

20 YEARS OF MPTP

Josef Urban

Czech Technical University in Prague

Mizar 50 Workshop
July 31, 2023, Bialystok



Outline

- As a tribute to Andrzej, this talk is full of side remarks
- "To prevent the discussion from the usual side remarks by the leader of this team" (P. Rudnicki, lunch discussion in the "Tent" ("Namiot"), somewhere around 2003-2005)
- As a tribute to Andrzej, this talk should be fully improvised without any slides (but ...)
- Probably more of the earlier history

MPTP - Mizar Problems for Theorem Proving

To: <mizar-forum@mizar.uwb.edu.pl>
Subject: [mizar] Mizar Problems for Theorem Proving
From: Josef Urban <urban@ktilinux.ms.mff.cuni.cz>
Date: Mon, 10 Mar 2003 14:04:34 +0100 (CET)

Mizar Problems for Theorem Proving (MPTP) version 0.1
is at <http://ktilinux.ms.mff.cuni.cz/~urban/MPTP/MPTP.tar.gz> .

The packed full distribution has about 70M and contains databases
and scripts for generating proving problems from

- 1) complete proofs of Mizar theorems (about 30000)
- 2) Mizar simple justifications - checker problems (about 630000).

If you want to generate the databases from your own Mizar
distribution, the necessary files only are at
<http://ktilinux.ms.mff.cuni.cz/~urban/MPTP/MPTPbasic.tar.gz>
(packed to about 230k). Generation of databases takes about 1 hour
on P4.

Some documentation can be found in the distro, paper and slides about it
are at

<http://link.springer.de/link/service/series/0558/papers/2594/25940203.pdf>
and <http://monet.nag.co.uk/mkm/bertinoro03/urban.pdf>.

- 2004 (JAR): MPTP – Motivation, Implementation, First Experiments
- “[...] We present results of first experiments with re-proving the MPTP problems with theorem provers.”
- “We also describe first implementation of the Mizar Proof Advisor (MPA) used for selecting suitable axioms from the large library for an arbitrary problem and”, ...
- ... “again, present first results of this combined MPA/ATP architecture on MPTP”

History

- around 1995: I decided to learn math/scientific reasoning automatically by mining large math corpora
- apparently, this idea has been recently (2018 - 2023) discovered by people at Google, Facebook, OpenAI, Microsoft, DeepMind, U. of Toronto, Stanford, CMU, etc., and according to NY Times, Quanta, New Scientist and other reliable outlets it seems very interesting ;-)
- **seriously**: Turing's 1950 AI paper: last chapter is already on "learning machines" - learning to play Go, etc.
- 1990s: The Munich AR group - Denzinger, Schulz, Fuchs, Fuchs, Goller ... ML for ATP
- Tree neural nets first introduced by Goller & Kuchler for logic, etc.
- many papers and several theses (I only came upon them after my MSc was finished in 1998)

History

- My Msc thesis - try learning reasoning from Mizar texts (reverse engineering Mizar)
- found Mizar early on the (new) World Wide Web, started to hack it
- “[...] we should be interested in the question how to find heuristics, and more generally a sufficiently large system of heuristics, that would give us recommendations for theorem proving similar to those of human.”
- “The most straightforward way of creating heuristics seems to be simply thinking of them and typing them in. It corresponds to teaching done completely by explanation.”
- “There is redundancy in this method, because there surely are some heuristics about finding heuristics, so why not let them find the rest for us. I think ‘learning from examples’ could be a (meta)heuristic of this kind and in this work experiments are carried out to find out what is possible to discover using this heuristic.”

History

- I wanted to define a "proof situation", "its characterizations", "the available actions"
- the plan was to make a rich evolving language for the "characterizations" and the "actions", going from low-level primitives to very high-level ones
- this is today mostly known as reinforcement learning
- email to Andrzej in 1999, asking for how the polymorphism resolution works in Mizar
- ... after my its reverse engineering in my MSc didn't go so well ...

