THE RANGE OF MODAL LOGIC

an essay in memory of George Gargov

Johan van Benthem

ILLC, University of Amsterdam

september 1997

Abstract

George Gargov was an active pioneer in the 'Sofia School' of modal logicians. Starting in the 1970s, he and his colleagues expanded the scope of the subject by introducing new modal expressive power, of various innovative kinds. The aim of this paper is to show some general patterns behind such extensions, and review some very general results that we know by now, 20 years later. We concentrate on simulation invariance, decidability, and correspondence. What seems clear is that 'modal logic' as a genre of logical systems has a much wider scope than originally conceived, and that we have not reached its limits yet.

to appear in D. Vakarelov, ed.,
Memorial Issue for George Gargov,
"Journal of Applied Non-Classical Logics"

1 Personal Reminiscences

My first vivid recollection of an encounter with George Gargov is that of a bus ride, somewhere in the Balkan country-side. His quiet voice managed to penetrate the usual shake-rattle-and-roll of Bulgarian public transportation, as he told me about the founding phases of the history of his country, which was then still wavering between the Latin and Greek varieties of Christianity, to the extent of sending repeated embassies to Rome. This was my first visit to Bulgaria, and I was fast revising my romantic expectations (based on the Balkan novels by my childhood author Karl May) of martial mountaineers carrying flint-lock muskets, followed at a distance by wives carrying the family capital as their jewelry. Evidently, solid traditional values were in decline in this part of Europe, just as in my western part. Balkan adventures, when they arose, were of a communist brand, like sudden power failures during winter lectures at the university. Teaching logic with passion became less of a communicative act, and more a way of keeping warm.

The reason for this visit, and subsequent ones, was my interest in the 'Sofia School' in modal logic. Through the 80s, George and his colleagues at Kliment Ohridski University were systematically expanding the scope of traditional modal logic, adding expressive power to its language in various ways. This move may seem natural now, but it was radical at the time. The initial paradigm of Lewis and Kripke had been so succesful that it seemed sacrilege to re-define it. I was intrigued by the new line, and this paper testifies to my later work in this spirit. The visit also showed that the modal logicians in Sophia were attracting gifted students, drawing upon the considerable intellectual resources in Bulgaria. George Gargov was a leader in this special research program. Toward the 90s, he also started pursuing a broader vision of 'logic, language, and information', creating a special laboratory for linguistic and logical modelling at the Bulgarian Academy of Sciences which brought together logicians, linguists, and computer scientists. Many colleagues from the West have visited this lively and hospitable centre of activity.

I could easily spend the pages of this paper on reminiscences. Lectures, conversations, meals – all with a mixture of academic pleasures and the wisdom of George's gentle personality. A favourite memory is a hike in the Balkan mountains with George and friends, which somehow captured the oscillating experiences of research. Up plateau after plateau, ever-changing views of the true lie of the country, a brief stay at the peak, then clambering down in an increasing fog, presaging new challenges for the day ahead. George was proud of the various beauties of his country – and confident that they would shine again, once Bulgarian talents had the political chance of reasserting themselves. He has died too early, at the hesitant beginning of this renaissance.

2 Extended Modal Logic

2.1 Basic Modal Logic

Basic modal logic has the usual language of propositional logic, extended with two operators $\Box$ (necessity) and $\Diamond$ (possibility). Modal models $\mathbf{M} = (W, R, V)$ consist of a set of worlds W, an accessibility relation R, and a valuation V for the proposition letters in each world $w \in W$. The truth conditions for the modalities read as follows:

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

Thus, modalities are like quantifiers over successors of the current world. (Given the obvious interdefinability via negation, we will often use the existential modality only.) The logical properties of this basic system are well-known. It behaves like first-order logic in many ways, including compactness, Löwenheim-Skolem, preservation theorems for model constructions, interpolation, as well as effective axiomatizability. But a crucial difference is its *decidability*, as opposed to the known undecidability of first-order logic. (Cf. Andréka, Németi & van Benthem 1996 for a recent survey of these properties, and a meta-analysis of what makes them tick.) Thus, basic modal logic has a good balance between expressiveness and complexity, which underlies its varied applications.

Over this basic system, the usual landscape of deductively stronger 'modal logics' arises. An extensive theory has emerged by now of richer logics for special frame classes, defined by various constraints on the accessibility relation. The latter are then reflected systematically in additional modal validities of the system (as brought out in the 'correspondence theory' of van Benthem 1984, 1985). Modal Logic in this wider sense by now has developed sophisticated general techniques for proving axiomatizability, definability, and decidability. (Cf. Goldblatt 1987, Wolter & Zakharyashev 1996, Chagrov & Zakharyaschev 1997 and the textbook Blackburn, de Rijke & Venema 1996.) The main thrust of investigations in the 'classical' phase of modal logic is as follows. The basic language does not change, and neither does the type of models, or the above truth definition. One rather makes special assumptions, restricting attention to particular model classes. This move is like the study of specialized *axiomatic theories*' in first-order logic. For instance, the modal logic S4 is not a competitor of the minimal logic K, but rather the special-purpose theory of quasi-orders. By contrast, what we are after in this paper is something else, namely changes in the underlying logical framework itself. Imposing further specialized frame constraints is an orthogonal issue, which can be raised whenever we have found a new base logic.

2.2 Extended Modal Logic: the Sofia Style

Instead of varying frame conditions with respect to the same modal language, boosting *deductive power*, one can also increase the *expressive power* of the modal language, trying to retaining the pleasant feature of low complexity for the new base system. This 'logical' strengthening allows us to say more about the same models. It is in the latter sense that the Sofia logicians have made their most original contributions. This move can be made in a number of ways. First, one can leave the similarity type unchanged, and merely increase expresssive power. Examples are additional universal modalities, 'Boolean Modal logic', or temporal logic, all to be discussed below. (This may be compared with boosting the expressive power of first-order logic, over the same extra-logical vocabulary, by adding, say, infinitary connectives or generalized quantifiers.) But other extensions involve a change in the *similarity type* of the original models, endowing them with more accessibility structure. Examples of this second route include 'polymodal logic', which increases the operator repertoire to include indexed modalities [i], <i> over families of relations R_i (with the obvious truth condition), or 'modal logic with names' for individual worlds, which rather increases the atomic repertoire. Also under this heading, we have 'polyadic modal logic', which employs n–ary modalities over (n+1)–ary accessibility relations (with truth conditions to be stated below). We will also discuss this second type of logical extension. But our list is still not exhaustive. There are further types of extension to the modal base. These include such different strategies as 'many-dimensional modal logic', which changes the format of interpretation to evaluate at tuples of worlds, rather than single ones – or 'neighbourhood semantics', which replaces the original models with first-order relations Rxy by new models with second-order relations RxY between worlds and sets of worlds, reinterpreting the same language. The latter have been less prominent in the Sofia School (though not entirely absent). We will need to refer to them occasionally, to get the proper general picture.

Sofia-style extensions did not arise through aimless variation. They reflect applied interests of Gargov and his colleagues. For instance, strengthened modal operators in the same similarity type arise naturally in temporal logic, which is interested in the betweenness structure of accessibility. Likewise, polymodal logic is the natural language for program logics reflecting the accessibilities of succesful computation. Actually, one concrete application may involve several extension strategies at the same time. For instance, propositional dynamic logic of programs interprets its program expressions at ordered pairs of states – in the semantic format $\mathbf{M}, s_{in}, s_{out} \models \pi$ – which makes a program operator like sequential composition ; a two-dimensional existential modality.

Other concrete interests in the Sofia School are knowledge systems (Vakarelov 1991) and geometrical structures (Vakarelov 1992, Balbiani, Fariñas del Cerro & Vakarelov 1997). For further reference, we list some typical modal extensions developed by Sofia logicians. (What follows is by no means complete – as we merely want to bring out general trends. Further samples are Gargov, Passy & Tinchev 1987, or Goranko 1990. De Rijke's 1993 dissertation on extended modal logic also references Bulgarian material. The special issue of the *Journal of Logic, Language and Information* 4:3 (1995) surveys more recently designed extended modal languages.) Once again, in this paper, we are mainly concerned with 'minimal modal logics' of such extended languages, interpreted over arbitrary models, without further special frame constraints.

Historical Digression 1 Intuitionistic Modal Logic

Every history is a rational reconstruction of recalcitrant facts. The actual genesis of the Sofia School lies around 1974, when George Gargov, Vladimir Sotirov and Dimiter Vakarelov were working with their students on intuitionistic modal logic. In particular, Gargov proved a completeness theorem for intuitionistic provability logic, and showed that it is incomplete under arithmetical interpretation in Heyting Arithmetic. Sotirov defended a dissertation on intuitionistic modal logic in Moscow (1977), Vakarelov one on negation over weak positive logics (Warsaw, 1976) and Gargov's dissertation was on intuitionistic mathematics and super-intuitionistic logics (Moscow, 1997). After the appearence of propositional dynamic logic (due to Pratt and co-workers around 1978), their interests changed to poly-modal logics oriented towards computer science and AI. Also, the above sketch imposes more systematicity than was observable at the time. See below for some delightful complications in the Sofia working habits...

2.3 Universal Modalities

The quantification in standard modalities is restricted to successors of the current world, and their iterations. Here are two unbounded variants, randomly accessing the whole W:

$\mathbf{M}, s \models E\phi$ iff for some $t \in W$, $\mathbf{M}, t \models \phi$

$\mathbf{M}, s \models U\phi$ iff for all $t \in W$, $\mathbf{M}, t \models \phi$

These operators allow us to start new processes of evaluation anywhere in a model.

2.4 Boolean Modal Logic

Standard modal logic refers to successors in a fixed alternative relation R. So-called polymodal logics have a family of relations R_a, each with an associated modality $\langle a \rangle$. These relations may carry further structure, e.g., that of some relational algebra, or a Kleene algebra of programs (as in propositional dynamic logic). *Boolean modal logic*

has just the Boolean operators of intersection, union and complement of relations in its language. Thus, modalities have the form $\langle\pi\rangle$ where π is some Boolean compound of basic relations, with a truth condition

$\mathbf{M}, s \models \langle\pi\rangle\phi$ iff for some t with $R_\pi st$, $\mathbf{M}, t \models \phi$

The additions to this repertoire that are needed to obtain full relational algebra are: converse, composition, and a special identity relation. We will not consider these here, but refer to van Benthem 1996, Marx & Venema 1996 for further details. It is worth noting that full relational algebra is undecidable: so that some of the 'modal spirit' is lost. Propositional dynamic logic stays on the good side of decidability by omitting some Booleans (cf. the very readable exposition of this result in the textbook Goldblatt 1987). It remains decidable when one adds intersection (as has been shown by Danecki), but adding negation will embed all of relational algebra. By contrast, Boolean Modal Logic stays decidable by having all Booleans, but dropping an explicit composition operator.

