Student name: Brando Miranda
E-mail: miranda9@illinois.edu

Mathematics is a powerful tool that has benefited humanity for millenia. Proficiency in this subject has far reaching implications for society because physics, mechanical engineering, computer science, statistics, machine learning and the scientific method itself depend on it. For example, being able to automate mathematical proofs would allow us to create safe and interpretable A.I. because a system with provable guarantees is a system we understand and can trust. This is why in this proposal we suggest a novel way of implementing the concept of sketching for automated theorem provers using powerful ideas from cognitive science and artificial neural networks; these ideas are: compositionality, learning-to-learn, learning as model building and curriculum-learn.

The principle of compositionality highlights that human knowledge, learning and reasoning are all organized as reusable (high-level) concepts. Research in cognitive science suggests that compositionality is crucial to learning because it enables efficient reasoning. If an intelligent system (artificial or not) processes concepts blindly at a too low-level of granularity, it can suffer from combinatorial explosion as there are simply too many possible solutions to naively check. If, however, the system is able to identify the main and re-occurring high-level concepts, it can benefit from the abstraction of unimportant low-level details -- and thus reach a conclusion more efficiently. I propose to implement this with sketching where a sketcher neural network learns to outline high-level concepts such that a synthesizer can efficiently fill out the low-level detail efficiently. In addition, with modern meta-learning techniques (like iMAML) the system can be made fully differentiable.

Another observation from cognitive science is that humans are able to learn rich representations from few examples and leverage past experience in future learning tasks. Thus, with modern learning-to-learn approaches, one can accelerate the learning abilities of these cognitively inspired theorem provers. In combination with compositionality, learning-to-learn further enables fast learning via shared representations. This also avoids repetitive re-learning in the system and allows for greater power of generalization since it's able to use shared ideas in different scenarios.

One of the most powerful ways that humans learn is by using structured data (like textbooks, Wikipedia, etc.) that already outline what the concepts are, therefore humans learn incrementally without having to discover everything from scratch. We suggest that new high-level concepts can be learned incrementally in the curriculum-learning paradigm. In the domain of theorem proving, not only are there thousands of textbooks outlining the order, but there are also mathematical libraries with proofs already in program form ready to be used in this learning paradigm. We suggest that this outline of the concepts are a source of semi-supervised data for the system to learn the compositionality of mathematics.

The last promising idea from cognitive science we will outline is Learning as Model Building. One of the reasons humans are so effective at learning programming and mathematics is because they learn beyond pattern recognition. They are able to build intuitive models that capture relationships and similarity between concepts. In other words, when people think about solving a new unseen problem they do not go through all the formal rules and "deduce the solution". Instead, we reason by analogies and predictions from the model we have built about the (real or mathematical) world. Those predictions simulate the world and sketch the solution. This is exactly what the proposed sketcher would implement. In addition with the unambiguous reward of provability, these systems can be trained in a self-supervised manner with reinforcement learning.

I believe this research direction is crucial because any progress in it would benefit any human endeavor that uses mathematical thought. Furthermore, it yields a system that gives provable guarantees of other systems, which inevitably contributes to a more transparent and safer A.I. Lastly, the approach I suggest drives A.I.'s development in a novel and promising direction that is currently underexplored.