Ingo Dahn and ILF

From urban@mail.kolej.mff.cuni.cz Wed Mar 3 11:02:12 1999 +0100
Date: Wed, 3 Mar 1999 11:02:11 +0100 (CET)
From: Josef Urban <urban@mail.kolej.mff.cuni.cz>
To: dahn@uni-koblenz.de
Subject: translating Mizar to Prolog

Dear prof. Dahn,

[...]

The goal of my master thesis [...]

"Experimenting with Machine Learning in Automatic Theorem Proving" was to learn from mathematical texts some heuristics and notions, that people use, when they prove theorems.

For this, I chose Progol as the learning system, and Mizar as the mathematical texts.

To be short, it simply turned out, that for the purpose of learning, we have to translate a large part of the Mizar system to Prolog, it was quite complicated, and so the result of the thesis was rather an environment, in which some nontrivial learnings could be done.

[...]

History - ILF

- 1999 - got the (then already abandoned) ILF system from Ingo Dahn and Christoph Wernhard
- a major piece of work (Prolog, ATPs, PostgreSQL) relying on a custom exporter written by C. Bylinski
- Dahn's group did a lot of interesting work e.g. on handling the Mizar type system as inclusion operators
- possibly predates even some of the work on similar HOL/hammer encodings by Joe Hurd
- I decided not to go there in my own later MPTP attempts and just use the straightforward type guards in the first MPTP releases
- here is a research opportunity - Hi MSc/PhD students ;-)

History - Getting serious

- email to Andrzej in 2000, asking what Mizar article I should write to get the sources
- so that I can see the reasoning steps done by the checker (Hi Mario! ;-)
- 2000: wrote the article, got the sources, came to Bialystok in summer 2000
- wrote another article and started to read and document the checker sources (Hi Mario ;-)
- team members at that time:
 - Andrzej Trybulec, Czeslaw Bylinski, Grzegorz Bancerek, Artur Kornilowicz, Adam Naumowicz, Adam Grabowski, Robert Milewski, Mariusz Zynel, Mariusz Giero, Bartek Skorulski, ...

From urban@mail.kolej.mff.cuni.cz Wed Mar 1 16:39:47 2000
Date: Wed, 1 Mar 2000 16:39:47 +0100 (CET)
From: Josef Urban <urban@mail.kolej.mff.cuni.cz>
To: Andrzej Trybulec <trybulec@math.uwb.edu.pl>
Subject: membership in Assoc. of Mizar Users

Dear prof. Trybulec,

I am the person who asked you a year ago about types in Mizar. Since then I worked with the ILF system you recommended to me [...]

I am trying to use the pl-files as examples for learning heuristics for theorem proving. [...] but the description ends at the level of simple justifications, so I do not know the proof steps carried out by the Mizar checker [...]

As you wrote me, the way to get the sources is to become a member of the Assoc. of Mizar Users, so I would like to apply for the membership.

[...]

So please could you let me know if you have an idea of some topic that is currently needed and could be written by a beginner?

[...]

First Steps with the Mizar Sources

- 2000 - continued the project of having fully detailed proof objects from the checker
- this blew up the checker size about twice
- (later, Bill McCune told me he had similar experience with documenting Otter and it led him to rewrite Otter into Prover9)
- unfinished (Hi Mario ;-)

First Steps to Get ATP Problems

- 2001: started to hack the sources to output ATP problems, mostly using SPASS and its DFG format
- 2001 - a detour: try to implement the Mizar soft type system as an extension to the SPASS soft typing mechanisms
- again unfinished - hard for me at that time
- btw. certain Diane Reynolds submitted a paper "Adding Mizar-style Soft Typing to Satallax" in 2018 - Hi Diane ;-)
- btw, some of this code has likely been re-used for the independent Mizar elaborator written by Chad for our autoformalization work over Mizar in 2019/20 - Hi Chad ;-)

