Search Heuristics and Term Encoding for Neural Theorem Proving

Automated theorem provers need to work with a variety of mathematical terms to prove any given statement. These terms include not only the definitions of functions and constants used to make the statement and specify its assumptions, but also previously proved theorems which may help prove the statement. We experiment with machine learning based automated theorem provers, exploring a variety of methods which expand their capacities to understand such mathematical terms. Additionally, we experiment with different search heuristics used to guide the theorem proving process. Experiments on CoqGym, a dataset of Coq theorems distinguished by the fact that repeat occurrences of terms between the training and testing set are guarded against, show the efficacy of these methods.