In: Abstracts of the 3rd Logical Biennal, Summer School & Conference in honour of S. C. Kleene, Varna, Bulgaria, June 1990

A NEW DECIDABLE FRAGMENT OF FIRST ORDER LOGIC

Andreas Herzig

IRIT - LSI, Université Paul Sabatier, F-31062 Toulouse Cédex
uucp : herzig@irit.fr
Tel.: (33) 61 55 67 64

Abstract

We define ordered formulas of first order logic as formulas where the order of the arguments of a predicate corresponds to that of the quantifiers governing this occurrence. We prove that ordered formulas are decidable by translating them into propositional modal logic Q.

INTRODUCTION

There are several decidable fragments of first-order logic. Two of them are well-known in literature. The first is monadic first-order logic ([Löwenheim 15], [Kleene 52]), i.e. only one-place predicates are allowed. In [Kripke 62], S. Kripke gave a translation of modal logic into monadic first-order logic plus one binary predicate expressing the properties of the accessibility relation of the possible-world models. Thus the resulting fragment is decidable, although monadic first-order logic together with one dyadic predicate is generally undecidable. Just as we will proceed, he used a bijection between a decidable propositional modal logic and the fragment. The second decidable fragment is due to Th. Skolem ([Skolem 20], [Kleene 52]) and consists of first order formulas without existential quantifiers in the scope of universal quantifiers. This result can be viewed as a corollary of the completeness of skolemization: In the case of this fragment, existentially quantified variables are replaced by skolem constants, entailing a finite Herbrand domain of the formula under concern.

The decidable fragment we present in this note neither contains one of these fragments, nor is it contained in them. There is no restriction on the number of arguments of a predicate, but the formula must be closed, and the ordering of the quantifiers must be that of the variables in the predicates they govern. In other words, for every occurence of a predicate $p(x_1,\ldots,x_n)$ and every i, $1 \leq i \leq n$, the existential or universal quantifier $Q_i x_i$ binding x_i must be exactly in the scope of the quantifiers $Q_1 x_1, \ldots, Q_{i-1} x_{i-1}$. A translation to propositional modal logic Q will ensure the decidability of the fragment. To achieve the proof we introduce deterministic modal logic DQ and define quantification over modal operators.

The note is divided into 5 sections. In section 1 we recall first-order logic and define the fragment of ordered formulas, and in section 2 we introduce the deterministic modal logic DQ. In section 3 we translate first-order logic to DQ. In section 4 we define a fragment of DQ containing the translated fragment of first-order logic, and in section 5 we show the latter to be decidable by translating it to propositional modal logic Q.

1 FIRST ORDER LOGIC

A **first-order language** is built on the following basic symbols :

- a set of variables VAR = {x, y, z, ...}
- for every $n \geq 0$, a set of n-ary predicate symbols $PRED^n$ = {p, q, r, ...}
- conjunction $\wedge$, negation $\neg$, universal quantifier $\forall$, and parantheses (,)

First order formulas are defined recursively as the least set such that

- $p(x_1,...,x_n)$ is a formula if $p \in PRED^n$
- $(\neg A)$ is a formula if A is a formula
- $(A \wedge B)$ is a formula if A and B are formulas
- $(\forall x\ A)$ is a formula if $x \in$ VAR and A is a formula

We introduce the existential quantifiers by defining $\exists x\ A$ as an abbreviation for $\neg \forall x \neg A$. We suppose the usual conventions and notions of scope, free and bound variables, subformula and closed formulas ([Andrews 86]). When we do not bother whether a quantifier is existential or universal, we note Qx A both $\forall x\ A$ and $\exists x\ A$.