2.5 Polymodal and Polyadic Logics

The preceding example involved extensions that change the semantic similarity type. Polymodal logics interpret their operators $\langle i\rangle$ as follows:

$\mathbf{M}, s \models \langle i\rangle\phi$ iff for some t with $R_i st$, $\mathbf{M}, t \models \phi$

More concretely, propositional dynamic logic interprets program modalities $\langle\pi\rangle$:

$\mathbf{M}, s \models \langle\pi\rangle\phi$ iff for some t with $R_\pi st$, $\mathbf{M}, t \models \phi$

making sure that the various accessibility relations R_π cohere in the right fashion. (E.g., $R_{\pi 1 ; \pi 2}$ must be the binary relational composition $R_{\pi 1} \circ R_{\pi 2}$.) For this purpose, propositional dynamic logic employs essentially a two-dimensional modal operator:

$\mathbf{M}, s, t \models \pi_1 ; \pi_2$ iff for some u, $\mathbf{M}, s, u \models \pi_1$ and $\mathbf{M}, u, t \models \pi_2$

Properly viewed, this is a concrete example of a more abstract polyadic modal logic, involving a higher arity. Treating ordered pairs (s, t) as objects x per se, the abstract scheme of interpretation exemplified here may be stated as follows:

$\mathbf{M}, x \models \Diamond\phi\psi$ iff there exist two (!) worlds y, z with $R^3 x, yz$ such that $\mathbf{M}, y \models \phi$ and $\mathbf{M}, z \models \psi$

We will encounter further polyadic modal operators in spatial and temporal logics.

Historical Digression 2 Mixed Polymodal Strategies

Section 2.4 was a modern reconstruction. The actual presentation in Gargov, Passy & Tinchev 1987 is a good deal more involved. Essentially, what happens there is this. One adds one new modal operator (a 'converse universal modality', cf. van Benthem 1979):

$\mathbf{M}, s \models [[]]\,\phi$ iff for all t with $\mathbf{M}, t \models \phi$, Rst holds

This is an interesting modal extension in its own right, over the same similarity type. Equivalently, one can add a Boolean complement modality $[-R]$. A decidable logic is then determined for this system, using a poly-modal variant with $[R]$, $[S]$. One imposes suitable constraints on $[S]$ to make its relation S behave as much as possible like the complement of R . (The finishing touch is an unraveling construction due to Vakarelov, which turns the 'quasi-models' obtained by (say) modal filtration into standard models.) Next, a similar technique is used for Boolean intersections $R \cap S$. One adds a third modality $[T]$ and forces T to behave like $R \cap S$ as far as possible in polymodal logic, using some model surgery afterwards to achieve the intended interpretation. This is more like a combination of the above-mentioned pure extension strategies with (after all) additional frame constraints. We will discuss some of the resulting issues later on.

We conclude with some further extensions demonstrating the scope of the Sofia School.

2.6 Names

Worlds are anonymous in the standard modal language. But one can introduce special propositional constants $a, b, c, \ldots$ whose interpretation is restricted to hold in exactly one world. These then serve as names for individual worlds, provided that we consider only models satisfying this constraint on their valuations for these propositional constants. This device has been discovered independently by many authors (including Prior; Blackburn 1990 uses it to name reference points in temporal logic): the Sofia logicians put it to use in obtaining axiomatizability results for hitherto intractable modal systems.

2.7 Two-Sorted Domains

In standard models, the domain W is 'flat'. But many actual applications have different sorts of objects. A good example are modal logics of time, where both 'points' and 'branches' are natural, or logics of space, which may have points and lines (and solids). Another good example is the paper Vakarelov 1992 on geometrically structured arrows. One can introduce heterogeneous models, having different domains, and modalities that either stay inside one domain, or cross between them. For instance, with points p and line segments l , one could have geometrical modalities like the following

M, p |= ON ϕ iff there exists a line l with p on l and **M**, l |= ϕ
M, l |= AT ϕ iff there exists a point p with p on l and **M**, p |= ϕ

Of course, models with arbitrary numbers of sorts can be dealt with in the same way. Note that these logics are again polymodal ones, possibly involving additional constraints to ensure the proper structural relations between objects of different sorts. More-sortedness also occurred in the above-mentioned abstract view of dynamic logic. Making ordered pairs into new objects amounts to introducing a new sort of primitive. This move is systematized in 'Arrow Logic', a two-sorted treatment of dynamic logic with 'state' and 'transition' domains on a par (cf. Venema 1996, van Benthem 1996).

2.8 Dyadic Modalities in Space and Time

Other natural geometrical modalities, even one-sorted, are more complex, and exemplify the earlier-mentioned polyadic type. Here is one simple example:

M, l |= AND $\phi\psi$ iff l is a disjoint sum of two lines l', l'' (in that order) such that **M**, l' |= ϕ, **M**, l'' |= ψ

Dyadic modalities occur across many applications – e.g., in categorial or relevant logics (van Benthem 1991, Kurtonina 1995), as well as in arrow logic (which treats binary relational composition through ternary composition of arrows).

But modalities can also be polyadic in the sense of having more than one argument, and yet fall under our heading of extended modal logic over the same accessibility relation. A good source for this is *Temporal Logic*. Standard tense logic is like basic modal logic, looking at both successors and predecessors in the alternative relation $<$. Its modalities F ("at least once in the future") and P ("at least once in the past") correspond somewhat to tenses in natural language. Natural binary modalities arise, however, when rendering temporal adverbs like "Until" or "Since":

M, t |= UNTIL $\phi\psi$ iff there exists t' with t<t' such that **M**, t' |= ϕ , and for every t'' with t<t''<t', **M**, t' |= ψ

The clause for SINCE is the past dual of that for UNTIL. These modalities are unlike the above in that they involve an interplay between existential and universal quantifiers. They provide a big jump in expressive power. By Kamp's Theorem, over complete linear orderings (though not over models of branching time), they yield full first-order logic: the limit to which many extended modal logics aspire...

2.9 Further Extensions

This concludes our sampling of extended modal languages. It is by no means complete. E.g., many-dimensional modal logics (Venema 1991, Marx & Venema 1996) show an immense variety. As a final counterpoint to the above cases, we mention neighbourhood semantics, which reinterprets the existential modality in a second-order fashion:

$\mathbf{M}, s \models \Diamond\phi$ iff there exists a set of worlds Y with RxY such that $\mathbf{M}, t \models \phi$ for all $t \in Y$

This subsumes the standard case by setting RxY iff $Y = \{y\}$ for some y with Rxy. We will use this case occasionally to test the limits of our over-all first-order approach.

By now, we have assembled a perhaps bewildering menagerie of extended modal logics. It is time to see if any unity remains, and if any significant questions survive at this level of generality. To this end, we first introduce a general perspective on modal languages, through translation into first-order logic (Section 3). After that, we introduce three major unifying themes of simulation, decidability and correspondence (Sections 4, 5, 6, resp.). Finally, we shall return to the issue of what defines Extended Modal Logic (Section 7).

3 A General First-Order Perspective

3.1 First-Order Translation

All modalities listed here fit into one uniform format. They may be translated into the natural first-order language over the earlier models (sorted as needed), which has binary (ternary,...) relation symbols for the alternative relations, unary symbols for the (predicates of worlds corresponding to) proposition letters, and (when needed) identity. Here are some concrete illustrations of how this works:

$\mathbf{t}(\Diamond\phi)$	=	$\exists y\,(Rxy \,\&\, \mathbf{t}(\phi)(y))$	(where y is new)
$\mathbf{t}(\Diamond\phi\psi)$	=	$\exists yz\,(Rx,yz \,\&\, \mathbf{t}(\phi)(y) \,\&\, \mathbf{t}(\psi)(z))$	(with y, z new)
$\mathbf{t}(E\phi)$	=	$\exists y\ \mathbf{t}(\phi)(y)$	(where y is new)
$\langle\!-R\cap S\rangle\phi$	=	$\exists y\,(\neg\, Rxy \,\&\, Sxy \,\&\, \mathbf{t}(\phi)(y))$	(where y is new)
$ON\ \phi$	=	$\exists l\,(Rpl \,\&\, \mathbf{t}(\phi)(l))$	(where l is new)
$UNTIL\ \phi\psi$	=	$\exists y\,(\ x<y \,\&\, \mathbf{t}(\phi)(y) \,\&\, \forall z\,((Rxz \,\&\, Rzy) \rightarrow \mathbf{t}(\psi)(z)\ )$	

When combined with the obvious translation of Booleans, formulas of modal languages translate into various *fragments* of first-order logic. A general question then arises at once. What first-order fragments may be called modal in some reasonable sense?

One answer involves *finite-variable restrictions*. Basic modal logic sits inside the so-called 'two-variable fragment'. More precisely, by suitable 'recycling' in its modal clause, the above translation can make do with just two world variables x, y. The same is true for e.g. the universal modality and Boolean Modal logic. (Semantically, this means that evaluation for these modal languages never stores more than two worlds at a time.) Three variables become essential with dyadic modalities, and also in temporal logic. (Temporal betweenness is an essential '3-configuration'.) Thus, we get a landscape of modal fragments of predicate logic, all the way up to the full (undecidable) system. Finite-variable levels are one way of structuring first-order logic, but we shall encounter an alternative in Section 5, based on *atomic bounds* for quantifiers – another prominent syntactic feature of the above translations. The landscape of 'modal fine-structure' has been studied in de Rijke 1993, Marx & Venema 1996, van Benthem 1996, Andreka, van Benthem & Nemeti 1996. In particular, the latter paper investigates in depth why most meta-properties of first-order logic seem to be inherited by natural modal fragments – but also, what it is that keeps the latter decidable all the same.

For the rest of this paper, we shall not emphasize syntactic structure so much as broad semantical and algorithmic concerns that run across extended modal languages in their first-order guise (and sometimes even beyond).

3.2 Invariance for Semantic Equivalences

The expressive power of a language can be measured by finding a matching semantic relation of 'indistinguishability'. First-order formulas are invariant for isomorphisms between models, in an obvious sense. But they are also invariant for coarser forms of comparison. Let us define a *potential isomorphism* between models $\mathbf{M}, \mathbf{N}$ as a non-empty set $\mathbf{I}$ of finite partial isomorphisms between $\mathbf{M}, \mathbf{N}$ which satisfy the following back-and-forth properties. Given any partial isomorphism $f \in \mathbf{I}$ and any object $d \in |\mathbf{M}|$, there exists an object $e \in |\mathbf{N}|$ such that $f \cup \{(d, e)\}$ is in $\mathbf{I}$, too. And the same must hold in the opposite direction. Infinite models can have a potential isomorphism without being isomorphic, but countable potentially isomorphic models are always isomorphic, by a Cantorian zigzag argument. A straightforward induction on formulas shows that

Proposition Let $\mathbf{I}$ be a potential isomorphism between $\mathbf{M}$ and $\mathbf{N}$, and let $f \in \mathbf{I}$. For each first-order formula ϕ and each assignment a from its free variables to the domain of f, we have the equivalence $\mathbf{M}, a \models \phi$ iff $\mathbf{N}, f \circ a \models \phi$

We express this fact by saying that "first-order formulas are *invariant for* potential isomorphisms". (The same result holds for formulas of $L_{\infty\omega}$ with arbitrary *infinite*

conjunctions and disjunctions.) This notion is very important, and it may be modified in various ways. In particular, a *k–potential isomorphism* is a non-empty set **I** as above, but with partial isomorphisms only up to size k , and back-and-forth properties only up to this length. (Moreover, the set **I** is to be closed under restrictions to subdomains.) The syntactic correlate are the above-mentioned *finite-variable fragments* L_k which use only k variables (free or bound). Formulas from the k–variable fragment are invariant for k–potential isomorphism, just as above. Moreover, we have a converse result, which is a semantic preservation theorem characterizing the syntactic finite-variable fragments:

Theorem A first-order formula $\phi = \phi(x_1, .., x_k)$ is definable inside the k–variable fragment using only $\{x_1, .., x_k\}$ iff it is invariant for k–potential isomorphism.

