Report from Dagstuhl Seminar 21192

Approaches and Applications of Inductive Programming

Edited by

Andrew Cropper[1], Luc De Raedt[2], Richard Evans[3], and Ute Schmid[4]

1 **University of Oxford, GB,** andrew.cropper@cs.ox.ac.uk
2 **KU Leuven, BE,** luc.deraedt@cs.kuleuven.be
3 **DeepMind – London, GB,** richardevans@google.com
4 **Universität Bamberg, DE,** ute.schmid@uni-bamberg.de

Abstract

In this report the program and the outcomes of Dagstuhl Seminar 21192 "Approaches and Applications of Inductive Programming" is documented. The goal of inductive programming (IP) is to induce computer programs from data, typically input/output examples of a desired program. IP interests researchers from many areas of computer science, including machine learning, automated reasoning, program verification, and software engineering. Furthermore, IP contributes to research outside computer science, notably in cognitive science, where IP can help build models of human inductive learning and contribute methods for intelligent tutor systems. Building on the success of previous IP Dagstuhl seminars (13502, 15442, 17382, and 19202), the goal of this new edition of the seminar is to focus on IP methods which integrate learning and reasoning, scaling up IP methods to be applicable to more complex real world problems, and to further explore the potential of IP for explainable artificial intelligence (XAI), especially for interactive learning. The extended abstracts included in this report show recent advances in IP research. The included short report of the outcome of the discussion sessions additionally point out interesting interrelation between different aspects and possible new directions for IP.

Seminar May 9–12, 2021 – http://www.dagstuhl.de/21192

2012 ACM Subject Classification Computing methodologies → Artificial intelligence; Computing methodologies → Machine learning; Software and its engineering → Compilers; Human-centered computing → Human computer interaction (HCI)

Keywords and phrases Interpretable Machine Learning, Explainable Artificial Intelligence, Interactive Learning, Human-like Computing, Inductive Logic Programming

Digital Object Identifier 10.4230/DagRep.11.4.20

Edited in cooperation with Jaimovitch López, Gonzalo

1 Executive Summary

Andrew Cropper
Luc De Raedt
Richard Evans
Ute Schmid

The goal of Inductive Programming (IP) is to provide methods for induction of computer programs from data. Specifically, IP is the automated (or semi-automated) generation of a computer program from an incomplete information, such as input-output examples, demonstrations, or computation traces.IP offers powerful approaches to learning from relational data and to learning from observations in the context of autonomous intelligent agents. IP is a form of machine learning, because an IP system should perform better given more data (i.e. more examples or experience). However, in contrast to standard ML approaches, IP approaches typically only need a small number of training examples. Furthermore, induced hypotheses are typically represented as logic or functional programs, and can therefore be inspected by a human. In that sense, IP is a type of interpretable machine learning which goes beyond the expressivity of other approaches of rule learning such as decision tree algorithms. IP is also a form of program synthesis. It complements deductive and transformational approaches. When specific algorithm details are difficult to determine, IP can be used to generate candidate programs from either user-provided data, such as test cases, or from data automatically derived from a formal specification. Most relevant application areas of IP techniques is end-user programming and data wrangling.

This seminar has been the fifth in a series – building on seminars 13502, 15442, 17383, and 19202. In the wake of the recent interest in deep learning approaches, mostly for end-to-end learning, it has been recognized that for practical applications, especially in critical domains, data-intensive blackbox machine learning must be complemented with methods which can help to overcome problems with data quality, missing or errouneous labeling of training data, as well as providing transparency and comprehensibility of learned models. To address these requirements, on the one hand, explainable artificial intelligence (XAI) emerged as a new area of research and on the other hand, there is a new interest in bringing together learning and reasoning. These two areas of research are in the focus of the 2021 seminar. Futhermore, recent developments to scale up IP methods to be more applicable to complex real world domains has been taken into account. Based on outcomes of the fourth seminar (19202), the potential of IP as powerful approach for explainable artificial intelligence ("IP for XAI") has been be elaborated. Bringing together IP methods and deep learning approaches contributes to neural-symbolic intergration research. While two years ago (seminar 19202) focus has been on IP as interpretable surrogate model, in the 2021 seminar explainability of different addressees of explanations and their need to different types of explanations (e.g. verbal or example-based) are considered. For many real world applications, it is necessary to involve the human as teacher and judge for the machine learned models. Therefore, a further topic of the seminar has been to explore IP in the context of new approaches to interactive ML and their applications to automating data science and joint human-computer decision making.

