the modern field is just too rich for that. Our purpose with this short article will have been served if the reader experiences a culture-shock, seeing the differences between reality and the picture still painted by many 'standard textbooks.'

References

Technical surveys

Bull, R. and Segerberg, K. (1984) Basic modal logic. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic* (pp. 1–88). Dordrecht: Reidel.

The chapters on modal logic by M. Fitting ('Modal logic') and C. Stirling ('Modal and temporal logics') in S. Abramsky, D. Gabbay and T. Maibaum (eds.) (1991) *Handbook of Logic in Computer Science*. Oxford: Oxford University Press.

Benthem, J. van (1991) *Manual of Intensional Logic*. Stanford: CSLI Publications.

Chagrov, A. and Zakharyashev, M. (1997) *Modal Logic*. Oxford: Clarendon Press.

Consecutive volumes in the recent series *Advances in Modal Logic*. Dordrecht: Kluwer.

Modern textbooks

Popkorn, S. (1992) *First Steps in Modal Logic.* Cambridge: Cambridge University Press.

Blackburn, P., de Rijke, M. and Venema, Y. (2000) *Modal Logic*. Cambridge: Cambridge University Press.

Modal predicate logic

M. Fitting's chapter in this volume.

Temporal logic

Benthem, J. van (1996) Temporal logic. In D. Gabbay, C. Hoggar and J. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming* (vol. 4, pp. 241–350). Oxford: Oxford University Press.

Gabbay, D., Hodkinson, I. and Reynolds, M. (1994) *Temporal Logic*. Oxford: Oxford University Press.

Epistemic logic

Fagin, R., Halpern, J., Moses, Y. and Vardi, M. (1995) *Reasoning about Knowledge*. Boston, MA: MIT Press.

Dynamic logic

Goldblatt, R.(1986) *Logics of Time and Computation.* Stanford, CA: CSLI Publications.

Harel, D., Kozen, D. and Tiuryn, J. (2000) *Dynamic Logic.* Boston, MA: MIT Press.

Provability logic

Boolos, G. (1993) *The Logic of Provability.* Cambridge: Cambridge University Press.

Modal set theory

Barwise, J. and Moss, L. (1999) *Vicious Circles.* Stanford, CA: CSLI Publications.

Philosophical foundations

Zalta, E. (1993) A philosophical conception of propositional modal logic. *Philosophical Topics*, 21, 263–81.

Modal grammar

Blackburn, P. and Meyer-Viol, W. (1997) Modal logic and model theoretic syntax. In M. de Rijke (ed.), *Advances in Intensional Logic* (pp. 29–60). Dordrecht: Kluwer.

Chapters on feature logic by W. Rounds ('Feature logics') and on categorial grammar by M. Moortgat ('Categorical type logics') in J. van Benthem and ter Meulen, A. (eds.) (1997) *Handbook of Logic and Language.* Amsterdam: Elsevier.

Modal logic and AI

See the Electronic Discussion Newsletter, *Reasoning about Action and Change*: Linköping, http://www.ida.liu.se/ext/etai/rac/notes/.

Modal logic and game theory

Benthem, J. van (1999) *Logic and Games.* Lecture notes. Amsterdam and Stanford.

Logical dynamics

Benthem, J. van (1996) *Exploring Logical Dynamics*. Stanford, CA: CSLI Publications.

Gerbrandy, J. and Groeneveld, W. (1997) Reasoning about information change. In J. van Benthem and Y. Shoham (eds.), *Cognitive Actions in Focus, Journal of Logic, Language and Information*, vol. 6 (pp. 147–69).

Baltag, A. (1999) *A Logic of Communication.* Amsterdam: Centre for Mathematics and Computer Science.

Ditmarsch, H. van (2000) Knowledge games. Dissertation. Groningen: Department of Computer Science and Institute for Logic, Language and Computation, Amsterdam.