Constraint Acquisition by Transformer

S. Chandra, S. D. Prestwich, G. Provan

School of Computer Science and Information Technology, University College Cork, Ireland
Email: 122117785@umail.ucc.ie, s.prestwich@cs.ucc.ie,
g.provan@cs.ucc.ie,

Constraint acquisition (CA) aims to learn problem constraints from datasets of examples. Research over the last 20 years has yielded a variety of approaches, often based on some form of machine learning. Most methods start from a set of candidate constraints called the bias, and apply tests to decide which should be learned.

A practical drawback with these methods is the existence of exponentially large biases, preventing the learning of models involving global constraints or high-arity SAT clauses. Most work in the literature tackles biases with a few tens of thousands of candidates. GROWACQ [4] incrementally focuses on increasing subsets of problem variables, and has successfully handled a bias of over a million candidates. SEQACQ [2] handled a bias of over a billion SAT clauses via fast statistical testing. However, scaling up to even larger biases is still a problem for CA using a generate-and-test approach.

Other approaches can handle very large biases. Constraint Synthesis [1] can learn mathematical models with real coefficients, thus handling an *infinite* bias, but this is only applicable to certain types of application. The Model Seeker [3] and COUNT-CP [5] methods can handle very large biases by making simplifying assumptions about the model (for example that it involves a rectangular matrix of variables) but such heuristics require human intervention.

We aim to handle very large biases automatically, without the use of hand-coded heuristics. To this end we use a *transformer* neural network to learn to predict constraint models from examples, given a large dataset of previously-learned models and examples. As a first experiment we attempt to learn SAT models.

Transformer neural networks, renowned for their attention mechanism, are highly useful in retaining and understanding the context of problems. They were initially developed for sequence transduction tasks, in other words, the task of transforming an input sequence to an output sequence. This type of sequence-to-sequence transformation requires some sort of memory and a clear definition of context, which transformers are uniquely equipped with. This capability is particularly valuable in problems like SAT, where long-range dependencies are crucial. Unlike traditional machine learning models that may struggle with long sequences, transformers excel due to their ability to process extensive datasets and their proficiency in pattern recognition. This makes them highly effective in deciphering the intricate patterns and dependencies inherent in SAT problems, where understanding the broader context is key to generating accurate models.

Methodology

This research addresses the issue of handling large sets of biases in CA without hand-coded heuristics. The primary focus is on predicting SAT models from given truth assignments and truth values of the target models. This task involves navigating through a complex landscape of biases and constraints. We train the transformer on a dataset of previously-learned models and examples, aiming to automate the prediction of constraint models, particularly in SAT.

Results

Initial experiments demonstrate transformers can accurately learn constraint models without simplifying assumptions and human intervention. In our experiment, we used two pre-trained models from the HuggingFace transformers library, a standard text-to-text transformer (t5-small) and Longformer Encoder-Decoder (LED) with 60 million and 460 million parameters respectively. The standard t5-model can only process 512 tokens at a time, hence it works well with 3-SAT problem. However, for more complex problems like 4-SAT, an LED model provides more accurate results.

t5-small

The t5-small model was specifically tested on the 3-SAT problem. We used two test-sets comprising 1000 and 10000 Conjunctive Normal Form (CNF) formulas (each formula having a set of 40 clauses distributed across 10 variables). The model demonstrated high efficiency, with all the predicted cnf formulas being equi-satisfiable to their respective ground truth target values.

LED Longformer

The LED model, which has a larger token capacity (16,384 max.) did not exhibit the same performance on the 3-SAT problem. However, it proved to be more effective in solving 4-SAT problem, at which the t5-small model failed to generalize well. In two tests, involving 1000 and 2000 cnf formulas for 4-SAT (each formula having 40 clauses and 10 variables), all the predicted formulas were equi-satisfiable to their respective ground truth target values. This outcome highlights the model's suitability for more complex SAT problems that require a larger token capacity for effective prediction.

Ongoing Experiments

With the promising results that have been achieved, we are now training our models on increasingly large SAT problems, with the aim of scaling up to problems beyond the reach of current methods. We will present the results at the upcoming Bridge event.

Acknowledgments

This work was conducted with the financial support of the Science Foundation Ireland Centre for Research Training in Artificial Intelligence under Grant No. 18/CRT/6223. We would also like to acknowledge the support of the Science Foundation Ireland under Grant No. 12/RC/2289-P2 which is co-funded under the European Regional Development Fund.

References

1. T. P. Pawlak, K. Krawiec. Automatic Synthesis of Constraints from Examples Using Mixed Integer Linear Programming. *European Journal of Operational Research* **261**(3):1141–1157, 2017.
2. S. D. Prestwich. Robust Constraint Acquisition by Sequential Analysis. *24th European Conference on Artificial Intelligence*, 2020.
3. N. Beldiceanu, H. Simonis. ModelSeeker: Extracting Global Constraint Models from Positive Examples. *Data Mining and Constraint Programming, Lecture Notes in Computer Science* vol. 10101, Springer 2016, pp. 77–95.
4. D. C. Tsouros, S. Berden, T. Guns. Guided Bottom-Up Interactive Constraint Acquisition. In Roland H. C. Yap, editor, *29th International Conference on Principles and Practice of Constraint Programming*, Vol. 280 of LIPIcs, Schloss Dagstuhl — Leibniz-Zentrum für Informatik, 2023.
5. M. Kumar, S. Kolb, T. Guns. Learning Constraint Programming Models from Data using Generate-and-Aggregate. CP'22.