A **model** of first-order logic is of the form N = (D,j), where

- D is a set called domain
- j is an valuation function mapping VAR into D and $PRED^n$ into D^n, for every $n \geq 0$

For a given first-order model N = (D,j),

$$N(x) = \{N' : N' = (D, j'), j'(p) = j(p) \text{ for every } n \geq 0 \text{ and } p \in PRED^n, j'(y) = j(y) \text{ for every } y \in VAR - \{x\}\}$$

is called the set of models **agreeing with N off x.**

Satisfaction is defined as usual, with

- N sat $p(x_1,...,x_n)$ if $(j(x_1),...,j(x_n)) \in j(p)$
- N sat $\forall x\ A$ if for every $N' \in N(x)$, N' sat A

A given formula A is **satisfiable** in first-order logic iff there is a first order model N such that N sat A, and A is **valid** in first-order logic iff for every first order model N we have N sat A.

Now we define a particular set of formulas of first-order logic as the set of closed formulas in which every argument of an atom is exactly in the scope of the quantifiers binding the arguments left of it. More formally, a first-order formula A is **ordered** iff A is closed and for every atom $p(x_1,...,x_n)$ occurring in A and for every x_i, $1 \leq i \leq n$, the quantifier $Q_i x_i$ binding x_i is in the scope of some quantifier Qy iff Qy is binding some x_j in $p(x_1,...,x_n)$, and $1 \leq j < i$.

Examples of ordered formulas are

$$\forall x\ (\ p(x) \wedge \exists y\ q(x,y)\ ),$$
$$\forall x\ \forall x'\ \exists u\ p(x,x',u)\ \wedge\ \forall y\ \exists y'\ \forall z\ \neg p(y,y',z).$$

Examples of formulas which are not ordered are

$$\forall x\ p(x,x),$$
$$\forall x\ \exists y\ (\ p(y) \wedge q(x,y)\ ),$$
$$\forall x\ \forall y\ (\ p(x,y) \wedge p(y,x)\ ).$$

Note that e.g. the non-ordered formula $Qy\ Q'x\ p(x,y)$ can be transformed into an ordered one by permuting the quantifiers Q and Q' if they are of the same type (i.e. $Q = Q'$). We may also permute the order of the arguments of a given predicate p in an appropriate manner and introduce thus a new predicate p' replacing p everywhere. Sometimes it is also possible to drop quantifiers in order to obtain ordered formulas. E.g. omitting $\forall y$ from the non-ordered formula

$$\forall x\ \exists x'\ p(x,x')\ \wedge\ \forall y\ \exists y'\ \forall z\ \neg p(y',z),$$

we obtain the equivalent ordered formula

$$\forall x\ \exists x'\ p(x,x')\ \wedge\ \exists y'\ \forall z\ \neg p(y',z).$$

Our aim is to show that ordered formulas of first-order logic are decidable.

2 DETERMINISTIC MODAL LOGIC DQ

In this section we introduce a particular multi-modal deterministic logic DQ which is similar to dynamic logic for deterministic programs ([Harel 79]). Propositional modal logic Q is contained in DQ. Moreover, we allow quantification over deterministic modal operators.

A **language** of DQ is built on the following basic symbols:

- a set of variables $VAR = \{x, y, z, \dots\}$
- a set of propositional variables $PROP = \{p, q, r, \dots\}$
- conjunction $\wedge$, negation $\neg$, and universal quantifier $\forall$
- parantheses (,) and brackets [,]

Formulas of DQ are defined recursively as the least set such that

- p is a formula if $p \in PROP$
- $(\neg A)$ is a formula if A is a formula
- $(A \wedge B)$ is a formula if A and B are formulas
- $(\forall x\ A)$ is a formula if A is a formula and $x \in VAR$
- $([x]\ A)$ is a formula if A is a formula and $x \in VAR$

We employ the usual conventions and notations. For every $x \in VAR$, $[x]$ is called a **modal operator**.

A **model** of DQ is of the form $M = (W, w_0, S, i)$, where

- W is a set called the set of worlds
- $w_0 \in W$ is a world called the actual world
- S is a set of total functions on W called the set of accessibility functions
- i is an valuation function mapping VAR into S and PROP into 2^W.

Just as the accessibility relations of the Kripke models for Q are all deterministic and serial, i.e. for every $w \in W$ there is exactly one $w' \in W$ such that $(w,w') \in R$ ([Hughes & Cresswell 86]), we may present the Kripke semantics in terms of total functions.

