Guarded Fragments of First-Order Logic: A Perspective for New Description Logics?

Erich Grädel
Mathematische Grundlagen der Informatik
RWTH Aachen
D-52056 Aachen
Email: graedel@informatik.rwth-aachen.de

It is well-known that description logics can be viewed as syntactic variants of propositional modal logics, and that they can be embedded into small fragments of first-order logic, or into extensions of such fragments by appropriate features like counting quantifiers, fixed-point operators etc.

Indeed an interpretation of atomic concepts and roles for a description logic is just a Kripke structure or, equivalently, a first-order structure with only unary and binary predicates. The translation of, say, the description logic $\mathcal{ALC}$ into first-order logic maps any concept C to a first-order formula $C^*(x)$ with one free-variable such that for all $\mathcal{ALC}$-interpretations $\mathcal{K}$ and all objects w in the domain of $\mathcal{K}$

$$w \in C^{\mathcal{K}} \text{ if and only if } \mathcal{K} \models C^*(w).$$

This translation takes any atomic concept A to the atomic formula Ax, commutes with the Boolean connectives, i.e.

$$\begin{aligned} (\neg C)^*(x) &:= \neg C^*(x) \\ (C \sqcap D)^*(x) &:= C^*(x) \wedge D^*(x) \end{aligned}$$

and translates the application of roles as follows:

$$\begin{aligned} (\exists R.C)^*(x) &:= \exists y(Rxy \wedge C^*(y)) \\ (\forall R.C)^*(x) &:= \forall y(Rxy \rightarrow C^*(y)). \end{aligned}$$

(Here $C^*(y)$ is obtained from $C^*(x)$ by replacing all occurences of x by y and vice versa).

Clearly, there is an equivalent translation of propositional modal logic into first-order logic. The image of $\mathcal{ALC}$ under this translation is called the *modal fragment* of first-order logic. It has turned out that the modal fragment has very interesting and useful *algorithmic* and *model-theoretic* properties. On the other side, the modal fragment is a rather small part of first-order logic. It is properly contained in FO^2, relational first-order logic with only two variables and the strictness of this inclusion can be pinned to a number of different restrictions on the modal fragment: It does not have equality, it does not have global quantification over objects and it does not have mechanisms for defining new binary predicates (i.e. new roles), not even at the quantifier-free level.

On the other side, although FO^2 is decidable for satisfiability and has the finite model property (see [10, 7]), it lacks the nice model-theoretic properties and the robust decidability of modal and description logics (see e.g. [1, 9, 8, 11]). The embedding in FO^2 therefore does not explain the good properties of description logics. Also with respect to complexity, description logics are simpler than FO^2: while the satisfiablility problems for the basic description logics are in PSPACE and in some cases EXPTIME, the satisfiability problem for FO^2 is NEXPTIME-complete [7].

A closer look at the translation of concepts into first-order formulae reveals that the quantifiers are used only in a very restricted way, and it has been suspected that this may be the real reason for the good properties of the modal fragments. To investigate this question Andréka, van Benthem and Németi [1] have introduced the *guarded fragment* of first-order logic. They dropped the restriction to use only two variables and only monadic and binary predicates, but imposed that all quantifiers must be relatived (or 'guarded') by atomic formulae.

Definition. The *guarded fragment* GF of first-order logic is defined by induction as follows:

(1) Every relational atomic formula belongs to GF.

(2) GF is closed under propositional connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$.

(3) If $\boldsymbol{x}, \boldsymbol{y}$ are tuples of variables, $\alpha(\boldsymbol{x}, \boldsymbol{y})$ is atomic and $\psi(\boldsymbol{x}, \boldsymbol{y})$ is a formula in GF, such that $\mathrm{free}(\psi) \subseteq \mathrm{free}(\alpha) = \{\boldsymbol{x}, \boldsymbol{y}\}$, then the formulae

$$\begin{gathered} \exists \boldsymbol{y}(\alpha(\boldsymbol{x}, \boldsymbol{y}) \wedge \psi(\boldsymbol{x}, \boldsymbol{y})) \\ \forall \boldsymbol{y}(\alpha(\boldsymbol{x}, \boldsymbol{y}) \rightarrow \psi(\boldsymbol{x}, \boldsymbol{y})) \end{gathered}$$

belong to GF.

The atom $\alpha(\boldsymbol{x}, \boldsymbol{y})$ that relativizes a quantifier as in rule (3) is the *guard* of the quantifier. Note that the guard must contain all the free variables of the formula that follows. Formulae in GF are called *guarded formulae.*

While the guarded fragment clearly contains the modal fragment of first-order logic, it seems not expressive enough for temporal logic. Indeed the straightforward translation of (ψ **until** φ) into first-order logic is

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

which is not guarded in the sense of the previous definition. However, the quantifier $\forall z$ in this formula is guarded in a weaker sense, which lead van Benthem [4] to the following generalization of GF.

