AI at CADE/IJCAR

Josef Urban[1]

Czech Technical University in Prague

Has first-order ATP flattened out? This was a discussion topic at the PAAR workshop at IJCAR'14. I had two papers at the workshop. One about a 25% improvement of E prover's auto-mode by BliStr – an evolutionary strategy invention. Another a system description of MaLARea 0.4 and 0.5, a system that won the CASC 2013 LTB competition by 77% (the largest margin in CASC since 2000). Both papers were rejects from IJCAR'14. Nobody cared for these numbers and the methods involved. The reviews were rather explicit about it. This has been happening over and over with the CADE/IJCAR conferences.

Machine learning, evolutionary methods, probabilistic inference guidance, probabilistic parsing: topics often uninteresting and unknown to the small and quite unchanging circle of CADE/IJCAR reviewers. Studying the motivation, taking a look at some ML 101 course to avoid elementary pitfalls in reviews, or getting an expert subreviewer is often too much to ask. Sometimes the rejections say explicitly: "who cares about competitions and benchmarks?". No idea that ATP is actually useful and that bechmarks, challenges and competitions like CASC LTB have been targeting (for a decade!) huge applications like ITP. IJCAR'16 did not have the rebuttal phase, removing one of the very few weapons for answering to very uninformed reviews.

Bob Veroff – one of the few people who use ATPs to prove hard open conjectures – has been gone from CADE/IJCAR long ago. His ongoing research on the hints guidance and proof sketches in related theories has been crucial to his success. A very interesting and practically working kind of symbolic machine learning targeted at saturation-style ATP. Not interesting to CADE/IJCAR.

Max Planck said that science progresses by opponents dying out. This will most likely have to happen to the current petrified CADE/IJCAR, which revels in completeness proofs and considers modern probabilistic AI methods to be "tuning" and "black magic". Reasoning topics are now more and more being taken up by the broader AI/ML community, which is emboldened by its recent huge successes. Nobody has solved Go, and nobody can "prove" that AlphaGo's strategy is "complete". But nobody cares. Nobody cares about manually encoding specialized solvers for ten (or so) possible openings in Go - total waste of time and human brain power. At FLoC'06 in Seattle, I wanted to scream when I heard that one of the hot AR challenges was going to be SMT (specialized solvers for ten or so theories). This is really not the way to build general AI for all of math and reasoning! Wake up! AI conferences like NIPS have grown to an army of six thousand young excited people who are willing to experiment and attack old tasks in new ways. Very often this includes some sort of learning.

I have long argued that deduction and learning need to be combined to get to strong AI and ATP. It is very naive to think that we will soon be able to fully manually encode all aspects and procedures of thinking and theorem proving. Mathematicians (and thinking people in general) learn, use, and improve a huge bag of context-dependent heuristics during their studies and work. Some of them can be made "crisp" after a while, resulting e.g. in provably correct decision procedures for various fragments. But a vast majority of them are fuzzy. Today's ATPs mostly lack the capability to learn and to use the learned knowledge. Most of today's ATP research is done outside of large theories where learning is most rewarding.

The AR domain has strong and solid semantics. It is really cool an mostly unseen in other AI areas. We (humans) can sometimes turn "fuzzy" ideas into "crisp" ideas that have formally

defined meaning and proofs, allowing computation, exact reasoning, proof verification, etc. But most of CADE/IJCAR has become so focused on producing crisp logic-oriented papers, that they forgot that 99% of human thinking and invention involves some fuzzy levels. One of the greatest opportunities for the AR field is today invention of methods that automatically go from the fuzzy understanding (e.g., in the form of some statistically learned knowledge), to its crisp formulation (in the form of a proved theorem or procedure working under exact assumptions). And also back: from crisp formulations to further fuzzy conjectures. This is going to be a huge AI topic that should be naturally taken up by the AR community. Unfortunately, the current AR community avoids statistical methods like a disease.

I am not going to argue and explain more here – I have written enough on this topic in the past fifteen years. Despite Dmitriy's and Giles's honest efforts, it is hard to believe that it would not be a wasted effort, as were my occasional replies to the most appalling reviews in the past. In 2004, I have tried something similar with the declining Mizar community, and it has been a complete failure. I wrote a wishlist[1], saying that they have a good chance to be a great project leading and benefiting the world (and themselves) IF they wake up to the 21st century. Instead, they have been doing the exact opposite of all of my points ever since, committing a slow honorable suicide in the diminishing circle of their small sect. It will take hundreds of years to get to the QED project's goals, because our great and unmistakable leaders have said so. So let's produce more incremental research.

I wish CADE/IJCAR had a better fate, but I am not going to put my money on it. I can see many of the symptoms. But I have a great belief in the future of AI and AR.

---

[1]`http://wiki.mizar.org/twiki/bin/view/Mizar/MizarWishlist`