2 Table of Contents

3 Overview of Talks

3.1 Beneficial and harmful explanatory machine learning

Lun Ai (Imperial College London, GB), Mark Gromowski, Céline Hocquette (Imperial College London, GB), Stephen H. Muggleton (Imperial College London, GB), and Ute Schmid (Universität Bamberg, DE)

Main reference Lun Ai, Stephen H. Muggleton, Céline Hocquette, Mark Gromowski, Ute Schmid: "Beneficial and harmful explanatory machine learning", Mach. Learn., Vol. 110(4), pp. 695–721, 2021.
URL https://doi.org/10.1007/s10994-020-05941-0

Given the recent successes of Deep Learning in AI there has been increased interest in the role and need for explanations in machine learned theories. A distinct notion in this context is that of Michie's definition of ultra-strong machine learning (USML). USML is demonstrated by a measurable increase in human performance of a task following provision to the human of a symbolic machine learned theory for task performance. A recent paper demonstrates the beneficial effect of a machine learned logic theory for a classification task, yet no existing work to our knowledge has examined the potential harmfulness of machine's involvement for human comprehension during learning. This paper investigates the explanatory effects of a machine learned theory in the context of simple two person games and proposes a framework for identifying the harmfulness of machine explanations based on the Cognitive Science literature. The approach involves a cognitive window consisting of two quantifiable bounds and it is supported by empirical evidence collected from human trials. Our quantitative and qualitative results indicate that human learning aided by a symbolic machine learned theory which satisfies a cognitive window has achieved significantly higher performance than human self learning. Results also demonstrate that human learning aided by a symbolic machine learned theory that fails to satisfy this window leads to significantly worse performance than unaided human learning.

3.2 A Declarative Framework for Knowledge-Based Explainable Link Analysis

Martin Atzmüller (Universität Osnabrück, DE)

Joint work of Martin Atzmüller, Cicek Guven, Dietmar Seipel

Generating explanations is a prominent topic in artificial intelligence and data science, in order to make methods and systems more transparent, interpretable and understandable for humans. We focus on link analysis: Here, link prediction and anomalous link discovery are challenging, e.g., in cold-start scenarios or when only sparse historic data is available [2, 3]. We discuss how to apply answer set programming (ASP) in a declarative framework for (1) formalizing knowledge-augmented link analysis in feature-rich networks [3], with (2) explanation generation using ASP [1, 4]. We exemplify this via simple link predictors on real-world network datasets.

References

1 Martin Atzmueller, Cicek Guven, Dietmar Seipel (2019) Towards Generating Explanations for ASP-Based Link Analysis using Declarative Program Transformations. Proc. International Conference on Applications of Declarative Programming and Knowledge Management (INAP 2019). CoRR abs/1909.03404
2 Martin Atzmueller and Cicek Guven (2019) A Multi-Perspective View on Model-Based Exceptional Link Analysis on Complex Interaction Networks, NPCS
3 Cicek Guven, Martin Atzmueller (2019) Applying Answer Set Programming for Knowledge-Based Link Prediction on Social Interaction Networks. Frontiers in Big Data
4 Cicek Guven, Dietmar Seipel, Martin Atzmueller (2020) Applying ASP for Knowledge-Based Link Prediction with Explanation Generation in Feature Rich Networks. IEEE Transactions on Network Science and Engineering 8(2):1305-1315

3.3 Inductively inferring human problem solving strategies from observed behavior

Thea Behrens (TU Darmstadt, DE) and Frank Jäkel (TU Darmstadt, DE)

When people solve Sudokus, they can apply many different inference rules. In a series of think-aloud studies we inferred these rules from participants' behavior and their verbalizations manually and implemented them as Prolog programs. In our studies just one general rule would have been enough to fill all cells of a Sudoku puzzle, but we still saw a lot of rule variability and flexibility in our participants. From the data we could estimate the preferences for each rule and each participant. We found that these preferences differ markedly between participants. The rules and rule preferences together form a probabilistic program that is a good description of a participant's problem solving strategy. Unfortunately, in our studies the observed behavior alone was not enough to allow us to infer the rules that participants used and we had to rely on think-aloud data. These natural language data are too unstructured to serve as input to available inductive programming systems. Therefore, we developed a user-interface for solving Sudokus that elicits all the information that participants use when they apply a rule. Our hope is that these new data will allow for a more formal approach to infer the rules that underlie the observed behavior.

3.4 Abductive Knowledge Induction from Raw Data

Wang-Zhou Dai (Imperial College London, GB) and Stephen H. Muggleton (Imperial College London, GB)

Main reference Wang-Zhou Dai, Stephen H. Muggleton: "Abductive Knowledge Induction From Raw Data", CoRR, Vol. abs/2010.03514, 2020.
URL https://arxiv.org/abs/2010.03514

For many reasoning-heavy tasks involving raw inputs, it is challenging to design an appropriate end-to-end learning pipeline. Neuro-Symbolic Learning, divide the process into sub-symbolic perception and symbolic reasoning, trying to utilise data-driven machine learning and knowledge-driven reasoning simultaneously. However, they suffer from the exponential computational complexity within the interface between these two components, where the sub-symbolic learning model lacks direct supervision, and the symbolic model lacks accurate input facts. Hence, most of them assume the existence of a strong symbolic knowledge base and only learn the perception model while avoiding a crucial problem: where does the knowledge come from? In this paper, we present Abductive Meta-Interpretive Learning (MetaAbd) that unites abduction and induction to learn neural networks and induce logic theories jointly from raw data. Experimental results demonstrate that MetaAbd not only outperforms the compared systems in predictive accuracy and data efficiency but also induces logic programs that can be re-used as background knowledge in subsequent learning tasks. To the best of our knowledge, MetaAbd is the first system that can jointly learn neural networks from scratch and induce recursive first-order logic theories with predicate invention.