For a proof and discussion of this result, cf. van Benthem 1991, Andréka, van Benthem & Németi 1996. The art is now to find similar characteristic semantic equivalences for *modal fragments* of first-order logic. For this purpose, we can make do with invariances for simpler correlations between the two models involved, that only measure relevant modal structure of accessibility. This theme will be taken up in Section 4 below, which introduces varieties of 'bisimulation' (van Benthem 1976). Insofar as modal languages translate into the two-variable fragment, the preceding theorem predicts that semantic equivalences matching only single worlds and ordered pairs of worlds will suffice. But the above bounded fragments suggest other kinds of equivalence (cf. Section 5 below), which work for a much wider range of modal languages. These 'guarded bisimulations' do not impose any upper bound on the size of the partial isomorphisms involved. They rather modify the back-and-forth conditions themselves, referring only to new choices of objects that stand in certain relationships to those already matched.

3.3 Decidable Fragments

Church's Theorem says that full first-order logic is undecidable. But many important fragments are *decidable* (cf. the recent survey Börger, Grädel & Gurevich 1996). 'Fragments' are often defined by restrictions on predicate arities, quantifier prefixes, or the above variable supplies. But the modal fragments owe their decidability not directly to any of these. For instance, even finite-variable fragments are not good enough as a measure. Basic modal logic and many of its variants translate into L_2 , which is indeed decidable (cf. Graedel & Kolaitis 1997), but many other decidable modal and temporal logics translate into L_3 , which is undecidable. (E.g., it suffices for encoding the above-mentioned undecidable full algebra of binary set relations.) We shall return to this issue in Section 5, starting from the analysis of Andréka, van Benthem & Németi 1996.

3.4 Second-Order Frame Conditions

All considerations so far concerned modal logic as a first-order language, over models **M** = (W, R, V) . But the typical correspondences between modal axioms and relational conditions rather concern *frames* **F** = (W, R) . A modal formula is true in a frame if it is true at each world of that frame for every valuation over it. This makes frame truth a *second-order* notion, and much of traditional correspondence theory is concerned with definability (first-order or otherwise) of modal formulas taken in this second-order sense (van Benthem 1984, 1985). There are algorithms determining (first-order) equivalents in this sense for wide classes of modal axioms, of which the Sahlqvist Theorem provides a prime example. (It covers the usual cases of S4, S5, as well as many further types of modal axiom that have been proposed since.) Again, these correspondence algorithms have been developed for the basic modal language. But we shall see in Section 6 how the gist of their correctness proof extends to a much wider variety of modal fragments.

3.5 Modal Fine-Structure of First-Order Syntax

Translating modal languages into first-order logic may seem like throwing them all into one amorphous bag. The vast expressive power of first-order syntax may drown the particularities of its less ambitious modal relative. But we can also look at this situation differently. Modal fragments are one way of studying interesting syntactic *fine-structure* of first-order logic, through natural categories like 'quantifier bounds', 'arities', etcetera. (Admittedly, we do not know yet what the most natural structure is.) Moreover, modal fragments show the true power of first-order techniques, which are not linked uniquely to full standard predicate logic. This 'modulation' will show in the following sections.

4 Modal Operators and Simulation Invariances

4.1 Bisimulation Invariance for Basic Modal Logic

Here is the basic semantic equivalence for modal logic (van Benthem 1976, Park 1981). It relates worlds in two models so that successor steps can be imitated reciprocally.

Definition Let $\mathbf{M}, \mathbf{N}$ be two possible worlds models. A *bisimulation* is a binary relation $\leftrightarrows$ between worlds in these two models satisfing the following two conditions:

- if $x \leftrightarrows y$, then $\mathbf{M}, x \models p$ iff $\mathbf{N}, y \models p$, for all proposition letters p
- if $x \leftrightarrows y$ and xR_az , then there exists u such that $R_a yu$ and $z \leftrightarrows u$
  if $x \leftrightarrows y$ and yR_au , then there exists z such that $R_a xz$ and $z \leftrightarrows u$

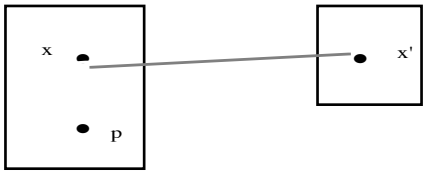
Bisimulations may be viewed as a species of 2–potential isomorphisms, with restricted back-and-forth conditions, and matches between single objects only. Note that this

definition does not require bisimulations to be total on $|\mathbf{M}|$, or surjective on $|\mathbf{N}|$. A straightforward induction shows that modal formulas are *invariant for bisimulation*:

Proposition Let $\equiv$ be a bisimulation between two models $\mathbf{M}, \mathbf{N}$ with $x \equiv y$. Then, for all modal formulas ϕ, $\mathbf{M}, x \models \phi$ iff $\mathbf{N}, y \models \phi$.

Proof The atomic step uses precisely the first clause in the definition of bisimulation. The Boolean steps are for free, as usual. The key step is that for modalities $\langle a \rangle \phi$, where the zigzag conditions provide precisely what is needed. ■

This observation subsumes the usual modal 'generation' and 'p-morphism theorems'. It may be used to show that certain notions are modally *undefinable*. For instance, an unrestricted existential statement $\exists x\, Px$ is not modally definable. It holds at x in the model pictured to the left, but not at x' in the model to the right – even though the drawn shaded link between x and x' defines a bisimulation (which trivially satisfies the back-and-forth conditions, as the alternative relation is empty in both models):



Bisimulations are ubiquitous in modal logic. For instance, the earlier-mentioned Vakarelov construction in the development of Boolean Modal Logic (Gargov, Passy & Tinchev 1987) is a new kind of 'unraveling' which gives a bisimulation between a 'quasi-model' and a standard model whose accessibilities behave as intended.

4.2 From Invariance to Definability

Conversely, one can prove various implications from invariance to definability. Here is one simple version (Hennessy & Milner 1985 – but the result also follows from a more general observation about modally equivalent saturated models in van Benthem 1976):

Proposition Let $\mathbf{M}, \mathbf{N}$ be *finite* models with $\mathbf{M}, x \models \phi$ iff $\mathbf{N}, y \models \phi$, for all modal formulas ϕ. Then there exists a bisimulation $\equiv$ between $\mathbf{M}, \mathbf{N}$ with $x \equiv y$.

Proof Define a relation between $\mathbf{M}$ and $\mathbf{N}$ by setting $x \equiv y$ iff x, y satisfy the same modal formulas in their models. Clearly, this satisfies the first condition of bisimulation ('atomic harmony'). We must now show it also satisfies the back-and-forth conditions. Here is one direction (the other is symmetric). Let $u \equiv v$, and assume that $u\, R\, s$ in $\mathbf{M}$. Enumerate v's R–successors as $\{s_1, ..., s_k\}$. Suppose these all fail as a match for s.

Then for each i (1≤i≤k) there exists a modal formula α_i true at **M**, s but false at **N**, s_i. But then the conjunction α of all the α_i is true at **M**, s , whence $\Diamond\alpha$ is true at **M**, u and hence also at **N**, v (since u≡v). By the truth definition, v must have a successor in **N** verifying α : which is a contradiction, as each s_i falsifies one of its conjuncts. ■

This result extends to all 'image-finite' (though not all infinite) models, but we omit details here. Another converse result expressing the match between the modal language and bisimulation is the *Modal Invariance Theorem* (van Benthem 1976, 1985):

Theorem A first-order formula $\phi(x)$ (of the right similarity type) is equivalent to the translation of a modal formula iff it is invariant for bisimulation.

Van Benthem 1996 has an up-to-date exploration of this kind of preservation result, which we will not prove here. (But we do note here that all results stated in this section could also be cast in the form of such model-theoretic preservation theorems.)

More finely-grained semantic analyses are possible as well. One can specialize the well-known *Ehrenfeucht-Fraïssé games* of first-order logic (cf. Doets 1996) to modal logic, with the usual Similarity and Difference players. A winning strategy for the 'Duplicator' in the game over n rounds will then correspond to modal equivalence between the two models *up to modal operator depth* n . This type of syntactic fine-structure is available in all Sofia fragments to be discussed – but we will not pursue it here.

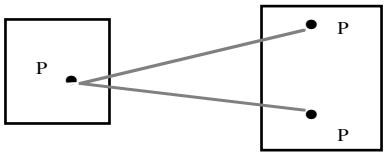
4.3 Universal Modalities and Total Bisimulations

The preceding results on bisimulation, and in particular their proofs, lend themselves easily to generalization. Thus, it is easy to see that adding unrestricted existential and universal modalities requires a new back-and-forth clause, referring to selection of some arbitrary new world to be matched, not necessarily an R-successor of the previous one. But this amounts to saying that bisimulations must now be *total* and *surjective*: for short, 'total'. It is easy to modify the above arguments to show that all formulas in the modal language extended with the universal modality are invariant for total bisimulations. And also, that on finite models, equivalence in this richer modal language implies the existence of a total bisimulation. Essentially the same arguments are involved as before.

A pleasant practical use of simulations are quick proofs of *undefinability*. We conclude with an example of something which is beyond the language of universal modalities, namely, an operator from Boolean Modal Logic:

<–R>p with first-order translation $\exists y (\neg Rxy \,\&\, Py)$

(Other cross-comparisons with extended modalities on our list can easily be supplied, finding undefinabilities with similar simple pictures.) It suffices to exhibit a total bisimulation between worlds in two models that disagree on the truth value of $<\!\!-R\!>\!p$:

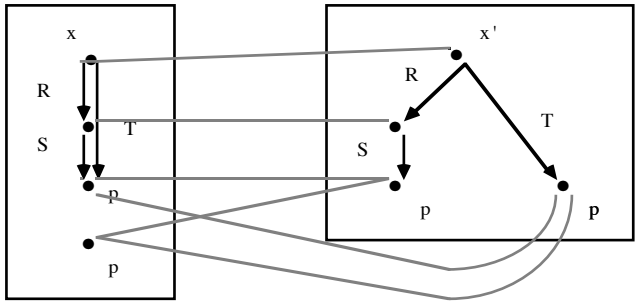


4.4 Boolean Modal Logic and Linked Pairs

By standard Boolean equivalences, Boolean Modal Logic has relations that may be written as finite unions of conjunctions of atomic relations and their complements. (For convenience, we assume a finite vocabulary of atomic relations.) The appropriate notion of a bisimulation has back-and-forth clauses for each type of successor world – one for each conjunction of (negated) atoms. (Boolean unions take care of themselves, via the valid equivalence $<\!R \cup S\!>\phi \leftrightarrow <\!R\!>\phi \vee <\!S\!>\phi$.) It is easy to prove, by the earlier inductive argument, that all BML-formulas are invariant.

