Combinatorial Aspects of Interpretability Logic

Vedran Čačić, University of Zagreb, Croatia

Gödel's theorems were a big breakthrough for matematical logic. With time, mathematicians started to wonder how they can be generalized, and what else, based on some simple facts we knew, could be deduced about provability predicates. Formalizing provability over some base theory T as a unary modal operator $\Box$, led to the theory GL (named after Gödel and Löb) which we know today is the provability logic of many base theories.

Provability is great for judging absolute strength of some formula against a theory. But what about relative strength? For some base theory T and two formulas F and G, is $T + F$ interpretable in $T + G$? That is, can we find a way of reinterpreting symbols of T, preserving provability of whole T, but such that (reinterpreted) formula F becomes a theorem, if we add G as an axiom? Here we don't just divide formulas into black and white, but try to order them in various shades of gray. In fact, various colors would be a better analogy, since the ordering is usually not total.

We can do something quite analogous here. Formalizing interpretability in the above sense as a binary modal operator $\rhd$, we are led to various interpretability logics, most basic of which is probably IL. Unfortunately, IL itself is, unlike GL, just a "lowest common intersection" of those interpretability logics, and different base theories add to IL different principles of interpretability, extending it in diverse ways.

However, we still can consider properties of GL, and ask ourselves if IL has something analogous. One well-known property of GL is that its closed formulas have very regular normal forms: every GL formula without variables is equivalent to a Boolean combination of formulas $\bot$, $\Box\bot$, $\Box\Box\bot$, and so on. That Boolean combination can be further normalized, taking into account that $\Box^n\bot \to \Box^m\bot$ whenever $n \leq m$.

Do IL formulas have something similar? The first hard question is: what are the basic blocks here? In GL, it was easy—repeating $\Box$ before $\bot$ gives a natural single-parameter countable family of "propositional variables", to be connected into Boolean combinations. There is not anything analogous in IL, except that family itself. Namely, it can be easily seen that $\Box$ can be emulated in IL, $\Box A$ being equivalent to $\neg A \rhd \bot$ ($\bot$ is invariant under interpretation, and $T + \neg A$ is inconsistent iff $T \vdash A$). So the same family $\{\Box^n\bot : n \in \mathbb{N}\}$ is available in IL, too.

In [1], we have shown that many IL formulas have GL equivalents, and by that, the same normal forms as GL formulas. Here we count those formulas, and find the exact share they have in the whole closed fragment of IL.

By studying the general forms of such families of IL formulas, we become aware of many interesting combinatorial problems. First, how to define "share" in the first place? Set of all closed IL formulas is infinite, but we use asymptotics based on complexity of formulas. Second, how to count formulas given recursively? Those recurrences often don't have closed form solutions, but asymptotic behaviour can be obtained via generating functions. And third, many of classes we have to count collectively aren't disjoint: exclusion-inclusion formula helps here.

References

[1] V. Čačić, M. Vuković, *A note on normal forms for the closed fragment of system IL*. Mathematical Communications, **17** (2012), 195–204