Side Note: MoMM - Most of Mizar Matches

- an E prover indexing's hack to add subsumption wrt Mizar type system
- (smart) memorization is incredibly useful (Hi LLM people ;-)
- over 50% of Mizar matches!
- side product of my unsuccessful hacking of SPASS's soft types
- and attempting to make SPASS work on large theories by adding indexing structures from E - Hi Stephan ;-)

MPTP 0.1

- 2003: MPTP 0.1 - DFG - direct SPASS, type guards used
- re-proving (fixpoint algo for adding the background/typing knowledge used by Mizar)
- advantage: I did it for all of MML, even if with bugs, allowed to measure the ATP potential
- ML systems trained for premise selection and hammering
- first alternative proofs found by the combined ML/ATP systems
- thousands/millions/zillions ATP problems (Hi Geoff!) easily generated
- cf. the "data scarcity problem" recently claimed by some ML/TP papers ;-)

Simplifications in MPTP 0.1

- higher-order constructs incorrectly translated
- all fraenkels mapped to just one constant
- no scheme handling?
- btw: full higher-order export reported here in Bialystok in 2016 by Chad (Hi Chad!)
- btw: great progress on higher-order ATP recently:
- Chad's Satallax finally beaten after a decade of rule -
- thanks to the Zipperposition and higher-order E (Hi Jasmin, Petar et al!)
- and the higher-order Vampire people (Hi Ahmed!)
- btw: see Mario's Metamath hammer talk tomorrow for the performance of higher-order ATPs on the Metamath set.mm library

- 2004/5: XML-ization of Mizar
- to be able to do MPTP
- and all sorts of other exports
- properly and without the usual bitrot of ad-hoc exporter tools (happened to ILF before)
- Andrzej took some convincing and beers in the jazz bar
- Czeslaw's help was crucial (Hi Czeslaw :-)

- 2005/6: the XML used to create intermediate TPTP-like representation - MPTP2
- btw. maybe TPTP will eventually get there by the time I retire:
- term-dependent types, intersection types and subtyping, undecidable type system, ...
- MPTP2 represents Mizar quite faithfully and also translates it to FOL quite faithfully:
- there should be no soundness issues and not many completeness issues
- proper handling of schemes, fraenkels, eventually also the "the" ;-), etc.
- Attempts to make it efficient - e.g. by making the first-order definitions as general/shared as possible for fraenkels

MPTP Challenges

- 2006 - MPTP Challenges
- each for \$100 - the bushy (small) and chainy (large) problems
- first large-theory global time limit ML competition:
- learn and solve many problems within a day in arbitrary order
- see the funny web (Hi Geoff!):
- <https://www.tptp.org/MPTPChallenge/>

- 2006: automated the interleaving of the ML and ATP done since 2005 over MPTP
- developed/upgraded since then
- semantic features in 2008
- btw. still - even today - semantics largely outside of the deep learning crowd's thinking box
- many different learners since, LSA, BliStr/Tune, ENIGMA, (conjecturing?, ...)
- btw., the MaLAREa feedback loop by now claimed couple of times by others - Hi Google, Deepmind, ... ;-)
- many people involved - J. Vyskocil, C. Kaliszyk, J. Jakubuv, D. Kuhlwein, etc.

ATP Cross Verification

- 2007: ATP cross-verification of the MPTP Challenge problems
- shown that ATPs can completely replay the Mizar proofs of the MPTP Challenge problems
- took some effort - needed Malarea - and found a rare completeness bug with the ATP export
- an alternative route to my old project of getting proof objects out of Mizar

- 2010: MizAR - the hammer and presentation tool
- both web-based and emacs based - see the demo:
- <http://grid01.ciiirc.cvut.cz/~mptp/out4.ogv>
- "cloud based": "The Mechanical Turk/Dutchman" suspected at my Mizar Japan talk - Hi AdamG and Mizar Japan ;-)

- 2010: Mizar Wiki - verification and XML-based htmlization as git hooks
- git allowed it all to be local and shared
- nice collaborative model (github actions today?)
- partial overlap with MizAR, but not tightly combined
- with Jesse Alama, Piotr Rudnicki and Herman Geuvers (Hi guys!)