As a consequence, we can show that e.g. relational *composition* is undefinable in BML. The shaded lines in the following picture give a BML-bisimulation (this may be checked case-by-case), while the worlds x, x' disagree on the truth value of the formula

$<\!(R \circ S) \cap T\!>\!p$:



The earlier argument on finite models (applied to each composite relational modality separately) provides a converse from BML-equivalence to bisimilarity in the new more demanding sense. There is also another matching bisimulation for this modal language. BML is like dynamic logic, in having both 'programs' and 'formulas'. One can also set up a characteristic bisimulation accordingly, as linking both (suitable) worlds and (suitable) pairs of worlds. This would be a variation on the earlier '2–simulations' for the two-variable fragment. (A complete semantic treatment along the latter lines would bring in a modification of 'safety for bisimulation', as introduced in van Benthem 1996, Ch. 5.)

4.5 Names and Simulations with Distinguished Worlds

Names give our modal models a number of distinguished worlds (say, the 'root' and other possible vantage points in a computational process.) The new bisimulations $\equiv$ will have to respect these distinguished states in the following sense:

whenever $x \equiv y$ and x satisfies a name letter p, then so does y – and vice versa.

This is just a modification of the atomic harmony clause in our notion of simulation. (Van Benthem 1996, Ch. 5 uses such modified simulations to obtain 'parametrized' versions of modal invariance theorems.) We omit further treatment of this case, as it involves no interesting changes in back-and-forth conditions.

4.6 Two-Sorted Cases

This case is totally straightforward. Bisimulations will only run between objects of the same sort in the two models being compared. Their back-and-forth clauses will follow whatever sortal structure is imposed by the relations underlying the modalities.

4.7 Simulations for Polyadic Modalities

Polyadic modal logic offers no great challenge to the preceding analysis. E.g., binary modalities do not change the above results. Their existential interpretation of the form $\exists y_1 y_2 \, (R^3 x, y_1 y_2 \, \& \, \phi_1(y_1) \, \& \, \phi_2(y_2))$ leads to the following semantic modification. Given any match $x \equiv x'$, the back-and-forth clauses refer to a ternary situation $R^3 x, y_1 y_2$ holding in one model, and require the simultaneous existence of a matching situation $R^3 x', y_1' y_2'$ in the other, with $y_1 \equiv y_1'$ and $y_2 \equiv y_2'$.

4.8 Simulations for Temporal Operators

Basic temporal logic is easily incorporated. One merely adds 'dual 'backward clauses' to what we had before. Things change with the more complex binary modalities SINCE and UNTIL. E.g. UNTIL involved an interchange of existential and universal quantifiers, which poses more of a challenge. These binary temporal operators refer essentially to the geometrical notion of *betweenness*. So, we modify the back-and-forth conditions for a standard modal bisimulation:

- if $x \equiv x'$ and $x < y$ in **M**, then there exists some y' with $x' < y'$ in **N** such that (i) $y \equiv y'$, (ii) whenever $x' < z' < y'$ holds in **N**, then there is a z in **M** with $x < z < y$ and $z \equiv z'$
- and vice versa.

This change in direction allows us to prove simulation invariance for UNTIL statements. Suppose that $\mathbf{M}, x \models \mathrm{UNTIL}\phi\psi$. Then there exists some y with $x<y$ and $\mathbf{M}, y \models \phi$, such that each intermediate world z satisfies ψ . Consider x' in $\mathbf{N}$, plus some $<$– successor y' given by the above clause. As $y \equiv y'$, y' satisfies ϕ in $\mathbf{N}$ by the inductive hypothesis. Moreover, consider any z' in between x' and y' . By the new simulation clause, it has a matching world z in between x and y in $\mathbf{M}$. The world z satisfies ψ – and hence so does z' , again by the inductive hypothesis. ■

We also prove a finite converse, to have a less trivial variation on the earlier argument.

Proposition Let the finite models $\mathbf{M}, x$ and $\mathbf{N}, y$ satisfy the same statements in the UNTIL tense logic. Then there exists an extended bisimulation between these two models of the above-defined kind, which links x to y .

Proof Again we show that tense-logical equivalence *itself* provides such a bisimulation. Suppose that x, x' satisfy the same tense-logical formulas. Let none of the finite set $\{y_1', ..., y_k'\}$ of $<$–successors for x' in $\mathbf{N}$ qualify as a match for y . Thus, for each i $(1 \leq i \leq k)$, we are in one of two cases. Either some formula α_i is true at y but false at y_i' (let us call these successors of type I), or no such difference exists, but there is a world z_i' between x' and y_i' which has no match in between x and y in $\mathbf{M}$ (successors of type II). Given the difference between the two types, each type II successor must satisfy the conjunction $\&_i \alpha_i$ of all the above type-I formulas α_i . Next, given any type II world y_i' , consider its 'defective' intermediate world z_i' . Let $\{z_1, .., z_m\}$ be the fixed finite set of worlds intermediate in $\mathbf{M}$ between x and y . For each such world z_j , there is a formula β_{ji} true at z_j but false at z_i' . Hence, all these intermediate worlds satisfy the disjunction $\beta_i = \vee_j \beta_{ji}$ of all the formulas β_{ji} . The crucial fact is that

$$\mathbf{M}, x \models \mathrm{U}\, \&_i \alpha_i \; \&_i \beta_i$$

But then, since x and x' verify the same temporal formulas, the same must be true at x' in $\mathbf{N}$. That is, there exists some successor y_m' of y where $\&_i \alpha_i$ holds, while all intermediate worlds between x' and y_m' satisfy $\&_i \beta_i$. But this contradicts the above construction. The first part tells us that y_m' is of type II, while the second contradicts the fact that 'its' defective intermediate world satisfies the negation of every β_{jm} . ■

This result is somewhat ugly, as the change to standard bisimulation seems very ad-hoc. A more natural characteristic simulation for temporal logic has been given only recently in Kurtonina & de Rijke 1997, using 'asymmetric' back-and-forth clauses. (But see the next section for a different take, criticizing the design of the temporal UNTIL language.)

4.9 From Languages to Simulations, and Vice Versa

We have seen how modal languages arise out of first-order logic, by varying the basic equivalence between models – in particular, by modifying back-and-forth conditions. One use of these simulations is practical: we can show that things are undefinable in the matching language by means of simple counter-examples. This illustrates a direction *from existing languages to appropriate simulations*.

But there is also a converse direction *from simulations to new languages*. E.g. we noted nothing much changes in simulation when we provide each relation with its *converse*. Hence all the above languages can be extended from 'modal' to 'temporal' ones. Simulation also suggests wholly new formats of definition. For instance, with both propositions and relations explicitly present in the modal language, it makes sense to work with modal simulations that relate both worlds ('length 1') and pairs of worlds ('length 2'). Thus, they become variations on the earlier-mentioned 2–potential isomorphisms for first-order logic (correlated with the use of just two world variables). Continuing along this line, temporal 'betweenness' even calls for inspection of matched triples (matches of 'length 3'). In this light, the standard temporal SINCE,UNTIL language seems somewhat ill-designed – and we may want to *strengthen* it to obtain better semantic behaviour (rather than twist our simulations). A much more elegant system of temporal logic would follow the natural structure of points, pairs, triples, and have a family of existential 'more-dimensional modalities' reflecting this. (Cf. van Benthem 1991, Ch. 17 for some proposals to this effect.) In other words, simulations naturally invite a passage to *many-dimensional modal logics* that 'complete' possibly awkward standard modal languages. (Marx & Venema 1996 presents a general theory.)

We conclude with an alternative modal design, which also fits well with bisimulation, and (as noted earlier) with the Sofia spirit. Instead of working with tuples as set-theoretic composite objects, one may also view pairs, triples, etc. as *primitive modal objects* in their own right. This happens in Arrow Logic, a decidable modal re-analysis of relational set algebra (van Benthem 1996, Venema 1996). Here we use many-sorted bisimulations relating (new-style) objects of the same sort. Back-and-forth conditions are still central in this alternative set-up, now w.r.t. intra-sortal predicates like "arrow a is a composition of a_1 and a_2 ", or cross-sortal predicates like "arrow a starts with point x ".

4.10 Applications and Extensions

In designing extended modal languages, one wants a balance between expressive power and complexity of valid inference. In addition, one presumably wants to preserve the major properties that make first-order logic so attractive, like the interpolation theorem,

or key semantic preservation theorems. Having a natural simulation characterization is an important step towards such meta-results. De Rijke 1993 has the meta-equation

"*modal logic : bisimulation = first-order logic : potential isomorphism*"

The preceding discussion has extended this to a pluralized fine-structure variant:

"*modal logics : simulations = first-order logics : potential isomorphisms*"

Andréka, van Benthem & Németi 1996 obtain many results via the latter methodology. (Cf. also van Benthem 1996, Ch. 4, for further illustrations in modal process theories.) But simulation analysis is not restricted to first-order logic. The symmetric form of invariance conditions typically allows arbitrary *infinitary* Boolean combinations. Barwise & van Benthem 1996 show how the above results may be extended to infinitary modal logics, where invariance theorems become generalized interpolation theorems. Things even become more natural there. (Two models are elementarily equivalent in the first-order logic $L_{\infty\omega}$ with arbitrary set conjunctions/disjunctions iff they are potentially isomorphic. Two models are bisimilar iff they share the same infinitary modal theory.)

Finally, *higher-order* modal features may also be treated through simulation analysis. We conclude with one example of this phenomenon, which is a good test for the reader's understanding. It also demonstrates one more similarity type for extended modal logic.

Example Finitary Accessibility Relations
Consider alternative relations Rx, Y which relate a single world x to a set of worlds Y (its 'support'). Recall that in this neighbourhood semantics, $\mathbf{M}, x \models \Diamond\phi$ iff there is some set Y with Rx, Y such that for all $y \in Y$, $\mathbf{M}, y \models \phi$. Here is one appropriate simulation condition for this language (its symmetric dual is obvious):

> Let $x \in |\mathbf{M}|$, $y \in |\mathbf{N}|$ with $x \equiv y$, and let Rx, Z . Then there exists a set U such that Ry, U and for each $u \in U$, there exists some $z \in Z$ with $z \equiv u$.

It is a nice exercise to prove that on finite models, the existence of such a bisimulation is equivalent to equivalence in neighbourhood semantics for the basic modal language.

5 The Bounds of Decidability

5.1 Filtration and Decidability

The basic modal language is decidable. It even has an effective finite model property, which may be established via the usual *filtration* technique. That is, given any model for

a modal formula, we can pass on to some related finite model, constructed using just truth values for subformulas of that formula. From a first-order point of view, filtration amounts to model contraction by identifying assignments modulo their behaviour on all formulas up to some fixed finite quantifier depth. When proving decidability for extended modal languages, one can add epicycles to the filtration method, stretching its range of applicability (cf. Marx 1995). We will not pursue this particular analysis here.

One can also analyze the situation 'from above', asking which fragments of first-order logic are decidable 'for modal reasons'. One answer here are finite-variable fragments. The *two-variable fragment* L_2 is decidable (cf. Grädel, Kolaitis & Vardi 1997; Mortimer in the 70s even showed an effective finite model property), and this accounts for basic modal logic, as well as many of the above extensions (e.g., with universal modalities, Boolean operations, or names). But the three-variable first-order fragment L_3 is undecidable, even though natural decidable temporal logics involve three world variables. (Moreover, L_3 lacks the usual first-order meta-properties of interpolation and preservation.) Therefore, Andréka, van Benthem & Németi 1996 provide an alternative, which rather focuses on the bounding patterns ('guards') used in modal quantification.

