Fully automated contributions to formal systems libraries

Josef Urban <josef.urban@gmail.com>   Sun, May 5, 2024 at 6:32 AM
To: Josef Urban <Josef.Urban@gmail.com>

Forwarded Conversation
Subject: Fully automated contributions to formal systems libraries
----

From: **Stanislas Polu** <spolu@openai.com>
Date: Sat, Aug 1, 2020 at 10:28 AM
To: Josef Urban <josef.urban@gmail.com>, <email@thibaultgauthier.fr>

Josef, Thibault,

Hope you're doing great. First of all, congratulations Josef for your
talk at CICM. I really enjoyed it and glad that you're using GPT-2 on
Mizar. The conjectures and cuts you report are really exciting. Also I
really hope we'll get to meet IRL later this year at AITP, fingers
crossed.

We've been working ourselves with GPT-type models on a variety of
formal systems (Hol light through Holist, Lean and Metamath). As you
may have seen on the Metamath google group, we've shared a series of
proof shortenings[0][1][2] that were very positively received by the
community:

> "I had a look at the proofs —very impressive results! Especially because we had a global minimization recently, and your method found much shorter proofsnevertheless."

> "Any ML-based system is impressive if it can find many shorter proofs than theones we already have. Nice work."

> "The shorter proof is easier to translate. It's more symmetric in that it treats Aand B identically. It's philosophically more concise in that it doesn't rely on theexistence of a universal class of all sets."

These new proofs were sampled from our system without any human assistance.

As we're working on a paper to share our work with the community, I'm
looking for references to other deep learning systems having
effectively contributed proofs back to a formal library. We obviously
would love to put this sentence in the paper but I'm pretty sure it's
a lie and would love to find out the references associated:

> To our knowledge, these shortened proofs[1][2] are the first effective contribution of a deep learning system to a formal mathematics library.

Do you have any references we could cite of other deep learning
systems similarly contributing proofs that were received positively by
maintainers?

Thank you so much for your help.

[0] https://groups.google.com/g/metamath/c/-FNsw2wyllI
[1] https://github.com/metamath/set.mm/pull/1547
[2] https://github.com/metamath/set.mm/pull/1561

From: **Josef Urban** <josef.urban@gmail.com>
Date: Sat, Aug 1, 2020 at 10:45 AM
To: Stanislas Polu <spolu@openai.com>
Cc: thibault gauthier <email@thibaultgauthier.fr>, Cezary Kaliszyk <cezary.kaliszyk@uibk.ac.at>

Hi Stanislas,

There is too much of this and I have only my phone.

Take a look at the references section of my CICM slides - http://grid01.ciirc.cvut.cz/~mptp/nnconj_pres.pdf . It should include the work on concept alignment for conjecturing and transfer of theorems between HOLs, autoformalization (PCFG/neural) based conjecturing, etc.

Also, finding of alternative (typically shorter) proofs has been going on with hammers for a long time - basically since 2003 and the first MPTP, Mizar Proof Advisor, Malarea, Sledgehammer, etc. Look at the ancient papers, our Mizar40 and Flyspeck papers - there should be examples of much shorter proofs. There is a summary in the "hammering towards QED" paper (2016). Then there are Tactictoe examples, etc. A collection of nice ENIGMA proofs is growing at https://github.com/ai4reason/ATP_Proofs .

Have fun,
Josef
(Disappearing into wilderness for a couple of weeks)

From: **Stanislas Polu** <spolu@openai.com>
Date: Sat, Aug 1, 2020 at 10:57 AM
To: Josef Urban <josef.urban@gmail.com>
Cc: Cezary Kaliszyk <cezary.kaliszyk@uibk.ac.at>, thibault gauthier <email@thibaultgauthier.fr>

Thank you so much!

While I'm sure that hammers have contributed numerous proofs and proof shortenings I was looking for references focused on deep learning systems.

Thibault, did TicTacToe (with deep neural networks) contribute proofs back to HOL4 that were accepted and merged into the main library?

Do you know of any such contributions by other deep learning systems (Deephol, ASTactic, ...)?

-stan

From: **Josef Urban** <josef.urban@gmail.com>
Date: Sat, Aug 1, 2020 at 11:03 AM
To: Stanislas Polu <spolu@openai.com>
Cc: Cezary Kaliszyk <cezary.kaliszyk@uibk.ac.at>, <thibault_gauthier@hotmail.fr>

I never get the "deep" distinction. There is good/bad, working/nonworking ML and AI, not deep and non-deep.
J.

From: **Stanislas Polu** <spolu@openai.com>

I presume I refer to "using neural networks" precisely. Which arguably is an non trivial abd meaningful distinction? ;-)

-stan

Stating that an ML/AI method A works differently than method B is Ok.

Omitting from an overview other/better ML/AI methods that solve the same problem as A just because they are not A is pseudoscience.

Completely agreed. The goal is not to omit but to state some capabilities that were not yet achieved by systems of type A while still reporting the overall capabilities of systems of types A, B, C, ...

-stan

In their GPT-f paper [1] OpenAI never cited the prior art they asked me about below. Neither the prior GPT-based work they "enjoyed" and found "exciting" when they watched me presenting it at CICM. I was quite amazed and only found the right word for this when I saw the quote on p.1 of Bobzien's article (discussed on FOM) [2].

[1]: https://arxiv.org/pdf/2009.03393.pdf
[2]: https://philarchive.org/archive/BOBFPT .