Definition. The *loosely guarded fragment* LGF is is defined similarly to GF, but the quantifier-rule is relaxed as follows:

(3)' If $\psi(\boldsymbol{x}, \boldsymbol{y})$ is in LGF, and $\alpha_1 \wedge \cdots \alpha_m$ is a conjunction of atoms, then

$$\exists \boldsymbol{y}((\alpha_1 \wedge \cdots \wedge \alpha_m) \wedge \psi(\boldsymbol{x}, \boldsymbol{y}))$$
$$\forall \boldsymbol{y}((\alpha_1 \wedge \cdots \wedge \alpha_m)) \rightarrow \psi(\boldsymbol{x}, \boldsymbol{y}))$$

belong to LGF, provided that every variable y_i coexists with every other variable of $\{\boldsymbol{x}, \boldsymbol{y}\}$ in at least one of the atoms α_j.

The guarded fragments turn out to have interesting properties [1, 6]:

(1) The satisfiability problem for GF and LGF is decidable.

(2) GF has the finite model property, i.e. every satisfiable formula in the guarded fragment also has a finite model.

(3) Many important model theoretic properties (like the interpolation property, the Beth definability property, the Łos-Tarski property etc.) which hold for first-order logic and the modal fragment, but not, say, for the bounded-variable fragments FO^k, do hold also for the guarded fragments.

(4) Both GF and LGF satisfy a generalized variant of the tree model property. Indeed, every satisfiable guarded formula has a model of small tree-width.

(5) The notion of indistinguishability by guarded formulae can be characterized by natural generalizations of bisimulation.

The guarded fragment thus displays properties which are rather similar to the modal fragment. An advantage of the guarded fragment with respect to other decidable fragments of first-order logic (see [5]) is that the usual restrictions of the latter on the number and arity of the relation symbols or on the quantifier pattern are avoided.

It thus is interesting to determine the power of the guarded fragment, or to put it differently: how much do the guards restrict expressiveness? We give a complexity theoretic answer by showing that the satisfiability problems for both GF and LGF are complete for deterministic double exponential time.

The upper complexity bound follows from a new decidability proof, which is based on extension properties of atomic types and also establishes the *tree-model property* for LGF and GF. The lower bound proof is based on the construction of a family of polynomial-size guarded formulae that force their models to contain a binary trees of double exponential depth.

Of course, double exponential complexity is very high and beyond practical possibilities. However, the reason for this high complexity is just the fact that there is no bound on the arity of the predicates that may occur in guarded formulae. Even if we have just one predicate of arity n over a domain of just two objects then there are 2^{2^n} possible interpretations. In practical applications, the arity of the predicates is usually bounded (recall that in description logics, only unary and binary predicates are used). We can prove that the bounded-variable or bounded-arity subfragments of GF and LGF have deterministic exponential time complexity, a complexity level we have to live with even for some rather modest extensions of $\mathcal{ALC}$. And indeed, we can often live with this level of complexity since it is in terms of the length of the formulae (the concept definitions), and in most practical application the formulae tend be rather small compared to the often huge size of the structures or knowledge bases.

Given that guarded formulae behave very much like modal formulae, established techniques in the fields of modal logic and description logics are applicable in this framework too. In particular this is the case for automata-theoretic techniques. I therefore believe that due to their nice syntactic, model-theoretic and algorithmic properties, the guarded fragments can be a basis for constructing new families of description logics and practical systems that are not limited to unary and binary predicates, that admit more flexibility in the definition of new roles while retaining most of the good features of known description logics.

References

[1] H. Andréka, J. van Benthem, and I. Németi, *Modal Languages and Bounded Fragments of Predicate Logic*, ILLC Research Report ML-96-03 (1996), 59 pages.

[2] J. van Benthem, *Modal Logic an Classical Logic*, Bibliopolis, Napoli 1983.

[3] J. van Benthem, *Exploring Logical Dynamics*, CSLI-Publications, Stanford 1996.

[4] J. van Benthem, *Dynamic bits and pieces*, ILLC Research Report (1997).

[5] E. Börger, E. Grädel, and Y. Gurevich, *The Classical Decision Problem*, Springer, 1997.

[6] E. Grädel, *On the restraining power of guards*, submitted to the Journal of Symbolic Logic.

[7] E. Grädel, P. Kolaitis, and M. Vardi, *On the decision problem for two-variable first-order logic*, Bulletin of Symbolic Logic 3 (1997), 53–69.

[8] E. Grädel and M. Otto, *On logics with two variables*, to appear in Theoretical Computer Science.

[9] E. Grädel, M. Otto, and E. Rosen, *Undecidability results on Two-variable logics*, Proceedings of 14th Symposium on Theoretical Aspects of Computer Science STACS'97, Lecture Notes in Computer Science Nr. 1200, Springer (1997), 249–260. A revised and expanded version will appear in Archive of Mathematical Logic.

[10] M. Mortimer, *On languages with two variables*, Zeitschr. f. math. Logik u. Grundlagen d. Math. 21 (1975), pp. 135–140.

[11] M. Vardi, *Why is modal logic so robustly decidable*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science 31, American Math. Society, 1997, 149–184.