First-order logic with self-reference

Reijo Jaakkola

Tampere University

July 13, 2022

In [1] a novel Turing-complete logic was introduced, which extends first-order logic FO with two natural features. The first one is the ability to modify the underlying structures, while the second one is the ability to use recursion (looping) via self-reference. Syntactically speaking, the self-referential statements in the Turing-complete logic are very analogous to the (in)famous goto statements that are permitted in several programming languages. The purpose of this presentation is to present very recent work on an extension FO(C) of FO with this type of self-reference (without the ability to modify the underlying model).

The semantics of FO(C)-formulas are based on semantical games and the looping based on self-referentiality is handled in these games as follows. Consider a formula $L_C(P(x) \vee C)$, where C is a symbol known as the *claim symbol* (intuitively in the previous formula C claims that the formula $(P(x) \vee C)$ holds). If during the evaluation game the players reach the novel atomic formula C, then the players jump back to a position where the current formula is $(P(x) \vee C)$.

The semantics of the reference symbols can be handled in several different ways. The semantics that we described above is called the unbounded semantics of FO(C), since there is no bound on the number of times players can jump back to an another subformula. In our work we also consider semantics that we call bounded semantics, where the two players need to first agree on a global bound n, which will work as an upper bound on how many times the players can jump back to an another subformula.

We emphasize that FO(C) contains sentences φ which are not determined everywhere, i.e., for which there exists a model M such that neither φ nor $\neg\varphi$ is true on M. A natural example of such a sentence is the sentence $L_C\neg C$, which intuitively speaking formalizes liar's paradox, since C is "claiming" that $\neg C$ holds.

Unsurprisingly, under both bounded and unbounded semantics the logic FO(C) becomes rather expressive. For example, under bounded semantics FO(C) is able to define the class of all graphs that have a bounded diameter, while under unbounded semantics FO(C) is able to define the class of all well-founded linear orders. Given the rather high expressive power of FO(C), it should not come as a surprise that there is no sound and complete proof system for FO(C), regardless of whether bounded or unbounded semantics are used.

In contrast to the incompleteness of FO(C), we will present a sound and complete proof system for the set of valid sentences of FO(C). The proof of the completeness of our proof system is based on the use of approximants, which intuitively speaking are FO-sentences that describe an initial portion of a semantical game played on a sentence of FO(C). If we are using bounded semantics, then it is straightforward to show that a sentence of FO(C) is valid if and only if one its approximants is. We are able to show that this equivalence continues to hold if we switch to unbounded semantics. Interestingly, the proof bears some similarity with the classical proof of Lindström's theorem.

Another interesting by-product of our completeness proof is that our proof system is able to prove that every sentence of FO(C) is equivalent with a sentence of FO(C) where no claim symbol C occurs in the scope of a negation. This is interesting, because the combination of negation and recursion has been challenging in other logics that use recursion, such as FO extended with least fixed points, where one needs to require that no fixpoint variables occurs in the scope of an odd number of negations (to guarantee the monotonicity of the operator for which the fixed point is computed).

If time permits, we will also go through other results that we have been able to obtain for FO(C). For instance, we have been able to show that under bounded semantics the satisfiability problem of FO(C) is Σ^0_2-complete while under unbounded semantics it is Σ^1_2-complete. We have also been able to show that, regardless of whether bounded or unbounded semantics is used, every sentence of FO(C) which is determined everywhere must be equivalent to one of its approximants (and in particular equivalent to a sentence of FO). We emphasize that similar results on everywhere determined sentences fail if we restrict our attention to finite models.

This talk is based on joint work with Antti Kuusisto.

Keywords: natural deduction, first-order logic, complexity

References

[1] Antti Kuusisto. "Some Turing-Complete Extensions of First-Order Logic". In: *Fifth International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2014*. Vol. 161. 2014, pp. 4–17. DOI: 10.4204/EPTCS.161.4.