5.2 The Guarded Fragment

The *Guarded Fragment* GF of first-order logic generalizes many modal languages, allowing all quantifier patterns $\exists \mathbf{y}\,(Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y}))$, where the atom $Q\mathbf{xy}$ is the 'guard'. Variables in the finite sequences $\mathbf{x}, \mathbf{y}$ may occur in any multiplicity and order. There exists a semantic notion of 'guarded simulation', which can be used to characterize the guarded formulas precisely as the first-order ones having the corresponding invariance. The proof of the latter result is like that of the Modal Invariance Theorem. We therefore omit it, and pass on to a less obvious observation, which explains the interest of GF.

Theorem Universal validity in the Guarded Fragment (GF) is decidable.

Under the obvious first-order translations for their semantic truth conditions, this result explains and extends the decidability of a large class of standard modal languages, from basic modal and tense logic to even the polyadic version of first-order relativized cylindric algebra (known as 'CRS'). We will check the above cases in a moment.

Proof We recall the basic steps. Any satisfiable GF-formula ϕ has a finite 'quasi-model', of 'types' consisting of subformulas of ϕ, of some effectively computable size, which also conversely generates a model for ϕ. Thus, whether a guarded formula is satisfiable is equivalent to its having a finite quasi-model – a decidable property.

From Standard Models to Finite Quasi-Models Suppose that formula ϕ is satisfiable in standard model $\mathbf{M}$. Let V be the set of variables occurring in ϕ (free or bound). Henceforth, we restrict attention to the finite set Sub_ϕ consisting of ϕ and all its subformulas, closed under simultaneous substitutions using only variables in V, that do not change syntactic forms. (This is feasible, by the cited references.) Each variable assignment verifies a 'type' Δ of finitely many formulas from this set. Our quasi-model has a universe consisting of the finitely many types realized in $\mathbf{M}$. In this structure, for each guarded formula $\exists \mathbf{y}\,(Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y})) \in \Delta$, there exists a type Δ' with (i) $Q\mathbf{xy}$, $\psi(\mathbf{x}, \mathbf{y}) \in \Delta'$, (ii) Δ, Δ' agree on all 'unaffected' formulas with only free variables in $\mathbf{x}$.

Definition (i) Let F denote the finite set of all guarded formulas of length $\leq |\phi|$ that use only variables from V. Note that $\phi \in F$ and F is closed under taking subformulas and 'alphabetic variants'. (ii) An F-*type* is a subset Δ of F for which we have

(a)	$\neg \psi \in \Delta$	iff	*not* $\psi \in \Delta$	whenever $\neg \psi \in F$
(b)	$\psi \wedge \xi \in \Delta$	iff	$\psi \in \Delta$ *and* $\xi \in \Delta$	whenever $\psi \wedge \xi \in F$
(c)	$[\mathbf{u}/\mathbf{y}]\psi$	implies	$\exists y\, \psi \in \Delta$	whenever $\exists y\, \psi \in F$

$[\mathbf{u}/\mathbf{y}]\psi$ comes from ψ by replacing each free variable in $\mathbf{y}$ with the corresponding variable in $\mathbf{u}$, simultaneously. (iii) Let $\mathbf{y}$ be a sequence of variables, and Δ, Δ' types. Write $\Delta =_{\mathbf{y}} \Delta'$ if Δ, Δ' have the same formulas with free variables disjoint from $\mathbf{y}$. (iv) A *quasimodel* is a set of F–types S such that, for each $\Delta \in S$ and each guarded formula $\exists \mathbf{y}\,(Q\mathbf{xy} \wedge \psi) \in \Delta$, there is a type $\Delta' \in S$ with $Q\mathbf{xy}$ and $\psi(\mathbf{x}, \mathbf{y})$ in Δ' and $\Delta =_{\mathbf{y}} \Delta'$. We say that ϕ *holds in a quasi-model* if $\phi \in \Delta$ for some Δ in this model. ■

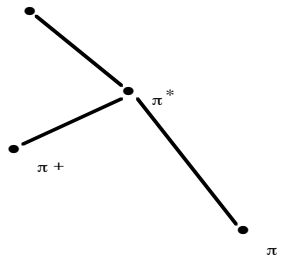
Clearly, if ϕ is satisfied by some model, then ϕ also holds in some quasi-model.

From Quasi-Models to Standard Models From any quasi-model $\mathbf{M}$, we can define a standard model $\mathbf{N}$. Call π a *path* if $\pi = < \Delta_1, \phi_1, ..., \Delta_n, \phi_n, \Delta_{n+1} >$ where Δ_1, Δ_{n+1} are types in $\mathbf{M}$, each formula ϕ_i is of the form $\exists \mathbf{y}\,(Q\mathbf{xy} \wedge \psi) \in \Delta_i$ and Δ_{i+1} is an alternative type as described above (i.e., $Q\mathbf{xy}$, $\psi(\mathbf{x}, \mathbf{y})$ in Δ_{i+1} and $\Delta_{i+1} =_{\mathbf{y}} \Delta_i$). We say that the variables in $\mathbf{y}$ *changed their values* from Δ_i to Δ_{i+1} (the others did not). Finally, variable z is called *new in path* π if either $|\pi| = 1$ or z's value was changed at the last round in π. *Objects* in $\mathbf{N}$ are all pairs (π, z) with π a path, z new in π. Next, we *interpret predicates* over these objects. I(Q) holds of the sequence of objects $<(\pi_j, x_j)>_{j \in J}$ iff the paths π_j fit into one linear sequence under inclusion, with a maximal path π^* such that (i) the atom $Q<x_j>_{j \in J} \in \Delta^*$ (the last type on π^*) and for no (π_j, x_j) does x_j change its value on the further path to the end of π^*. Finally, we define an *assignment* s_π for each path. We set $s_\pi(x) =_{def} (\pi', x)$ with π' the unique

subpath of π^* at whose end x was new, while it remained unchanged afterwards. The correctness of this model construction shows at *last*(π), the last type on the path π :

Truth Lemma For all paths π in $\mathbf{N}$, and all formulas $\psi \in F$,
$\mathbf{N}, s_\pi \models \psi$ iff $\psi \in last(\pi)$.

Proof Induction on ψ . ***Boolean cases*** are immediate, by the closure conditions for $\neg$ and $\wedge$ on types. ***Atoms***: involve a straightforward calculation, via the linearity condition in the interpretation function I, plus the '$=_{\mathbf{y}}$-clause' in quasi-models ensuring transfer of 'unaffected formulas' along paths. For later reference, we repeat the full argument for bounded ***Existential Quantifiers*** $\exists \mathbf{y} (Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y}))$. (i) First, suppose that $\exists \mathbf{y} (Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y})) \in last(\pi)$. Then there is an extended path $\pi^+ =_{def} \pi$ *concatenated with* $< \exists \mathbf{y} (Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y})), \Delta' >$, where Δ' is a successor type for Δ chosen as above with $Q\mathbf{xy}, \psi(\mathbf{x}, \mathbf{y}) \in \Delta'$ (satisfying the transfer condition for unaffected formulas with free variables $\mathbf{x}$). All objects (π^+, y_i) with y_i in $\mathbf{y}$ are new here. By definition, the atomic guard $I(Q)$ holds for the object tuples $s_{\pi+}(\mathbf{y})$, $s_{\pi+}(\mathbf{x})$ ($= s_\pi(\mathbf{x})$) . Also, by the inductive hypothesis, $\mathbf{N}, s_{\pi+} \models \psi(\mathbf{x}, \mathbf{y})$. Therefore, $\mathbf{N}, s_{\pi+} \models \exists \mathbf{y} (Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y}))$. By $\mathbf{x}$-invariance in the standard model $\mathbf{N}$, then, indeed $\mathbf{N}, s_\pi \models \exists \mathbf{y} (Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y}))$. (ii) Conversely, suppose that $\mathbf{N}, s_\pi \models \exists \mathbf{y} (Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y}))$. By the truth definition, there are objects $d_i = (\pi_i, u_i)$ with $\mathbf{N}, s_\pi{}^{\mathbf{y}}_{\mathbf{d}} \models Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y})$. (Here, $s_\pi{}^{\mathbf{y}}_{\mathbf{d}}$ is the assignment which is like s_π except for setting all y_i to d_i .) In particular, $I(Q)$ holds of the objects $s_\pi(\mathbf{x})$, $\mathbf{d}_i$. This leads to a picture of forking paths. The $s_\pi(\mathbf{x})$ were all introduced by stage π^* inside π , and then the $\mathbf{d}_i$ were (either interpolated, or) added to form a maximal sequence π^+ with the atom $Q\mathbf{xy}$ true at the end. The fork is such that $\mathbf{x}$-values do not change any more from π^* onward, whether toward π or π^+ . (This is the only case where the atomic guard on our quantifiers comes in essentially.) We now analyse this situation a bit more carefully:



Now, the variables u_i do not have to be the y_i . Say, π^+ has $s_{\pi+}(u_i) = (\pi_i, u_i) = d_i$. Thus, the assignments $s_\pi{}^{\mathbf{y}}_{\mathbf{d}}$ and $s_{\pi+}$ agree on $\mathbf{x}$, and for all $y_i \in \mathbf{y}$ we have $s_\pi{}^{\mathbf{y}}_{\mathbf{d}}(y_i) = d_i = s_{\pi+}(u_i)$. Then, by $\mathbf{N}, s_\pi{}^{\mathbf{y}}_{\mathbf{d}} \models Q\mathbf{xy} \wedge \psi$ and the above observations, we have $\mathbf{N}, s_{\pi+} \models [\mathbf{u}/\mathbf{y}]Q\mathbf{xy}$, $\mathbf{N}, s_{\pi+} \models [\mathbf{u}/\mathbf{y}]\psi$. By the inductive hypothesis, $[\mathbf{u}/\mathbf{y}]\psi \in last(\pi^+)$.

Also, from the initial description of π^+ , we see at once that $[\mathbf{u}/\mathbf{y}]Q\mathbf{xy} \in last(\pi^+)$ (by the interpretation of atomic predicates). By closure conditions (b), (c) for types, one gets $\exists \mathbf{y}\,(Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y})) \in last(\pi^+)$. Finally, since no changes in $\mathbf{x}$-values occurred on the fork from π^* , the transfer condition for unaffected formulas along successor types along paths ensures that this same formula is in $last(\pi)$. ■

Thus having a quasi-model implies having a real model, and the Theorem is proved. ■

Grädel 1997 has just determined the exact complexity class for the Guarded Fragment: it is complete for *doubly exponential* time. Such complexity results elaborate the fine-structure of extended modal logics. Boosts in expressive power that do not affect decidability may still increase the *complexity* of an old decision procedure. We do not have a full picture yet in this area: but most Sofia extensions do seem to pay a price: e.g., going from the known PSPACE completeness for the basic modal logic to EXPTIME. (Cf. Blackburn, de Rijke & Venema 1996 for a treatment of some extended modalities, using the techniques developed in Spaan 1993.) We will forego this fine-structure here.

5.3 Decidability of the Sofia Fragments

Let us now analyze the Sofia fragments in the guarded perspective, bearing in mind that all our first-order translations for modal truth conditions were effective. What we find is a diverse picture, which raises some new questions, both algorithmic and semantic.

