Andreas Herzig
IRIT, Univ. Toulouse
Toulouse, France
Tel.: +49 56155 8123
www.irit.fr/~Andreas.Herzig

Report on the PhD thesis of Dimiter Todorov Georgiev "Algorithmic Methods for Non-Classical Logics"

Faculty of Mathematics and informatics, Sofia University "Kliment Ohridski"

Actuality of the problem

Dimiter Georgiev's PhD thesis is about correspondence and completeness results for normal modal logics. These logics have modal operators $\Box$ and $\Diamond$ that are interpreted in Kripke frames with a set of possible worlds W and binary accessibility relations r on W. Starting from the basic modal logic K, various modal logics are obtained by imposing constraints on the accessibility relations, such as reflexivity, transitivity, etc. In order to further understand and analyse the validities in specific classes of frames they are then characterised by means of specific axioms. The *correspondence problem* is to obtain, given a modal formula A, a formula in the language of first-order logic with equality and binary predicates r that has the same frames as A. For example, the modal formula $p \to \Diamond p$ corresponds to the first-order formula $\forall x r(x,x)$, and $p \to \Box p$ corresponds to the first-order formula $\forall x, y\ r(x,y) \to x{=}y$. In order to establish such correspondences, a modal axiom such as $p \to \Diamond p$ is first translated into the second-order formula with free variable x

$$\forall p\big(p(x) \to \exists y(r(x,y) \land p(y))\big)$$

and it is then established that the latter is equivalent to the first-order formula $r(x,x)$. The fact that a modal axiom A corresponds with a first-order constraint φ does not imply completeness: it might still happen that a modal formula B is valid in the class of frames defined by φ but cannot be proved from A. Completeness is however guaranteed when A is canonical, i.e., when the canonical model satisfies constraint φ.

These problems are of utmost importance for logic-based artificial intelligence. Indeed, the dominant methodology in that area—sometimes called logic engineering—consists in modelling a domain by isolating its main concepts (such as the concepts knowledge, belief, action, time, obligation, permission, preference, proximity in space, etc.), designing a semantics for these concepts and analysing the mathematical properties of the resulting formalisms. This methodology typically resorts to logics extending classical logic by so-called modal operators, with semantics in terms of Kripke models. One of the most prominent parts of the mathematical analysis is the identification of axioms and the proof of their soundness and completeness w.r.t. the semantics: every formula that is provable in the axiomatics is valid in the semantics, and vice versa. Many such modal logics have been designed in the last 50 years in AI and the problem of establishing correspondence and completeness theorems pops up each time, urging for general results and techniques.

In the early years of modal logic, correspondence and completeness results existed only for very few logics: mainly the so-called basic modal logics whose axioms are among D, T, 4, 5, and B, as exemplified e.g. in the textbooks by Hughes and Cresswell (1968) and Chellas (1980). This was generalised to Lemmon-Scott axioms $\Diamond^k\Box^l p \to \Box^m\Diamond^n p$ and their multi-modal versions, which correspond to generalised confluence constraints on the relation r; see e.g. the 1984 textbook by Hughes and Cresswell. In 1975, Henrik Sahlqvist provided a further generalisation by identifying a class axioms that is named after him. In 1991, Lilia Chagrova showed a fundamental limitation of correspondence theory: she proved that it is undecidable whether a first-order equivalent exists. In 2005 and 2006, Philippe Balbiani and Tinko Tinchev proved that the correspondence problem is however decidable for extensions of modal logic S5. Dimiter's contribution in the present thesis (chapters 4, 5, and 6) is to push Balbiani and Tinchev's approach further: he shows that the problem is also decidable for KD45 ('weak S5'), while it is undecidable for K5.

Shortly after Chagrova's undecidability result, researchers started to look for algorithms solving the correspondence problem. The first was Gabbay and Ohlbach's 1994 SCAN algorithm which uses the standard translation of (the negation of) the modal formula, skolemisation, resolution, and un-skolemisation. The resolution part of the SCAN algorithm does not necessarily terminate and if so, may produce a formula that cannot be un-skolemised. It was only established in 2004 that the SCAN algorithm successfully terminates for all Sahlqvist formulas. Andrzej Szalas and colleagues proposed to replace the resolution principle by Ackermann's Lemma, which provides a way of eliminating second-order existential quantification. Just as the SCAN algorithm, Szalas DLS algorithm uses a translation into predicate logic; contrarily to SCAN it always terminates, but un-skolemisation may fail. Besides a further approach due to Simmons, Conradie and colleagues proposed an approach that avoids translation into predicate logic and instead is based on the use of hybrid logic nominals, which play the role of Skolem functions. Their algorithm is thus able to work within the language of modal logic. Dimiter's 2006 Master thesis was dedicated to its implementation. Dimiter's contribution in the present thesis (Chapter 3) is a deterministic version of the original algorithm.