For a given DQ-model $M = (W, w_0, S, i)$,

$$M(x) = \{M' : M' = (W, w_0, S, i'),\ i'(p) = i(p) \text{ for every } p \in \text{PROP},\ i'(y) = i(y) \text{ for every } y \in \text{VAR} - \{x\}\}$$

is called the set of models **agreeing with M off x.**

Satisfaction is defined as usual for the classical connectors, plus:

- M,w sat p if $w \in i(p)$
- M,w sat $\forall x$ A if for every $M' \in M(x)$, M' sat A
- M,w sat [x] A if M,i(x)(w) sat A

A given formula A is **satisfiable** in DQ iff there is a DQ-model $M = (W, w_0, S, i)$ such that M, w_0 sat A, and A is **valid** in DQ iff for every DQ-model $M = (W, w_0, S, i)$ we have M, w_0 sat A.

Normal-Form-Lemma. The following formulas of DQ are valid.

$$[x]\,A \wedge [x]\,B \leftrightarrow [x]\,(A \wedge B)$$
$$\neg [x]\,A \leftrightarrow [x] \neg A$$
$$\forall x\,[y]\,A \leftrightarrow [y]\,\forall x\,A \quad \text{if } x \neq y$$
$$A \wedge \forall x\,B \leftrightarrow \forall x\,(A \wedge B) \quad \text{if x does not occur in A}$$
$$A \wedge \exists x\,B \leftrightarrow \exists x\,(A \wedge B) \quad \text{if x does not occur in A}$$

Other deterministic logics are a recent subject of research. They can be defined by closure conditions or other restrictions on the set of accessibility functions S. It may be closed by function composition, or contain the identity function ([Ohlbach 88], [Fariñas & Herzig 88], [Auffray & Enjalbert 88], [Auffray 89], [Herzig 89]). It has been shown that powerful normal form theorems can be obtained for deterministic logics: by the above equivalences, deterministic modal operators can cross the classical connectors and may thus be shifted into positions where there is no more classical connector in their scope. Thus, at the end modal operators only occur in what can be called modal literals, i.e. literals prefixed by sequences of modal operators. Calling 'clause' a disjunction of modal literals, every formula of deterministic logics can be rewritten as a conjunction of clauses. Now e.g. a resolution principle can be defined ([Robinson 65]), using appropriated algorithms to find the unifiers of two sequences of modal operators.

3 EMBEDDING FIRST-ORDER LOGIC INTO DQ

In this section we shall translate first-order logic into DQ. The translation function $trans_1$ is defined as follows.

- $trans_1(p(x_1,\ldots,x_n)) = [x_1]\ldots[x_n]\, p$
- $trans_1(A \wedge B) = trans_1(A) \wedge trans_1(B)$
- $trans_1(\neg A) = \neg\, trans_1(A)$
- $trans_1(\forall x\, A) = \forall x\, trans_1(A)$

$trans_1$-Lemma. A a first-order formula A is first-order-satisfiable iff $trans_1(A)$ is DQ-satisfiable.

Proof. First we augment the language of DQ in order to englobe first-order-language by the following basic symbols:

- for every $n \geq 0$, a set of n-ary predicate symbols $PRED^n$, $PRED^n = PROP$

The formulas of the extended language are constructed as before, plus

- $p(x_1,\ldots,x_n)$ is a formula if $p \in PRED^n$

Given a model $M = (W,w_0,S,i)$ we extend the valuation function i such that

$$i : PRED^n \rightarrow S^n$$
$$i(p) = \{ (s_1, \ldots ,s_n) : i(s_n)\circ\ldots i(s_1)(w_0) \in i(p) \} \quad \text{for } p \in PRED^n$$

Then the equivalences

$$trans_1(p(x_1,\ldots,x_n)) \leftrightarrow [x_1]\ldots[x_n]\, p \leftrightarrow p(x_1,\ldots,x_n)$$

are DQ-valid, and hence $A \leftrightarrow trans_1(A)$ for every first-order formula A, too.