5.3.1 *Basic modal logic*

Basic modal logic translates into the Guarded Fragment GF, as may be checked from the guarded form of the existential modality:

$$\exists y\,(Rxy \,\&\, \phi(y))$$

Given the above conjecture about the complexity of GF, and the known (lower) PSPACE-completeness of basic modal logic, this is a costly reduction complexity-wise. But we are now after general reasons for decidability, disregarding such fine-structure.

Indeed, we can use the GF reduction to motivate a passage to modal extensions with greater expressive power. For instance, here is a point reminiscent of temporal logic. Note that GF allows even more complex matrix formulas $\phi(x, y)$ referring to the initial world of evaluation (temporal "now"), without losing guardedness and decidability. Also, note that there is no restriction on the atoms occurring in the matrix. That is, unlike in basic modal formulas, R–atoms are admissible in addition to unary formulas Px .

Incidentally, this observation explains the innocence of another Sofia favourite, namely the modal statement *loop(a)* , which is true at the 'fix point' worlds s where $R_a ss$.

5.3.2 *Universal modalities*

Next, *universal and existential modalities* are typically non-guarded:

$$\forall y\, \phi(y)\ ,\ \exists y\, \phi(y)$$

Nevertheless, we can show decidability of the modal language with universal modalities in various ways – for instance, by means of a simple extended filtration argument. What we give here instead is an effective reduction to GF. This is over-kill, of course, but it does tell us about a general safe extension of GF.

Proposition Extending GF with *closed* existential quantification over unary formulas $\exists y\, \phi(y)$ leaves the system decidable.

Proof Consider any formula α in the extended language, with some distinguished occurrence of a formula $\exists y\, \phi(y)$. We show how to reduce satisfiability of this formula to that of a guarded one. First, α is logically equivalent to a disjunction by cases, viz. $(\exists y\, \phi(y) \,\&\, \alpha(T)) \vee (\neg \exists y\, \phi(y) \,\&\, \alpha(\bot))$. The latter is satisfiable iff one of its disjuncts is.

Claim $\exists y\, \phi(y) \,\&\, \alpha(T)$ is satisfiable iff $\exists y\, (Uy \,\&\, \phi^U(y) \,\&\, \alpha^U(T))$ is satisfiable.

Here the ^{U} stands for relativization of every quantifier to some new unary predicate U . Any guarded quantifiers will remain guarded under this operation. But the total formula to the right has also become guarded. Now, any model for $\exists y\, \phi(y) \,\&\, \alpha(T)$ becomes one for $\exists y\, (Uy \,\&\, \phi^U(y) \,\&\, \alpha^U(T))$ by setting U equal to the whole domain. Conversely, if $\exists y\, (Uy \,\&\, \phi^U(y) \,\&\, \alpha^U(T))$ is satisfiable, then we can pass to the submodel defined by U to satisfy $\exists y\, \phi(y) \,\&\, \alpha(T)$. Next, a similar argument works for satisfiability of the second disjunct $\neg\, \exists y\, \phi(y) \,\&\, \alpha(\bot)$, but this time, using the corresponding guarded form $\neg\, \exists y\, (Uy \,\&\, \phi^U(y)) \,\&\, \alpha^U(\bot) \,\&\, \exists x Ux$. ■

This argument extends to all closed quantifications $\exists y_1 ... y_k\, \phi(y_1 ..., y_k)$. Similar tricks work for the 'difference modality' $D\phi$ (" ϕ holds in at least one other world than x " . What is their general range, beyond the expressive power of the Guarded Fragment?

5.3.3 *Boolean modal logic*

Under the obvious translation, formulas of BML need not land up in GF as they stand. Basic modalities are indeed guarded, and so are their disjunctions and *conjunctions*. The former is because of the logical equivalence

$$\exists y ((Rxy \vee Sxy) \& \phi(x, y)) \leftrightarrow \exists y (Rxy \& \phi(x, y)) \vee \exists y (Sxy \& \phi(x, y))$$

The latter point deserves some emphasis, as adding conjunction is sometimes seen as complex, calling for essentially new modal constructions. But the translation

$$\exists y (Rxy \& Sxy \& \phi(y))$$

is guarded as it stands, with 'Rxy' for a guard, and 'Sxy & ϕ(y)' the admissible matrix. Complement relations are unguarded, however: $\exists y (\neg Rxy \& Py)$ is not in GF. (This may be proved using guarded bisimulations: cf. Andréka, van Benthem & Németi 1997.) Nevertheless, one of the major results in Gargov, Passy & Tinchev 1987 is this

Proposition Boolean Modal Logic is decidable.

Here is a quick proof: BML translates into the decidable two-variable fragment L_2. Indeed, most Sofia logics live there, which gives a uniform reason for their decidability. This first-order analysis suggests easy decidable generalizations. E.g., adding relational converses to BML stays inside L_2. On the other hand, relational *composition* would change things, as two variables no longer suffice. This is essential, as we know that the resulting system is undecidable. We would also like to find a direct reduction to GF, but we have been unable to find one. This is an instance of a more general

Question To give an effective reduction from L_2 to GF.

The analysis in the next subsection shows a strong analogy between BML and all of L_2. Using the observations made there, the following reduction can be proved effectively. A two-variable formula ϕ is satisfiable iff some conjunction $\alpha \& \phi'$ is satisfiable, with α a GF-formula providing atomic definitions for (suitable) subformulas of ϕ, and ϕ' a BML–formula using only basic modalities $\langle R \rangle$ and complement modalities $\langle -R \rangle$.

5.3.4 *Historical excursion: 'balancing' modal languages and frame conditions*

But there is more to Gargov, Passy & Tinchev 1987. It is a pleasure re-reading the many ideas in this paper – and be reminded at the same time of the golden sands of the 1987 Gödel Conference at Druzhba. The authors present an indirect proof of decidability (and effective axiomatizatizability) for BML. Here are its main features in a nutshell.

First, the GPT system is a polymodal logic with flat modalities which enforces special connections $S = -R$ and $T = R \cap S$ for basic modalities $\{R], [S], [T]$. Thus the language as such does not contain explicit BML-forms like $\exists y (Rxy \& \neg Sxy \& \phi(y))$. But there is a two-way reduction of validity. A polymodal formula ϕ is a GPT theorem

iff its first-order translation $\mathbf{t}(\phi)$ follows from the (translated) special connections. And the latter holds iff the first-order formula resulting from plugging in definitions for 'complex' modalities in $\mathbf{t}(\phi)$ (which is a BML-form) is universally valid as it stands. The converse is true, too. (This is a well-known trick in complexity arguments, encoding successive subformulas inside out by means of atoms.) But despite this connection, GPT thinking and our guarded analysis seem different. E.g., consider a Boolean complement $S = -R$. In GF, it is easy to express that two relations R, S are disjoint, but we cannot say in a guarded manner that their union is the universal relation. By contrast, using polymodal axioms, it is easy to enforce that some relation T (which GPT make into an equivalence relation) is contained in the union of R and S, but we cannot express the disjointness of the latter. (The latter has to be provided later on, up to bisimulation, via Vakarelov's model surgery.) Incidentally, $\forall xy\,(Txy \rightarrow (Rxy \vee Sxy))$ is indeed a guarded formula, but we cannot express in GF that T is an equivalence relation (see below).

The GPT approach increases expressive power in a way outside of our earlier strategies. One stays poly-modal in the working language, enforces the right connections between modalities via frame conditions, and then proves correctness through model-theoretic 'surgery'. Thus, there is a delicate balance between frame conditions and expressiveness of the modal language. In particular, we must consider the complexity of modal logics over extra frame conditions after all. There are many open questions here, as neither two-variable restrictions nor guardedness take care of all cases. For instance, *transitivity* of a binary relation is essentially three-variable, and it is also essentially non-guarded. (GF with a transitivity statement even becomes undecidable.) But then, why is modal K4 decidable? Our analysis so far provides no answer. We think that there is more to be had from analyzing modal filtration arguments, to see how they transfer accessibility structure, in combination with eventual 'model surgery'.

We conclude by formulating one more issue here, which seems of independent interest. Surgery constructions usually have the following features:

> (i) there is a model $\mathbf{M}, s$ for a modal formula ϕ whose underlying frame satisfies some first-order condition γ, (ii) we find a bisimilar model $\mathbf{N}, t$ whose underlying frame satisfies some additional condition δ.

Because of bisimulation invariance, the modal formula ϕ will still be satisfied at $\mathbf{N}, t$ – but the environment has improved to $\gamma \& \delta$. Let us say that first-order formula γ *can be boosted to* δ *via bisimulation* if each model for γ has a bisimulation-equivalent model where δ holds. One can think of boosting as a new kind of modal consequence relation.

This notion makes sense for arbitrary model relations B . It has itself a modal pattern, viz. $\gamma \rightarrow \langle B \rangle \delta$. (A similar modal pattern $\gamma \rightarrow [B] \delta$ is used to introduce a new notion of 'consequence along a model relation' in Barwise & van Benthem 1996.)

Question Is there a syntactic characterization for the relation of 'boosting by bisimulation' among first-order formulas?

5.3.5 *Names and sorts*

The case of *names* for individual worlds does not affect decidability. Admittedly, their relevant semantic model condition $\forall xy ((Px \,\&\, Py) \rightarrow x=y)$ is not in a guarded form. (GF does remain decidable with identity: provided that = never serve as a guard atom.) But singleton predicates denote individual worlds, and we get guarded translations with free variables for these. (Cf. the use of 'parametrized bisimulations' in van Benthem 1996, Ch. 5.) Finally, the move to *many-sorted* modal logics is orthogonal to the GF analysis. It does not matter if we have sorted variables: all that matters is whether we have guards. In the above geometrical examples, the predicates ON and AT serve just this purpose, and hence the minimal modal logic of incidence geometry is decidable.

5.3.6 *Poly-modality and polyadic modalities*

The basic pattern of interpretation is guarded by the relevant (n+1)-ary accessibility:

$$\exists yz (Rxyz \,\&\, \phi(y) \,\&\, \psi(z))$$

and GF sees no difference between single and polyadic quantifier prefixes. Even more trivially, polymodal logic is decidable as no special guard needs to be distinguished.

5.3.7 *Temporal operators*

So far, we saw how all Sofia-style extended logics with unary modalities are decidable by translation into larger decidable first-order fragments – while the latter also suggested some easy decidable generalizations. (We did run into some open issues of complexity.) Things get more complicated in temporal logics. Basic tense logic is guarded under translation (conversion of guards for past tenses is irrelevant). But *binary* modalities provide a challenge. Interval summation or arrow composition translate into essentially non-guarded formulas, as do "Since", "Until". We need a new line, to be provided now.

5.4 Loose Guards Remain Decidable

Some natural decidable modal logics are beyond the scope of the Guarded Fragment. Recall our example of Temporal Logic. The binary modality UNTIL $\phi\psi$ said that

$$\exists y (x<y \wedge \phi(y) \wedge \forall z ((x<z \wedge z<y) \rightarrow \psi(z)))$$

Here, the "betweenness" clause has a *composite guard* $x<z \wedge z<y$, which cannot be replaced by a single one through any obvious logical equivalence. But, it is easy to show that conjunctive guards are problematic, at least in principle.