Short overview of the thesis

Apart from rather short introduction and conclusion chapters, the thesis has 6 main chapters (which are designated as "sections" in the document). Chapter 2 sets the stage: it contains the main definitions and results that are needed. Chapter 3 is the longest chapter of the thesis and contains what is, in my view, its main contribution: a description of deterministic SQEMA together with a proof of its correctness, success for Sahlqvist and inductive formulas, as well as several examples (in particular pre-contact logics). The last three chapters contain three results about the definability problem for logics beyond K5: decidability of both modal definability and first-order definability for KD45 (Chapter 4) and for KD45 plus the universal modality (Chapter 5) and undecidability for K5 (Chapter 6).

In the sequel I will describe the approach and the contributions in more detail.

Evaluation of the main achievements in the thesis

The main contribution of Dimiter's thesis is a deterministic version of Conradie et al.'s originally nondeterministic SQEMA algorithm, described in Chapter 3 of the thesis. The algorithm consists in first negating the axiom and then applying a set of transformation rules with the aim of producing a pure formula: a formula having nominals $c, d, \ldots$, but no propositional variables $p, q, \ldots$. The rules may fail to produce such a formula, in which case it remains unknown whether there is a first-order correspondent; otherwise, if they output a pure formula then the latter can straightforwardly be turned into a frame constraint by means of the standard translation. The transformation rules contain in particular the following:

- the $\Box$-rule which, by replacing conjuncts $A_1 \vee \Box A_2$ by $\Box^{-1} A_1 \vee A_2$, allows to 'shift boxes around' (thereby introducing converse modal operators; this is actually nothing but the conversion rule that is familiar from temporal logics);
- the $\Diamond$-rule which introduces nominals, replacing conjuncts $\neg c \vee \Diamond A$ by $(\neg c \vee \Diamond c') \wedge (\neg c' \vee A)$ where c' is fresh (thus performing a kind of skolemisation);
- a hybrid version of the Ackermann rule which, collecting all conjuncts of the form $\alpha_1 \vee p, \ldots, \alpha_n \vee p$, basically replaces conjuncts β that are negative in p by $\beta[p/\neg\alpha_1 \vee \ldots \vee \neg\alpha_n]$.

Following the original proof due to Conradie et al., Dimiter proves that these rules preserve local frame equivalence and either local equivalence over descriptive frames (if there are no nominals in the original formula) or otherwise local equivalence over discrete frames (Proposition 50; uses Esakia's Lemma). He then introduces an arrangement of these rules which make them a deterministic algorithm and proves that it terminates. Dimiter then discusses some example axioms, including a case where there is a solution (that can be obtained with the SQEMA variant from [18]) but where deterministic SQEMA fails. He finally proves that the algorithm succeeds when executed on modal axioms from the Sahlqvist class as well as from the more general class of inductive formulas. He moreover studies pre-contact logics via a translation to the Sahlqvist class. A final short section briefly describes the implementation of deterministic SQEMA that is available through several mirror webpages.

The second part of the thesis is made up of chapters 4-6. They explore the 'danger zone' of undecidability of the problem of finding first-order formulas corresponding to modal formulas and, the other way round, the problem of finding modal formulas corresponding to first-order formulas. Dimiter's approach is based on earlier work by Philippe Balbiani and Tinko Tinchev who had established decidability results for definability problem for extensions of S5 as well as for extensions of S5 plus universal modality. He succeeds in adapting their technique to KD45: in Chapter 4 he proves for the basic modal language that

- every modal axiom extending KD45 is definable in the language of first-order logic;
- it is decidable in polynomial space whether for a given first-order formula there is a modal formula that is equivalent on KD45 frames.

The proofs use that KD45 frames can be reduced to what Dimiter calls daisies with one petal: frames with 'stamen worlds' that can see each other and a 'petal world' that can see all stamen worlds. Furthermore, it uses that when a daisy frame validates a modal formula A then every daisy frame with less worlds validates A, too. Dimiter shows that the modal axioms extending KD45 are basically

$$\Big(\bigwedge_{1\leq i\leq n} \Diamond p_i\Big) \rightarrow \bigvee_{1\leq i<j\leq n} \Diamond(p_i \wedge p_j)$$

expressing that there are less than n stamen worlds. In Chapter 5 Dimiter generalises these results to the extension of the basic modal language by the universal modality. Both chapters also contain PSPACE completeness complexity results for the problem of modal definability of first-order formulas.