It remains to show that for every first-order formula A, A is first-order-satisfiable iff A is DQ-satisfiable.
Given a DQ-model $M = (W,w_0,S,i)$ satisfying A, it is easy to show that $N = (S,j)$ is a first-order model satisfying A iff j is i restricted to PRED and VAR.
In the opposite sense, to a given first-order model $N = (D,j)$ we associate a DQ-model $M = (D^*, nil,D,i)$ such that

- $D^* = \{D^n : n \geq 0\}$ is the set of lists over D
- nil is the empty list
- every $d \in D$ is considered to be an application from D^* into D^* such that $d((d_1, \ldots ,d_n)) = (d_1, \ldots ,d_n,d)$
- The restriction of i to VAR and PRED is j, and for every $p \in PROP^n$, $i(p) = \{ (d_1, \ldots ,d_n) : (d_1, \ldots ,d_n) \in i(p) \}$.

M is a DQ-model, and M,nil sat $p(x_1,\ldots,x_n)$ iff N sat $p(x_1,\ldots,x_n)$ for every atom $p(x_1,\ldots,x_n)$. Consequently M,nil sat A iff N sat A for every first-order formula A.

4 ORDERED AND STRONGLY ORDERED DQ-FORMULAS

In the first section we have defined ordered formulas of first-order logic. In this section we shall define the corresponding property in DQ, and we show that ordered formulas of DQ can be transformed into formulas having a more restricted form, close to the traditional modal language. First, we need an auxiliary definition.

We introduce a relation **balances** between variable lists and formulas. It is defined recursively as the least relation such that

- $(x_1,\ldots,x_n)$ balances p if $p \in$ PROP and $n = 0$
- $(x_1,\ldots,x_n)$ balances $B \wedge C$ if $(x_1,\ldots,x_n)$ balances B and $(x_1,\ldots,x_n)$ balances C
- $(x_1,\ldots,x_n)$ balances $\neg B$ if $(x_1,\ldots,x_n)$ balances B
- $(x_1,\ldots,x_n)$ balances $Qy\ B$ if $(x_1,\ldots,x_n,y)$ balances B
- $(x_1,\ldots,x_n)$ balances $[y]\ B$ if $y = x_1$ and $(x_2,\ldots,x_n)$ balances B

Now a DQ-formula A is **ordered** iff A is closed, and the empty list (noted "nil") balances A.

Examples of ordered DQ-formulas are

$$\forall x\ (\ [x]\ p \wedge \exists y\ [x]\ [y]\ q\ )$$
$$\forall x\ (\ [x]\ p \wedge [x]\ \exists y\ [y]\ q\ )$$
$$\forall x\ [x]\ (\ p \wedge \exists y\ [y]\ q\ )$$

Fact. Let A a first-order formula. If A is ordered then $\text{trans}_1(A)$ is ordered.

Let A a DQ-formula. A is **strongly ordered** iff A is closed, and for every subformula of A of the form $Qx\ B$, B is of the form $[x]\ C$ and x does not occur in C.

E.g. the last formula of the preceeding examples is strongly ordered, whereas the first two are not.

Fact. Strongly ordered DQ-formulas are ordered.

In the rest of the section we shall define a translation from ordered DQ-formulas into strongly ordered DQ-formulas and prove their equivalence. First we need an auxiliary definition.

We define recursively a function **factor** mapping a variable list and a formula into strongly ordered formula as follows:

- $\text{factor}(\ (x_1,\ldots,x_n)\ ,\ p) = p$ if $p \in$ PROP and $n = 0$
- $\text{factor}(\ (x_1,\ldots,x_n),\ B \wedge C) = \text{factor}(\ (x_1,\ldots,x_n),\ B) \wedge \text{factor}(\ (x_1,\ldots,x_n),\ C)$
- $\text{factor}(\ (x_1,\ldots,x_n),\ \neg B) = \neg\ \text{factor}(\ (x_1,\ldots,x_n),\ B)$
- $\text{factor}(\ (x_1,\ldots,x_n),\ Qy\ B) = \text{factor}(\ (x_1,\ldots,x_n,y),\ B)$
- $\text{factor}(\ (x_1,\ldots,x_n),\ [y]\ B) = \text{factor}(\ (x_2,\ldots,x_n),\ B)$ if $y = x_1$