Proposition GF extended with arbitrary conjunctions of guards is undecidable.

Proof Recall that the three-variable fragment L_3 is undecidable. Now observe that a formula ϕ in this language is satisfiable iff its guarded relativization $(\phi)^U$ to some new ternary predicate U is satisfiable for some *full Cartesian product* $U = DxDxD$. The latter can be expressed as the satisfiability of a formula $(\phi)^U$ & CART(U) , where

> CART(U) $=_{def}$ (i) $\exists xyz\, Uxyz$ & (ii) $\forall xyz\, (Uxyz \rightarrow$ & U-followed -by "all permutations and identifications among {x, y, z}") & (iii) $\forall xyzuvw\, ((Uxyz \& Uuvw) \rightarrow$ & U-followed-by "all selections of three variables from among {x, y, z, u, v, w}")

Note that the latter formula is in GF with added conjunctions of guards. ■

Yet, there is another natural decidable modal logic beyond GF, viz. 'Pair Arrow Logic', i.e., relational algebra over arbitrary top relations (not just full Cartesian squares). Here, the general arrow strategy would use ternary guards $Uxyz$ for relational composition, whereas pair models use binary relativization with a composite guard $Uxy \wedge Uyz$:

$$R \circ S =_{def} \lambda xy \bullet \exists z\, ((Uxz \wedge Uzy) \wedge Rxy \wedge Szy)$$

Here is the proper generalization of GF which covers both positive examples. Call a quantification *loosely guarded* if it has the following format:

$$\exists \mathbf{y}\, (\, \& Q\mathbf{xy} \wedge \psi(\mathbf{x}, \mathbf{y}))$$

where $\& Q\mathbf{xy}$ is a conjunction of atoms with free variables $\mathbf{y}, \mathbf{x}$ in which every variable y in $\mathbf{y}$ co-occurs with every other variable in $\mathbf{y} \cup \mathbf{x}$ in at least one of the listed atoms, conjoined with a matrix formula $\psi(\mathbf{x}, \mathbf{y})$ from the Loosely Guarded Fragment. Single atomic guards exemplify this, and so does the above $x<z \wedge z<y$ (with z in the role of $\mathbf{y}$). A typical non-example is transitivity

$$\forall y_1 y_2 y_3\, ((y_1<y_2 \wedge y_2<y_3) \rightarrow y_1<y_3)$$

without co-occurrence of y_1, y_3 in a guard atom. The property CART(U) is similar.

Theorem The Loosely Guarded Fragment is decidable.

Proof We analyse the above representation argument. The definition of quasi-models carries over without major changes, as does their representation via 'path models'. Here, we now allow path extensions via the new generalized form of bounded quantification. Again, the crucial result is the Truth Lemma, saying that guarded formulas hold under the assignment induced by a path iff they occur in the last set encoded in that path. The step from right to left here is as before. Thus, the key is a combinatoric aspect of the converse direction, whose main step was illustrated in the above picture. The argument for true existential formulas still works with a conjunction of atomic guards like above. We look at the maximal position π^* as before. For each new variable y, again given the truth condition for atomic statements, loose guardedness requires that the path of the new y-value fits linearly with the original path on which the x-values occurred. Therefore, it either lies on the latter, or it extends it starting from π^*. Moreover, the condition also applies to all new values y amongst each other - and hence, these form at worst some linear path π^+ extending π^*, up to some maximal node where the highest new y-value has been introduced. The rest of the argument is as before, since all relevant y-atoms hold at π^+, and no y-values change in going back towards π^*. Cases of mere interpolation of the new y-values on the old path π are merely simpler. (Here, we heavily use the constancy of relevant variable values in an atom along the path up to the highest variable mentioned. This requires some checking of cases.) ■

By the earlier-mentioned translations, then, we have a new application. The above proof provides a uniform new argument for the following assertions.

Corollary Basic Until/Since Temporal Logic and Pair Arrow Logic are decidable.

But there are further interesting open questions concerning the Guarded Fragment as a classical mirror of modal logic. For instance, modal and dynamic logic can be extended with arbitrary *fixed-point operators* $\mu p \bullet \phi(p)$ (where p occurs only positively in ϕ) to obtain the earlier-mentioned μ*–calculus*, which is known to remain decidable.

Question Is GF still decidable when we add modal fixed-point operators?

The guarded analysis might even extend to second-order semantics, like the earlier neighbourhood models for the basic modal language. It is reminiscent of the 'restricted' second-order formulas of Orey 1959, Gallin 1975, which lead to very simple fragments of second-order logic. This analogy remains to be understood.

6 Correspondence Generalized

6.1 Sahlqvist Correspondence Via Substitutions

Finally, we consider correspondence arguments. One of the main attractions in modal logic are the systematic correspondences between modal axioms and frame properties. Textbook authors tend to view these as an art, not to be spoilt by routine algorithms – but there are effective methods for this purpose. One instance is the substitution algorithm for Sahlqvist forms. Without formulating the Sahlqvist Theorem as such, let us analyze this method by means of an example, seeing what ingredients are involved.

Proposition The modal K4-axiom $\Box p \rightarrow \Box\Box p$ corresponds to *transitivity* of the alternative relation in frames (W, R) .

Proof For precise formulations of the algorithm and its correctness proof, we refer to van Benthem 1984, de Rijke 1993. Here we merely describe the essence of the method.

(1) We translate the modal axiom so as to bring out its second-order frame form:

$$\forall x\, \forall P:\ (\forall y\,(Rxy \rightarrow Py) \rightarrow \forall y\,(Rxy \rightarrow \forall z(Ryz \rightarrow Pz)))$$

(2) Now, the antecedent in this implication is true for any predicate P iff it is true for a certain *minimal predicate* $Pu =_{def} \lambda u.Rxu$. Note that we can define this 'minimal valuation' as a first-order formula in R .

(3) We now *substitute* this first-order description for P in the above formula, to obtain the first-order substitution instance

$$\forall x:\ (\forall y\,(Rxy \rightarrow Rxy) \rightarrow \forall y\,(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz)))$$

(4) Here the antecedent is tautologically true, whence after some *simplification*, we get the desired property of transitivity:

$$\forall x:\ \forall y\,(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz)))$$

The algorithm consists in performing these substitutions. Now we turn to its *correctness*.

I In one direction, any universal second-order formula implies all its (first-order) substitution instances, and hence the frame form of $\Box p \rightarrow \Box\Box p$ implies transitivity.

II Normally, there will be no converse implication: specific instances do not imply their universal generalizations. But in the present special case, one may argue as follows.

Let (W, R) be any transitive frame, and let V be any propositional valuation over it satisfying the antecedent $\Box p$. By observation (2), the same antecedent will also be true under the *minimal valuation* described above. But then, since transitivity was a substitution instance of $\Box p \rightarrow \Box\Box p$ under this valuation, the substitution instance of the consequent $\Box\Box p$ must hold under the minimal valuation, too. But then, that consequent will still hold for the original valuation V with which we started, for two reasons: (i) V extends the minimal valuation, (ii) the consequent is a syntactically *positive* formula, which is upward *monotonic* under extending predicate denotations. ■

Now we have the key to Sahlqvist-type correspondence results. What we need are modal axioms which allow for minimal valuations in a 'triggering part', which can then be substituted into a 'positive part'. A surprising number of axioms of the form $\alpha \rightarrow \beta$ in both old and new modal logics fall under this heading, with a triggering antecedent α and a positive consequent β – provided one defines these notions appropriately. (As for *limitations* to the substitution method, cf. the counter-examples in van Benthem 1985, Ch. 10 – which uses them to present some best-possible versions of the above result.)

6.2 The General Scope of Minimal Valuations

The above analysis extends far beyond the basic modal logic. Note for a start that all that matters for the consequent is its *semantic monotonicity*. Apart from that its syntax is irrelevant, and we might allow all extended modal, first-order, higher-order, or even infinitary shapes. (The normal shape, of course, will be some sort of 'positive form'.) With antecedents, things are more complicated, as it is harder to say when we have a 'minimal valuation'. Moreover, there are some small bureaucratic details. In general, we do not process the translated antecedent as it stands, but we move out as many *existential quantifiers* as we can, turning them into universal quantifiers in the prefix of the axiom, which only serve to supply free variables for the substitutions. Only after that, semantic 'minimality' comes into play. Here is one more precise formulation of the latter property, which guarantees the existence of a smallest satisfying denotation:

> *Total Conjunctivity*
> for any family of predicates $\{ P_i \mid i \in I \}$, $\alpha(P_i)$ for all $i \in I$ iff $\alpha(\bigcap_i P_i)$

At least for first-order formulas α, there exists a complete characterization of all the syntactic shapes that have this semantic property.

Theorem A first-order formula $\alpha(P)$ is totally conjunctive in the predicate P iff it is definable as a conjunction of 'Horn forms' $\forall \mathbf{x} (\phi(\mathbf{x}) \rightarrow P\mathbf{x})$, where $\mathbf{x}$ is a finite sequence of variables, and formula ϕ does not contain the predicate P.

Proof Cf. van Benthem 1986 (Ch. 9), 1996 (Ch. 11). ■

We will not discuss further generalizations of the substitution method as such. By now, there are also extended algorithms which process non-minimal antecedents, using choice functions to pick a verifying valuation, and substituting first-order formulas with that. (The 1997 update of van Benthem 1984 has references, often from computer science.) Naturally, these produce higher-order correspondents – which can still be interesting.

6.3 Applications

The above analysis should suffice to show that correspondence is not a local peculiarity of the basic modal language. Extended modal languages support precisely the same interplay (over their minimal logic) between extra axioms and special frame conditions. In particular, all earlier Sofia fragments support extended Sahlqvist forms. Positive consequents are available, where all new modalities create positive contexts, except for Until/Since, which are positive in their first argument, but negative in their second. (Thus, the latter language calls for even parity checks to compute over-all positivity.) As for antecedents $\alpha(P)$, the universal format of the preceding theorem allows many new modal operators, since no restriction is placed on the form of the formula $\phi(\mathbf{x})$. Standard Sahlqvist arguments look at boxes $\Box = \forall y\ (Rxy \rightarrow Py)$, but one might just as well allow more complex antecedents, say, with complex program modalities $[\pi]p$:

$$[R \cap -(R)^{\smile}]p \quad = \quad \forall y\ ((Rxy \wedge \neg\ Ryx) \rightarrow Py)$$
$$[R \cap S]p \quad = \quad \forall y\ ((Rxy \wedge Sxy) \rightarrow Py)$$

Binary statements UNTILpq may also be allowed – at least in top position – as their first-order forms $\exists y\ (x<y \wedge Py \wedge \forall z\ ((x<z \wedge z<y) \rightarrow Qz))$ allow for prefix exportation of the first existential, after which a universal of the above form remains.