Mizar Licensing

- btw, this triggered the 2010/2011 Mizar licensing work
- also triggered e.g. by inclusion of Mizar problems in TPTP
- L. Mamane, M. Kohlhase, S. Schulz, J. Alama, A. Naumowicz, P. Rudnicki
- I finally got to chair a committee: "Mizar Licensing Committee" ;-)
- btw.: We have in that 2011 licensing paper (thinking about the MizAR hammer trained on all of MML) pretty much predicted the current legal troubles with training AIs/LLMs on all of github, wikipedia, etc.:
- quote "For example, Wikipedia is today used for data and algorithm extraction, used in advanced algorithms by, e.g., the IBM Watson system. With sufficiently advanced extraction algorithms (which we are clearly approaching), many "documents" can become "programs" [with all the legal trouble attached]."

More ATP Fun with MPTP

- 2010: Large MPTP/ATP eval with Krystof Hoder and Andrei Voronkov (Hi Krystof/Andrei)
- Vampire tuned quite nicely
- all ATPs solved over 40% of the bushy (small - reproofing) problems

More ATP Fun with MPTP

- 2010/11: first experiments with training internal guidance of simpler ATPs:
- MaLeCoP - led to a whole family of learning-guided connection-based systems (J. Vyskocil, C. Kaliszyk, M. Olsak, Z. Zombori, M. Rawson, F. Romming, J. Heemstra, ...)
- 2012: first experiments with automated synthesis of targeted ATP strategies on the MPTP problems:
- BliStr/Tune (Hi Jan, Martin, ...) - took off for all sorts of datasets and ATPs

More ATP Fun with MPTP (and Cezary, Jan, ...)

- 2012/13: - we got the 40% hammering (larg problems) results and 56% bushy results
- on Flyspeck and MML with Cezary - lots of optimism (Hi Cezary!)
- 2014 - the 3 AITP Bets/Challenges for 10k eur from my talk at IHP:
<http://ai4reason.org/aichallenges.html>
- one of them discharged by us in 2021
- (3 years ahead of schedule - Jan's talk tomorrow - Hi Jan!)

And Even More Fun with MPTP

- 2015-20: first learning-based autoformalization experiments over Mizar and Flyspeck
- for Mizar using Grzegorz's FM work
- updated to the XML-based processing and updated by Adam Naumowicz during his visit to Prague
- large autoformalization LM experiment using 1M synthetic LaTeX/Mizar pairs
- with Shawn Wang and Cezary
- pretty encouraging results - I started to much more believe my autoformalization propoganda then ;-)

MizAR60 for Mizar 50

- 2020/2021 - 75% on the bushy problems
- almost 60% in the hammering setting
- deploying the newly developed heavy ML guns ENIGMAs/Deepires
- strong premise selectors, many learning/proving feedback loops
- many people involved – see Jan's talk!

MPTP problems in the Proofgold blockchain

- 2018: Andrew Appel complained at Hales60 in 2018:
- that my AITP-and-Hales talk missed the blockchain buzzword - Hi Andrew ;-)
- 2021: hard MPTP problems as bounties in the Proofgold blockchain - you can get rich by doing ATP!
- talk to Chad to learn about blockchains for formal math
- (and Proofgold specifically)

Conclusion and Thanks

- Great things in AI/TP have happened thanks to the Mizar/MPTP project
- I don't know how much Andrzej (and others) was expecting and aiming for it
- But he was a sci-fi expert and a visionary of a rare kind
- He had a vision that seemed crazy to others,
- and also put massive practical effort towards it, followed by the whole Mizar team
- Quite uninterested in money, PR, recognition, etc., doing it out of the scientific curiosity, for the fun
- I am very grateful for having been for 20+ years part of the Mizar project
- MPTP and related Mizar-based projects have been good fun and a great scientific adventure
- Thanks everybody!