The importance of applying computational creativity to scientific and mathematical domains Co-author feedback

March 6, 2019

Daniel Arnold

Some comments, mostly where I think you need to back up points with references.

Im not sure you tie the calls to examples from maths and geology. Do you want to do this? If so I can think of some examples. Experiments we could try so to speak.

Daniel Winterstein

Overall: monkey read, monkey like :)

Re. the Saturation Problem I'm not convinced the Saturation Problem exists. E.g. the internet has created a saturation of many (human made) art forms - but society is still keen for more. I can't see a saturation problem really happening unless both history and creativity stagnate. The explosion in productivity of art *does* have profound implications for art – and not all bad: e.g. photography, by flooding the market for images, forced artists to redefine their value and led to the creation of modern art. Also, I can't see "look at science" would be a solution. If AI has reached the point where it can auto-generate so much great art, writing, music, and film that no more is wanted – then switching focus to science as a relatively neglected area would be only a short-term respite. Science too would become saturated.

Extra citations? "It has certainly resulted in some original and significant proofs" "Deep learning and ML are making inroads everywhere" – an eg for each domain? I suggest removing "windfarms" – I know the story – it's DeepMind's PR team putting a misleading spin on a perfectly decent project. A more accurate write up would have been: "By using weather reports, we can sell electricity futures, and you can make a bit more money that way. We needed a bit of ML to link the weather reports with the wind-farm output. Linear regression probably would have worked, but we like neural nets."

s/focus on science/have a balanced approach that doesn't neglect science/ – I don't think we're against CC for art.

Define "framing".

A mega sentence, which I had to read a couple of times to parse, especially the 2 x generate: "meta-level processes which can *generate* for instance the means by which an artefact is generated"

For neural networks – I think the sentence re convolutional layers is subtly wrong. We encode structure into neural networks by choosing network layers that fit the task. Examples are: A convolutional node is a 2D matrix, inspired by filters from graphics programs, and is well-suited to edge detection or texture analysis. An LSTM network is designed for modelling sequences (e.g. sentences), using a recurrent structure with some memory nodes.

In the next paragraph: Understanding a neural net is, as you say, an active research area (i.e. we don't know how), but saliency maps are one way. Another way (better IMHO) is "interrogating" the network, via generating inputs aligned to deep features (e.g. you specify a deep-level state, then "train" the input to get close – this is how style transfer works and can be used for creativity, but you can also use it to explore what the neural net has learned). There's also example-based exploration of decision boundaries. I think there's an interesting analogy between these approaches, and how we use introspection and analysis to understand human learning. We're gradually becoming cognitive scientists and psychologists for the robots.

life-critical domains – include legal e.g. AI used to draft legal decisions (in real use today in Buenos Aires) https://www.bloomberg.com/news/articles/2018-10-26/this-ai-startup-generates-legal-papers-without-lawyers-and-suggests-a-ruling http://ojs.imodev.org/index.php/RIDDN/articl – suggests this is a black-box system (wtf).

Finally: the paper opens by saying maths and geology will be used for reference. And they are, but not noticeably more than other areas. Maybe add a section to discuss your work in those two areas? Or adjust the statement that they will be the focus.

Anyway – good writing, lots of interesting stuff in here, including many examples I hadn't previously known. I feel wiser.

All the best, - Daniel

update

A minor comment on my comment: Re. saliency maps vs "interrogating" the network, via generating inputs aligned to deep features. – each has its place. Saliency maps (and more generally, local linear approximations to arbitrarily complex models) let us examine why an AI might have made a specific decision. Exploring deep features gives us some insight into what it's learned and how it thinks in general.

Chris Warburton

Done? check.

Mike Cook

This paper is really good! All the examples in it are so interesting, and it really made me think a lot more broadly about the world of problems out there and the complexity of the impact on knowledge technology is going to (continue to) have. It's really food for thought, I enjoyed reading it a lot.