We may even take this observation as a desideratum for new modalities. For instance, with binary modalities, one usually defines the universal variant $\Box pq$ of $\Diamond pq$ to be

$$\neg \Diamond \neg\ p \neg\ q \quad = \quad \forall\ yz\ (Rxyz \rightarrow (Py \vee Qz))$$

This is not a Horn form as required above, because of the positive disjunction. The latter may lead to complications in correspondence (van Benthem 1985). What would work better for correspondence purposes is another dual operator, namely:

$$\neg \Diamond p \neg\ q \quad = \quad \forall yz\ (Rxyz \rightarrow (\neg Py \vee Qz))$$
$$\text{or equivalently} \quad \forall yz\ ((Rxyz\ \&\ Py) \rightarrow Qz))$$

Interestingly, the latter is precisely the standard choice made in Relevant Logic for 'relevant implication'. So, correspondence arguments seem hardly affected by the passage to extended modal logics – and they even better unfold their true generality.

What is not supplied by the preceding analysis are syntactic *axiomatization* results that form counterparts to the above semantic correspondences. (Venema 1991 shows how the above substitution arguments may be mimicked in Henkin models, provided that one adds suitable *new* rules of inference.) Also, infinitary or higher-order correspondences may be computed by the above substitution technique, transforming first-order frame properties into less familiar, but potentially interesting *generalizations*. A case in point are frame correspondences for the second-order neighbourhood semantics of sections 2.9, 4.10. For instance, the earlier modal axiom $\Box p \rightarrow \Box\Box p$ will come to express this natural generalization of transitivity for relations Rxy to a Cut Rule for relations RxY :

$$\forall x \forall Y \forall \{Z_y \mid y \in Y\}: ((R\, x\, Y \wedge R\, y\, Z_y \ (\textit{for all}\ y \in Y)) \rightarrow R\, x \bigcup \{Z_y \mid y \in Y\})$$

Thus, techniques for dealing with extended modal logics reach into second-order logic.

7 Once Again: What is Extended Modal Logic?

The landscape of extended modal logics turns out to be quite rich. But is it anything more than a quilt of diverse pieces? As the history of syncretism tell us, mixing elements from various religions may lead to all-embracing systems, but not necessarily to stable or substantial ones. Let us see, therefore, what main lines may be discerned in Modal Logic more broadly conceived. We can think of a logic here in the sense of abstract model theory, as specified by a triple $(L, M, \models)$ of a language, a model class, and a truth relation. We have first noted a general division of labour between a logical part setting up the basic framework, and the study of special-purpose constraints on models in M giving rise to particular theories in this logic. The latter direction varies deductive power, while changing the logical framework varies expressive power. Next, we noted that changes of the latter kind may affect various components of the triple $(L, M, \models)$. For instance, one might change L only, without changing M, or vice versa. (But any change in L or M will bring about a change in $\models$.) The component $\models$ is independent from the other two, however. One can change the truth definition for a fixed language over its original model class. Many extended modal logics change all parameters at once, witness cases where the similarity type changes, and the truth definition generalizes.

These are still semi-empirical observations. They reflect a mathematical landscape of all modal logics, which is further structured by various relations, such as (conservative) extension, (faithful) interpretation, as well as various operations. (E.g., polymodal logic is a 'direct sum' of basic modal logics.) One could locate the unity of the field in some category-theoretic structure of this kind. But this seems too unfocused as an inspiring definition. Another general approach proceeded via translation into first-order logic, as a 'universal language' for modal truth conditions. Then, modal logic might be defined as the study of first-order syntactic fine-structure. But this leaves out non-first-order systems with 'modal behaviour', like dynamic logic, or the μ–calculus. Therefore, in this paper, we have opted for analyzing the field more 'dynamically', in terms of certain broad techniques, namely simulation invariances, complexity analysis, and computable correspondences. But this too does not give us the essence of Modal Logic. If pressed, we would say its core is the proper balance between expressive power and complexity. For the moment, we must leave matters at that.

Digression Those Exasperating Twists

What aggravates these matters is the Protean nature of logic. Distinctions are seldom hard and fast. Here are some 'modal' strategies that cross the above neat boundaries. (1) Instead of imposing a model constraint, one can change a truth definition. The modal logic T arises without frame constraints by reading $\Box p$ as $\forall y\,(Rxy \rightarrow Py) \wedge Px$. (2) Dynamic logic is either a minimal logic for its modal vocabulary, or a poly-modal one with special axioms enforcing appropriate connections between accessibility relations. (3) Many-dimensional modal logics generalize unary modal ones, but they are also (many-sorted) special cases of them, when we make tuples into new primitive objects. (4) Second-order neighbourhood semantics becomes first-order with 'sets' as a new sort. The unity of Logic lies in invariances across these twists, which may be hard to see...

Even without grand definitions, there can be progress. Our analysis has brought to light a number of unifying themes. Moreover, it has generated some interesting new questions, e.g. concerning complexity of extended modal logics, or their infinitary versions. Thus, we can end on a positive note. What we have learnt from the Sofia School is how modal logic is an open-ended enterprise, where more expressive operators can be added without qualms of conscience as the need arises. George Gargov and his colleagues have given us various techniques for exploring these new systems. The purpose of this paper has been to demonstrate some more general perspectives on the resulting landscape of modal logics, thereby providing the next plateau of observation. But further surprises are no doubt waiting beyond our current horizon of vision. Unlike the mountains of this world, modal ranges are not conquered once and for all: they provide a life-time occupation.

8 References

H. Andréka, I. Németi & J. van Benthem, 1996, 'Modal Logics and Bounded First-Order Fragments', Report ML–96–03, Institute for Logic, Language and Computation, Amsterdam. To appear in *Journal of Philosophical Logic*.

Ph. Balbiani, L. Fariñas del Cerro, D. Vakarelov, 1997, 'Modal Logics for Incidence Geometries', *Journal of Logic and Computation* 7:1, 59-78.

J. Barwise & J. van Benthem, 1996, 'Interpolation, Preservation, and Pebble Games', Report ML–96–12, Institute for Logic, Language and Computation, Amsterdam.

J. van Benthem, 1976, *Modal Correspondence Theory*, doctoral dissertation, Mathematical Institute, University of Amsterdam.

J. van Benthem, 1979, 'Minimal Deontic Logics', *Bulletin of the Section of Logic* 8:1, 36-42.

J. van Benthem, 1984, 'Correspondence Theory', in D. Gabbay and F. Guenthner, eds., 1984, *Handbook of Philosophical Logic, Vol. II.*, Reidel, Dordrecht, 167-247. (Updated version in new 1997 edition of the Handbook.)

J. van Benthem, 1985, *Modal Logic and Classical Logic*, Bibliopolis, Napoli.

J. van Benthem, 1986, *Essays in Logical Semantics*, Reidel, Dordrecht.

J. van Benthem, 1991, *Language in Action.Categories, Lambdas and Dynamic Logic*, North-Holland, Amsterdam. (Second edition with addenda, MIT Press, 1995.)

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford. Distributed by Cambridge University Press, Cambridge.

J. van Benthem, 1997, 'Dynamic Bits and Pieces', Report LP-97-01, Institute for Logic, Language and Computation, University of Amsterdam.

J. van Benthem and J. Bergstra, 1995, 'Logic of Transition Systems', *Journal of Logic, Language and Information* 3:4, 247 - 283.

P. Blackburn, 1990, *Nominal Tense Logic, and Other Sorted Intensional Frameworks*, Ph.D. thesis, Centre for Cognitive Science, University of Edinburgh.

P. Blackburn, M. de Rijke & Y. Venema, 1996, *A Course in Modal Logic*, Department of Computer Science, University of Warwick.

P. Blackburn & J. Seligman, 1995, 'Hybrid Languages', *Journal of Logic, Language and Information* 4:3, 251-272.

E. Börger, E. Grädel & Y. Gurevich, 1996, *The Classical Decision Problem*, Springer, Berlin.

A. Chagrov & M. Zakharyaschev, 1997, *Modal Logic*, Oxford University Press, Oxford.

K. Doets, 1996, *Basic Model Theory*, CSLI Publications, Stanford, (distributed by Cambridge University Press).

D. Gallin, 1975, *Systems of Higher-Order Modal and Intensional Logic*, North-Holland, Amsterdam.

G. Gargov, S. Passy & T. Tinchev, 1987, 'Modal Environment for Boolean Speculations'. In D. Skordev, ed., *Mathematical Logic and its Applications*, Plenum Press, New York, 253-263.

R. Goldblatt, 1987, *Logics of Time and Computation*, CSLI Lecture Notes, Stanford University, Chicago University Press.

V. Goranko, 1990, 'Modal Definability in Enriched Languages', *Notre Dame Journal of Formal Logic* 31, 81–105.

V. Goranko, 1995, 'Hierarchies of Modal and Temporal Logics with Reference Pointers', department of mathematics, Rand Afrikaans University. To appear in *Journal of Logic, Language and Information* .

E. Graedel, 1997, 'On the Restraining Power of Guards', Department of Informatics, Technical University of Aachen.

E. Grädel, P. Kolaitis & M. Vardi, 1997, 'On the Decision Problem for Two-Variable First-Order Logic', Bulletin of Symbolic Logic 3:1, 53-69.

M. Hennessy & R. Milner, 1985, 'Algebraic Laws for Nondeterminism and Concurrency', *Journal of the ACM* 32, 137-161.

N. Kurtonina, 1995, *Frames and Labels. A Modal Analysis of Categorial Deduction*, Ph.D. Thesis, Onderzoeksinstituut voor Taal en Spraak, University of Utrecht & Institute for Logic, Language and Computation, University of Amsterdam.

N. Kurtonina & M. de Rijke, 1997, 'Simulating Without Negation', *Journal of Logic and Computation* 7:4, 501-522,

M. Marx, 1995, *Arrow Logic and Relativized Algebras of Relations*, Ph. D. thesis, CCSOM and ILLC, University of Amsterdam.

M. Marx & Y. Venema, 1996, *Multi-Dimensional Modal Logic*, Kluwer, Dordrecht.

S. Orey, 1959, 'Model Theory for the Higher-Order Predicate Calculus', *Transactions of the American Mathematical Society* 92, 72-84.

D. Park, 1981, 'Concurrency and Automata on Infinite Sequences', *Proceedings 5th GI Conference*, Springer, Berlin, 167-183.

M. de Rijke, 1993, *Extending Modal Logic*, Dissertation, Institute for Logic, Language and Computation, University of Amsterdam.

E. Spaan, 1993, *Complexity of Modal Logics*, dissertation, Institute for Logic, Language and Computation, University of Amsterdam.

D. Vakarelov, 1991, 'A Modal Logic for Similarity Relations in Pawlak Information Systems', *Fundamenta Informaticae* 15, 61–79.

D. Vakarelov, 1992, 'A Modal Theory of Arrows I', Report ML-92-04, Institute for Logic, Language and Computation, University of Amsterdam.

Y. Venema, 1991, *Many-Dimensional Modal Logics,* Dissertation, Institute for Logic, Language and Computation, University of Amsterdam.

Y. Venema, 1995, 'Cylindric Modal Logic', *Journal of Symbolic Logic* 60, 591-623.

Y. Venema, 1996, 'A Crash Course in Arrow Logic', in M. Marx, M. Masuch & L. Pólos, eds., *Arrow Logic and Multi-Modal Logic*, CSLI Publications, Stanford, 3-34, (distribbuted by Cambridge University Press).

F. Wolter & M. Zakharyaschev, 1996, 'Advanced Modal Logic', School of Information Science, JAIST, Kanazawa. To appear in D. Gabbay, ed., *Handbook of Philosophical Logic*, second enlarged edition, Kluwer, Dordrecht.