Fuchs (1995) – heuristic learned to be chosen with Gen Algo

Denzinger and Schulz (1996) – inference control heuristics for equational deduction. Data from prev proofs, select equations that are likely to be used in new situations. 1st eval fn works by symbolic retrieval of generalized patterns from a kkn base, 2nd eval fn compiles the knowledge into abstract term evaluation trees. **Analyzed proof protocols** by representing knowledge about protocols'n'proofs

Denzinger et al. (1997) – case-based reasoning, similarity concept, cooperation concept, reactive planning; still 'learn' from previous successful proof attempts'

Fuchs (1998) – Learn search-guiding heuristics by employing features in a simple, yet effective manner. Features used to adapts a heuristic to a solved problem. Utilize heuristic profitably for related target problems. **Prediction of usefulness of a fact.**

SNoW Carlsonn et al. (1999) – learning program that can be used as a general purpose mulit-class classifier and is specifically tailored for large number of features. Sparse Network of Winnows (not a typo). Sparse network of sparse linear functions over a pre-defined or incrementally acquired feature space. Several update rules may be used – sparse variations of the Winnow update rule, the Perceptron or Naive Bayes. Multi class learner. Decisions either binary or continuous (confidence in [0, 1]).

Proof General Aspinall (2000) – tool for developing proofs with ITP. Interaction based around proof script (seq of commands sent to ITP). Provides UI.

Mizar proof advisor Urban (2004) – MPTP (Mizar Problems for Theorem Proving) is system described; translates MML into FOL for ATPs and for generating thm proving problems corresponding to Mizar Mathematical Library. Mizar proof advisor used for selecting suitable axioms from the large library for an arbirtrary problem. Feature based ML framework, symbols are the features that characterise formulas. They had 40k targets and about 7k features. **SNoW** learning architecture used mainly (NLP archit, designed for large num of feat and targets).

MizarMode Urban (2006) – Emacs authoring environment. Code-generating Code-Browsing Code-searching methods. Auto gen proof skeletons, semantic browsing of articles, structured viewing, proof advice using **machinee learning tools** like Mizar Proof Advisor.

Lightweight relevance filtering ... Meng and Paulson (2009) – relevance filtering methods, based on counting fn symbols in clauses. Signature based relevance filter. Not exactly ML?...

The use of Data-Mining... Duncan (2007) – evaluate the applicability of data-mining techniques for tactics from large corpuses of proofs. Data mine information to find occuring patterns. Patterns are then evolved into tactics. Variable Length Markov Models used to predict next proof step.

MaLARea Urban (2007) – simple metasystem iteratively combining deductive Automated Reasoning tools (now the E and the SPASS ATP systems) with a machine learning component (now the SNoW system used in the naive

Bayesian learning mode). Intended use – large theories, i.e. large num of problems which in a consistent fashion use many axioms, lemmas, thms, etc. The system works in cycles of thm proving followed by ML from successful proofs, using the learned information to prune the set of available axioms for the next cycle. MPTP challenge - 142/252. Learning could be stated as creating an assoc of some features of the conjecture with proving methods. Features – just symbols appearing in them. "Proving method" – ordering of all av axioms. Goal – given symbols, produces ordering of axioms, according to expected relevancy wrt the set of symbols. Sufficiently simple to implement and quite efficient in the first experiments with thm proving over Mizar library. Deduce, learn, loop implemented via growing axiom set and growing timelimit policy. First try to solve cheaply (min num axioms / most relevant, lowest timelimit). On success, learning performed on the newly available solution and axiom/time limit dropped to min values. No success – increase limits. Details in paper. MLMLMLMLML SNoW used in NB mode, bcs of speed. One training example contains all the symbols of a solved conjecture w/ names of axioms needed. A bayes network is trained. Easier to just relearn every time. Trained classifier is used to prune the axiom set for the next runs – we take all the unsolved conjectures and create a testing example from each by taking all its symbols. The classifier run on this printing (ordered) axioms. This is then used to select req num of axioms. Usage of previous results exists.