Now we define a function **$trans_2$** translating ordered DQ-formulas into strongly ordered DQ-formulas by:

$$trans_2(A) = factor(nil,A)$$

$trans_2$-Lemma. Let A an ordered DQ-formula. Then $trans_2(A)$ is strongly ordered and equivalent to A.
Proof. By induction on the structure of A it is possible to establish that for every variable list $(x_1,\ldots,x_n)$ balancing A we have that

- factor($(x_1,\ldots,x_n)$, A) is strongly ordered
- $[x_1]\ldots[x_n]$ factor($(x_1,\ldots,x_n)$, A) $\leftrightarrow$ A

by induction on the structure of A, using for each of the cases in the induction step one of the corresponding equivalences of the Normal-Form-Lemma.

5 FROM STRONGLY ORDERED FORMULAS OF DQ TO MODAL LOGIC Q

Traditionally, modalities are expressed by means of the well-known operators [] and <>. We are going to show that the propositional modal logic Q corresponds to the fragment of strongly ordered formulas of DQ. The operators [] and <> can be expressed in DQ by replacing [] A by $\forall x$ [x] A and <>A by $\exists x$ [x] A. Now there is an immediate way to establish a semantical correspondency: It suffices to consider the traditional Kripkean accessibility relation to be expressed in terms of a set of accessibility functions, and it is easy to see that [] A and $\forall x$ [x] A are interpreted in the same manner.

First, we add [] as a basic symbol to the language of DQ. **Formulas** of the extended language are defined as before, plus

- [] A is a formula if A is a formula

As usual, <> A is defined as an abbreviation for $\neg$ [] $\neg$A. Thus, the formulas of propositional Q are just those where neither deterministic modal operators nor quantifiers occur. Now the set of accessibility functions S of a DQ-model M can be interpreted as a relation in the sense that $(w,w')\in S$ iff there is $s \in S$ such that $s(w) = w'$.

Hence we are able to define **satisfaction** as

- M,w sat [] A if for every $w' \in W$ such that $(w,w') \in S$, M,w' sat A
- M,w sat <> A if there is $w' \in W$ such that $(w,w') \in S$ and M,w' sat A

Fact. The formulas

$$[]\, A \leftrightarrow \forall x\, [x]\, A$$
$$<>\, A \leftrightarrow \exists x\, [x]\, A$$

are valid in extended DQ-models.

An accessibility relation R of a Kripke model of Q must be serial, i.e. for every $w \in W$ there is some $w' \in W$ such that $w R w'$ ([Hughes & Cresswell 86]). Hence there can be defined a set of total functions S forming a partition of R, in the sense that

$$R = \{ (w,w') : \text{there is } s \in S \text{ such that } w' = s(w) \},$$

and converse, and there is equivalence with respect to satisfaction between the traditional Kripke models of Q and the models of DQ. Consequently we have the following fact.

Fact. Let A a formula of propositional modal logic Q. A is satisfiable in a Kripke model with serial accessibility relation iff A is DQ-satisfiable.

Now we may define a translation $trans_3$ of strictly ordered formulas of DQ to formulas of propositional Q as follows.

- $trans_3(p) = p$ for every $p \in PROP$
- $trans_3(\neg A) = \neg\, trans_3(A)$
- $trans_3(A \wedge B) = trans_3(A) \wedge trans_3(B)$
- $trans_3(\forall x\, [x]\, A) = []\, trans_3(A)$
- $trans_3\ (\exists x\, [x]\, A) = <>\, trans_3(A)$

Thus, the translation $trans_3$ consists in replacing $\forall x\, [x]$ by $[]$ and $\exists x\, [x]$ by $<>$.