This is not really a suggestion to add, but I feel like there are two other underrated reasons why we see little scientific domain work in ICCC/CC:

It's very easy to be a hobbyist game designer/artist/composer or a lapsed one, and thus be an AI researcher in your day job with a side interest in another domain. A lot of us in CC are not-so-secretly deeply involved in the domains we work in. By contrast, it's harder to be an AI researcher and also an occasional physicist, or vice versa. I think this puts a higher value on collaboration, which is a good thing, but also I don't meet many geologists. Possibly we should think about inviting a few total, total outsiders to our conference occasionally, or something? Low-hanging fruit in most scientific fields is very easy to dismiss as uninteresting. By which I mean, when ANGELINA makes a game, it's not a good game and it's not really considered a contribution to the field, but it's comparable to a kid making their first game maybe, in quality terms. The equivalent in mathematics is something like reinventing pythagoras' theorem (this is a bad example but you get what I mean) - you might find it encouraging that you managed to do it, but there's little interest past the first one, and little implication this will build into anything. So getting the good vibes from a first step feels harder in the sciences. This one is maybe not as good a point as the first one. I'm not 100

I found the section on what understanding means fascinating. Azalea actually uses Coq in her work, and even though it's theoretically a tool to make proof easier, it requires a lot of custom work to be done to set up the environment to attack a particular class of problem. Users write "tactics" which are designed to perform certain kinds of tasks on certain kinds of problems, and often there are only a handful of people in the world who know how to leverage Coq against a certain subarea (although this is growing, Azalea's actually planning to teach a course on Coq when she gets a lectureship job). Papers at her conference get stamped (literally, in the proceedings version, there's a big digital stamp on the first page) to indicate their solutions have been "Mechanised in Coq" i.e. you don't need to read the proofs in the appendices, because we've had a Coq proof object submitted along with the paper. I find that so fascinating. I implicitly do this for every CC paper I read (I trust that they did the thing they described, and I probably shouldn't!) but here it's got a different tone.

It's a fascinating set of ideas and kind of made me eager to do work in scientific areas (while simultaneously not knowing how to begin). I have a friend who wanted to use Danesh to analyse generative models of star formation, maybe I should encourage him to get more bold and start inventing new models!

Thanks for sending the paper over, it's really fascinating.

Simon Colton

I read through the science paper this morning on the (first) flight, which

seems like a thousand years ago now! It's always great to read your papers, as I learn a lot! I think the paper is good. Please find attached an annotated version of it, and feel free to take into account any suggestions there (if there is still time to submit it). There aren't many annotations, actually, as I'm not knowledgable on much of the material. There are some additions that I can suggest:

1. In the "What is Science?" section, we could mention attempts to standardise terminology and concepts, with SI units being the obvious ones, but there are really big attempts elsewhere, such as the Gene Ontology.

2. We could add to the bottom of the Saturation Problem section a further warning that CC output in the arts might run up to a natural boundary in areas like poetry, as the poems won't be taken seriously as valuable given the lack of authenticity of life experiences they have. Given that we wrote a paper about this last year, I think we could put a paragraph in about authenticity, as part of the Saturation Problem.

3. Resolution theorem proving does actually produce a proof (as a set of resolution steps), but it's very difficult for people to read.

4. I think we could perhaps put in something from Doron Zeilberger along the lines of "The only interesting maths is the maths that people can't understand". He was quite controversial in his day for saying things like that. This would go in the section on the Understandability problem.

5. You do hint at the psychology of ANNs in the conclusions, but I could expand on this in the "forgoing understandability" approach section. To my mind, we could treat a deep-learned ANN with great predictive/generative power as a scientific discovery, like the discovery of a new type of mammal, etc. The advice here would be to encourage ML researchers to think of their DL output (the nets) as such, and ready for scientific scrutiny. It gets interesting of course, to realise that these discoveries are (albeit tangentially) akin to discovering a functioning brain. It would then be natural to investigate the discovery with methods similar to how we examine brain function in humans (and animals), namely: neuroscience, e.g., giving the ANN a stimulus and seeing which parts of the network "light up", and psychology, e.g., given the ANN a series of carefully chosen stimuli, and seeing what it predicts/generates for them. Note that I have a sentence or two about this in the HR3 paper!

Let me know if you're interested in these contributions, and I'll send some text over tomorrow morning. But if the window has closed, then we'll have to put it in the CRC, if the paper is accepted.

PS. I've noted the irony of us talking about the lack of rigour in science publishing, (i.e., non-repeatability, incorrectness of results), and then citing arxiv papers, which aren't even peer reviewed!