SRASS Sutcliffe and Puzis (2007) – selection determined by semantics of the axioms and conjecture, heuristically ordered by a syntactic relevance measure. Many problems more solved. At each iter the process looks for a model of selected axioms and the neg of the conjecture. If no model found, then the conjecture is consequence. Otherwise, then an unselected axiom that is false in the model is moved to the set of selected axioms. Newly selected axiom excludes the model from the models of the selected axioms and neg conj, eventually leading to a situation where there are no models of the selected axioms and the negated conjecture. Unselected axioms selected in decreasing order of usefuleness. Syntactic relevance score for usefulness. Direct relevance is ratio of how many predicates and/or functors the have in common to how many they have overall. Contextual direct relevance uses 'contextual intersection'.

MaLARea SG Urban et al. (2008) – combines model-based and learning based methods for automated reasoning in large theories. The implementation is based on MaLARea. Extended by taking into account semantic relevance of axioms, similar to SRASS. Combined system outperforms both. Three extensions to selection of axioms. 1. check for countersatisfiability in runs where this is probable. Allows for countersatisfiability precheck to detect more cases when more axioms need to be added. 2. Use models found when a problem is found to be countersatisfiable, as an additional criterion for computing axiom relevance. Need to efficiently evaluate formulae in the models. 3. Extend axiom specification using a logical criterion: the set of axioms should exclude as many known models of the negated conjecture as possible. **Weird combination. Review!**

References

Alama, J., Heskes, T., Kühlwein, D., Tsivtsivadze, E., and Urban, J. (2014). Premise selection for mathematics by corpus analysis and kernel methods. *J. Autom. Reasoning*, 52(2):191–213.

Aspinall, D. (2000). Proof General: A generic tool for proof development. In *Tools and Algorithms for Construction and Analysis of Systems, 6th International Conference, TACAS 2000, Held as Part of the European Joint Conferences on the Theory and Practice of Software, ETAPS 2000, Berlin, Germany, March 25 - April 2, 2000, Proceedings*, pages 38–42.

Blanchette, J. C., Greenaway, D., Kaliszyk, C., Kühlwein, D., and Urban, J. (2016). A learning-based fact selector for Isabelle/HOL. *J. Autom. Reasoning*, 57(3):219–244.

Carlsonn, A. J., Cumby, C. M., Rosen, J. L., and Roth, D. (1999). SNoW user guide. Technical report, University of Illinois.

Denzinger, J., Fuchs, M., and Fuchs, M. (1997). High performance atp systems by combining several ai methods. In *Proceedings of the 15th International Joint Conference on Artifical Intelligence - Volume 1*, IJCAI'97, pages 102–107, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Denzinger, J. and Schulz, S. (1996). Learning domain knowledge to improve theorem proving. In *Automated Deduction - CADE-13, 13th International Conference on Automated Deduction, New Brunswick, NJ, USA, July 30 - August 3, 1996, Proceedings*, pages 62–76.

Duncan, H. (2007). *The use of data-mining for the automatic formation of tactics*. PhD thesis, University of Edinburgh, UK.

Fuchs, M. (1995). Learning proof heuristics by adaptive parameters. In *Machine Learning, Proceedings of the Twelfth International Conference on Machine Learning, Tahoe City, California, USA, July 9-12, 1995*, pages 235–243.

Fuchs, M. (1998). A feature-based learning method for theorem proving. In *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence*, AAAI '98/IAAI '98, pages 457–462, Menlo Park, CA, USA. American Association for Artificial Intelligence.

Irving, G., Szegedy, C., Alemi, A. A., Eén, N., Chollet, F., and Urban, J. (2016). DeepMath - deep sequence models for premise selection. In *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pages 2235–2243.

Kaliszyk, C., Chollet, F., and Szegedy, C. (2017). Holstep: A machine learning dataset for higher-order logic theorem proving. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*.