$trans_3$-Lemma. Let A a strongly ordered formula of DQ. Then $A \leftrightarrow trans_3(A)$ is valid in extended DQ-models.
Proof. The proof is obtained by induction on the number of steps done in the translation, using the fact that $[]\, A \leftrightarrow \forall x\, [x]\, A$ and $<>\, A \leftrightarrow \exists x\, [x]\, A$ are valid.

For example, $\forall x\, [x]\, (p \wedge \exists y\, [y] q)$ is translated to $[]\, (p \wedge <>\, q)$.

Decidability Theorem. Let A an ordered formula of first-order logic. Then A can be translated into a formula D of propositional modal logic Q such that A is satisfiable in first-order logic iff D is satisfiable in Q.
Proof. Let A an ordered first-order formula. Let $B = trans_1(A)$. By the $trans_1$-Lemma of section 3 and the fact of section 4, B is an ordered DQ-formula such that A is satisfiable in first-order logic iff B is satisfiable in DQ. Let $C = trans_2(B)$. By the $trans_2$-Lemma of section 4, C is a strongly ordered formula of DQ such that B is equivalent to C. Now by the previous $trans_3$-Lemma, C is is satisfiable in DQ iff $trans_3(C)$ is is satisfiable in Q.

CONCLUSION

In this note, we have proven the decidability of ordered formulas of first-order logic. Our method to decide whether a given ordered formula A is satisfiable can be summarized as follows: First translate A into a formula B of deterministic modal logic DQ. Second, transform B into a strongly ordered formula C of DQ. Third, translate C into a formula D of propositional modal logic Q. Finally, apply to D some decision procedure for Q.

ACKNOWLEDGEMENTS

I wish to thank Philippe Balbiani, Pierre Bieber, Luis Fariñas del Cerro, Gabriella Crocco and Christel Seguin for their useful comments.

REFERENCES

[Andrews 86] Peter B. Andrews, An Introduction to Mathematical Logic and Type Theory: To Truth through Proof. Academic Press, 1986.

[Auffray & Enjalbert 88] Yves Auffray et Patrice Enjalbert, Démonstration de théorèmes en logique modale - Un point de vue equationnel. *Journées Européennes Logique et Intelligence Artificielle*, Roscoff, 1988.

[Auffray 89] Yves Auffray, Résolution modale et logique des chemins. *Ph.D. thesis, Université de Caen*, 1989.

[Fariñas & Herzig 88] Luis Fariñas del Cerro et Andreas Herzig, Quantified Modal Logic and Unification Theory.- *Rapport LSI n°293*, jan. 1988.

[Fariñas & Herzig 89] Luis Fariñas del Cerro et Andreas Herzig, Automated Quantified Modal Logic. *Machine Learning, Meta Reasoning and Logics* (eds. P. Brazdil and K. Konolige), Kluwer Academic Press, 1990.

[Harel 79] David Harel, Dynamic Logic. *Handbook of Philosophical Logic* (eds. Dov Gabbay et Franz Günthner), D. Reidel, 1984.

[Herzig 89] Andreas Herzig, Raisonnement automatique en logique modale et algorithmes d'unification. *Ph.D. thesis, Université Paul Sabatier*, 1989.

[Hughes & Cresswell 86] G. Hughes et M.J. Cresswell, A Companion to Modal Logic. Methuen & Co. Ltd., London 1986.

[Kleene 52] S. C. Kleene, Introduction to Metamathematics. Van Norstrand, New York, 1952.

[Kripke 62] Saul Kripke, The undecidability of monadic modal quantification theory. *ZML vol. 8*, 1962, 113-116.

[Löwenheim 15] Leopold Löwenheim, Über Möglichkeiten im Relativkalkül. *Math. Ann., vol. 76*, pp. 447-470.

[Robinson 65] J.A. Robinson, A machine oriented logic based on the resolution principle. *JACM 12*, 1965, 23-41.

[Skolem 20] Thoralf Skolem, Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen. *Skrifter utgit av Videnskapsselskapet i Kristiania, I. Matematisk-naturvidens- kabelig klasse 1919*, no. 3, 37 pp.