Title
Thoughts on the Prospective MML-TIP: A Mental MetaLogic-Based Theorem Prover

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 23(23)

ISSN
1069-7977

Authors
Bringsjord, Selmer
Yang, Yingrui

Publication Date
2001

Peer reviewed

Thoughts on the Prospective MML-TP: A Mental MetaLogic-Based Theorem Prover

Selmer Bringsjord & Yingrui Yang
selmer@rpi.edu • yangyri@rpi.edu
Dept. of Philosophy, Psychology & Cognitive Science (S.B. & Y.Y.)
The Minds & Machines Laboratory (S.B. & Y.Y.)
Department of Computer Science (S.B.)
Rensselaer Polytechnic Institute
Troy NY USA

Why do automated reasoning systems fail to even *hint* at the power of a human mathematician or logician? (See, e.g., the comparison between the "real" Gödel and machine versions of proofs of his famous incompleteness theorems: Bringsjord 1998.) Why is it that the best machine-proved theorems are shallow when matched against minds? From the standpoint of psychology and cognitive science, this is an exceedingly difficult question to answer, in no small part because these fields don't offer a mature theory of mature reasoning. What they *do* offer are theories of *im*mature reasoning, that is, theories of — as it's said — untrained or ordinary reasoning. (We have no precise definition of the difference between expert and novice reasoning, of course; but the distinction seems to make intuitive sense, and we can resort to characterization by example. E.g., *reductio ad absurdum* is probably *generally* the province of expert deductive reasoners. Certainly, say, diagonalization is the province of such reasoners.) If we assume that these theories can scale to the expert case, then it's no surprise at all that automated reasoning systems, when stacked against mathematical minds, look pretty bad. The reason is that theories of ordinary reasoning, before the advent of our Mental MetaLogic, invariably put their eggs in one basket. For example, the representational system of current mental logic theory can be viewed as (at least to some degree) a psychological selection from *only* the syntactic components of systems studied in modern symbolic logic. For example, in mental logic *modus ponens* is often selected as a schema while *modus tollens*[1] isn't; yet both are valid inferences in most standard logical systems. Another example can be found right at the heart of Lance Rips' system of mental logic, PSYCOP, set out in (Rips 1994), for this system includes conditional proof (p. 116), but *not* the rule which sanctions passing from the denial of a conditional to the truth of this conditional's antecedent (pp. 125–126). So, a theorem prover based exclusively on mental logic would of necessity fail to capture human mathematical reasoning that is "semantic." And while it's certainly true that expert reasoners often explicitly work within a system that is purely syntactic, it's also undeniable that such reasoners often work on the semantic side. Roger Penrose has recently provided us with an interesting example of semantic reasoning: He gives the case of the mathematician who is able to see via an image of the sort that is shown in Figure 1 that adding together successive hexagonal numbers, starting with 1, will always yield a cube.

Now, as a matter of fact, nearly all theorem provers are essentially incarnations of mental logic, so it is unsurprising that such provers are impoverished relative to trained human reasoners. Mental MetaLogic (MML) (Yang & Bringsjord under review) is a theory of reasoning which draws from the proof theoretic side of symbolic logic, the semantic (and therefore diagrammatic) side (and hence owes much to Johnson-Laird's 1983 mental model theory), and the content in between: metatheory. Furthermore, while theories of reasoning in psychology and cognitive science have to this point been restricted to elementary reasoning (e.g., the propositional and predicate calculi), MML includes psychological correlates to modal, temporal, deontic, conditional, ... logic. Accordingly, MML-TP will be theorem prover that ranges over a diverse range of representational schemes and types of reasoning.

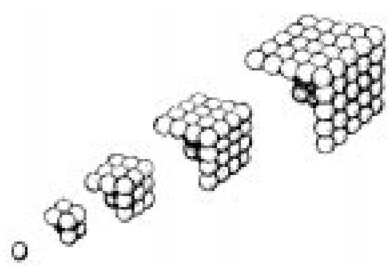

Figure 1: Viewed a Certain Way, the Cubes are Hexagons

References

Bringsjord, S. (1998), 'Is Gödelian model-based deductive reasoning computational?', *Philosophica* **61**, 51–76.

Johnson-Laird, P. N. (1983), *Mental Models*, Harvard University Press, Cambridge, MA.

Rips, L. (1994), *The Psychology of Prooof*, MIT Press, Cambridge, MA.

Yang, Y. & Bringsjord, S. (under review), 'Mental metalogic: A new paradigm in psychology of reasoning'.

[1] From $\phi \rightarrow \psi$ and $\neg\psi$ one can infer to $\neg\phi$.