Kaliszyk, C. and Urban, J. (2013). Mizar 40 for mizar 40. *CoRR*, abs/1310.2805.

Kaliszyk, C. and Urban, J. (2014). Learning-assisted automated reasoning with Flyspeck. *J. Autom. Reasoning*, 53(2):173–213.

Kaliszyk, C. and Urban, J. (2015). FEMaLeCoP: Fairly efficient machine learning connection prover. In *Logic for Programming, Artificial Intelligence, and Reasoning - 20th International Conference, LPAR-20 2015, Suva, Fiji, November 24-28, 2015, Proceedings*, pages 88–96.

Kaliszyk, C., Urban, J., Michalewski, H., and Olsák, M. (2018). Reinforcement learning of theorem proving. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada.*, pages 8836–8847.

Kaliszyk, C., Urban, J., and Vyskocil, J. (2014). Machine learner for automated reasoning 0.4 and 0.5. In *4th Workshop on Practical Aspects of Automated Reasoning, PAAR@IJCAR 2014, Vienna, Austria, 2014*, pages 60–66.

Komendantskaya, E., Heras, J., and Grov, G. (2012). Machine learning in Proof General: Interfacing interfaces. In *Proceedings 10th International Workshop On User Interfaces for Theorem Provers, UITP 2012, Bremen, Germany, July 11th, 2012.*, pages 15–41.

Kühlwein, D., Blanchette, J. C., Kaliszyk, C., and Urban, J. (2013). MaSh: Machine Learning for Sledgehammer. In *ITP*.

Kühlwein, D., van Laarhoven, T., Tsivtsivadze, E., Urban, J., and Heskes, T. (2012). Overview and evaluation of premise selection techniques for large theory mathematics. In *Automated Reasoning - 6th International Joint Conference, IJCAR 2012, Manchester, UK, June 26-29, 2012. Proceedings*, pages 378–392.

Loos, S. M., Irving, G., Szegedy, C., and Kaliszyk, C. (2017). Deep network guided proof search. In *LPAR-21, 21st International Conference on Logic for Programming, Artificial Intelligence and Reasoning, Maun, Botswana, May 7-12, 2017*, pages 85–105.

Meng, J. and Paulson, L. C. (2009). Lightweight relevance filtering for machine-generated resolution problems. *J. Applied Logic*, 7(1):41–57.

Rocktäschel, T. and Riedel, S. (2017). End-to-end differentiable proving. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 3788–3800.

Sutcliffe, G. and Puzis, Y. (2007). SRASS - A semantic relevance axiom selection system. In *Automated Deduction - CADE-21, 21st International Conference on Automated Deduction, Bremen, Germany, July 17-20, 2007, Proceedings*, pages 295–310.

Urban, J. (2004). MPTP - motivation, implementation, first experiments. *J. Autom. Reasoning*, 33(3-4):319–339.

Urban, J. (2006). MizarMode - an integrated proof assistance tool for the mizar way of formalizing mathematics. *J. Applied Logic*, 4(4):414–427.

Urban, J. (2007). MaLARea: a metasystem for automated reasoning in large theories. In *Proceedings of the CADE-21 Workshop on Empirically Successful Automated Reasoning in Large Theories, Bremen, Germany, 17th July 2007*.

Urban, J., Sutcliffe, G., Pudlák, P., and Vyskočil, J. (2008). MaLARea SG1 - machine learner for automated reasoning with semantic guidance. In Armando, A., Baumgartner, P., and Dowek, G., editors, *Automated Reasoning*, pages 441–456, Berlin, Heidelberg. Springer Berlin Heidelberg.

Urban, J., Vyskocil, J., and Stepánek, P. (2011). MaLeCoP machine learning connection prover. In *Automated Reasoning with Analytic Tableaux and Related Methods - 20th International Conference, TABLEAUX 2011, Bern, Switzerland, July 4-8, 2011. Proceedings*, pages 263–277.