In the last chapter, Chapter 6, Dimiter finally focusses on extensions of the modal logic K5, a modal logic that is slightly weaker than KD45. He again proves two definability results:

- just as for KD45, every modal axiom in the basic language is definable in the language of first-order logic;
- it is undecidable whether a first-order formula has a corresponding modal axiom (that is equivalent on the class of K5 frames).

This is unpublished work in collaboration with Philippe Balbiani and Tinko Tinchev.

Critical remarks

The thesis is well and carefully written: all the concepts that are used are defined precisely, the proofs are rigorous and correct as far as I was able to check, and I spotted only very few typos and other minor issues (that I will send separately to the author). However, the exposition is rather compact and minimalistic. This is mainly due to a presentation that follows a 'definition-lemma-theorem' schema, without too much explanations in between: it would be nice if at least some of the definitions and propositions were illustrated by simple examples. It would also be good if acronyms such as SQEMA were spelled out somewhere.

The presentation of the deterministic SQEMA algorithm in Chapter 3 is quite involved. This is clearly due to its non-trivial nature; however, at least in part it is also due to the effort to keep rule application deterministic. As a consequence, the termination proof in this chapter is perhaps more complicated than they could be. I wondered whether rewriting theory with concepts such as confluence of a Noetherian system of rewriting rules could make things easier to present.

While the chapters 4, 5, 6 study the complexity of the respective decision problems in detail, the complexity of the deterministic SQEMA algorithm of Chapter 3 is not mentioned. In [17] it is stated that the nondeterministic SQEMA runs in nondeterministic polynomial time and it is conjectured that a polynomial time complexity could be established by means of "additional rules which determine the right order of elimination (if any)". Given that the deterministic SQEMA as defined in Section 3.5 uses backtracking, I suppose that the present thesis only supports this claim for the Sahlqvist and the inductive class (where no backtracking is needed, see corollaries 86and 91), but not in the general case; In any case, it would be good to discuss this issue.

The thesis is a bit short on related work. As to Chapter 3, what are the advantages of SQEMA over SCAN and DLA? What are the advantages of deterministic SQEMA over the original, non-deterministic SQEMA? It is mentioned that deterministic SQEMA performs better than the original SQEMA at least in the case of formula 3.6.2: on page 51 it is said that "the classical SQEMA fails" for that formula. This (and potentially other examples such as 3.6.4 on page 53, or the explanations on top of page 59) might be used to argue for the approach in the thesis and the argument could, if possible, be worked out in more detail. As to chapters 4 and 5, to which extent do the techniques from Balbiani and Tinchev's paper transfer?

The conclusion is presently rather short. It would gain from a more systematic exposition of research avenues and possible future work that are discussed a bit hidden here and there throughout the chapters. For example, when deterministic SQEMA is run on Formula 3.12.2 (page 81) it is observed that "[i]n the future, the author would like to explore a matching definition of inductive formulas in the PCL language". Similarly, at Formula 3.6.4 where deterministic SQEMA fails, the integration of a trick from [19] that guarantees success is mentioned.

Finally, the layout of the document could be improved by having chapters instead of sections; Latex environments such as 'itemize', 'enumerate', and 'tabular' would be useful to structure the definitions and other text in many places (to witness, Definition 79 is a single sentence running over 14 lines); tables such as those on page 44 and 45 should be provided with a number and caption. pages could have a header with chapter title.

Correspondence between thesis and author's abstract

The abstract duly summarises Dimiter's contributions, pointing out the original parts of his work and mentioning the (ongoing or accomplished) publications corresponding to each chapter.

Summary

Overall, I believe this is a very good PhD thesis that provides a significant contribution to the literature on modal logics. I would like to highlight that the SQEMA algorithm has been implemented in an automatic tool that is freely accessible to the community through a website[1]. The description of that substantial piece of work only takes two pages in the thesis; however, I believe this to be a highly valuable service to logic-based AI that is easily under-estimated. It moreover appears from Chapter 3 that the implementation helped the author to test the properties of his algorithm.

To sum it up, I am happy to recommend the defense of Dimiter's thesis without any hesitation.

Andreas Herzig
Directeur de recherche CNRS
IRIT, Université de Toulouse

[1] `http://www.fmi.uni-sofia.bg/fmi/logic/sqema` and several mirror sites