3.5 From Statistical Relational to Neuro-Symbolic AI

Luc De Raedt (KU Leuven, BE)

Joint work of Luc De Raedt, Sebastijan Dumancic, Robin Manhaeve, Giuseppe Marra
Main reference Luc De Raedt, Sebastijan Dumancic, Robin Manhaeve, Giuseppe Marra: "From Statistical Relational to Neuro-Symbolic Artificial Intelligence", in Proc. of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020, pp. 4943–4950, ijcai.org, 2020.
URL https://doi.org/10.24963/ijcai.2020/688

Neural-symbolic and statistical relational artificial intelligence both integrate frameworks for learning with logical reasoning. This survey identifies several parallels across seven different dimensions between these two fields. These cannot only be used to characterize and position neural-symbolic artificial intelligence approaches but also to identify a number of directions for further research.

3.6 Knowledge Refactoring for Inductive Program Synthesis

Sebastijan Dumancic (KU Leuven, BE), Andrew Cropper (University of Oxford, GB)

Joint work of Sebastijan Dumancic, Andrew Cropper, Tias Guns
Main reference Sebastijan Dumancic, Tias Guns, Andrew Cropper: "Knowledge Refactoring for Inductive Program Synthesis", Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35(8), pp. 7271–7278, 2021.
URL https://ojs.aaai.org/index.php/AAAI/article/view/16893

Humans constantly restructure knowledge to use it more efficiently. Our goal is to give a machine learning system similar abilities so that it can learn more efficiently. We introduce the knowledge refactoring problem, where the goal is to restructure a learner's knowledge base to reduce its size and to minimise redundancy in it. We focus on inductive logic programming, where the knowledge base is a logic program. We introduce Knorf, a system that solves the refactoring problem using constraint optimisation. We evaluate our approach on two program induction domains: real-world string transformations and building Lego structures. Our experiments show that learning from refactored knowledge can improve predictive accuracies fourfold and significantly reduce learning times.

3.7 On Conditional Teaching Size and Minimal Curricula

Manuel Garcia-Piqueras (University of Castilla-La Mancha, ES) and José Hernández-Orallo (Technical University of Valencia, ES)

Main reference Jan Arne Telle, José Hernández-Orallo, Cèsar Ferri: "The teaching size: computable teachers and learners for universal languages", Mach. Learn., Vol. 108(8-9), pp. 1653–1675, 2019.
URL https://doi.org/10.1007/s10994-019-05821-2

Machine teaching, under certain integrative prior knowledge, enables the instruction of any concept expressed in a universal language. Latest experiments show that there are instructional sets surprisingly shorter than the concept description itself [1]. We delineate a border for these remarkable experimental findings through *teaching size* and concept complexity. Also, we study teaching curricula and find a new phenomenon that we call *interposition*: certain prior knowledge generates simpler compatible concepts which increase the teaching size of the concept that we want to teach. Far beyond, we provide an algorithm which builds *optimal curricula* based on interposition avoidance. These results reveal innovative curriculum design strategies for machines, but also for animals and humans.

References

1 Telle, Jan Arne and Hernández-Orallo, José and Ferri, Cèsar. The teaching size: computable teachers and learners for universal languages, Machine Learning, 108(8), 1653–1675 (2019).

3.8 IP vs Humans: Learning from Machine Teaching Examples

Gonzalo Jaimovitch López (Technical University of Valencia, ES), Cesar Ferri Ramirez (Technical University of Valencia, ES), and José Hernández-Orallo (Technical University of Valencia, ES)

Main reference Gonzalo Eduardo Jaimovitch López: "Comparison between machine learning and human learning from examples generated with machine teaching", 2020.
URL http://hdl.handle.net/10251/152771

Inductive programming has been singled out as one important approach to learning, where background knowledge and simplicity priors together play a key role to infer patterns, including algorithmic ones, from very few examples. Other machine learning techniques, especially deep learning, require thousands, if not millions of examples, to reach the same inference. This duality seems to be challenged by new massive models based on transformers that are able to be pretrained from large datasets. These models capture vast amounts of knowledge with very abstract representations and then make inferences from very few examples, with little tuning or no retraining needed. In particular, large language models have shown an impressive ability for few-shot learning. It seems relevant to ask now what kind of patterns these models can capture and how many examples they need in their prompts. We present this question as a machine teaching problem with strong priors [1, 2], and test whether language models can learn simple algorithmic concepts from small witness sets. In particular, we explore how several GPT architectures, inductive programming systems (the inductive functional programming system MagicHaskeller and the inductive logic programming system Louise) and humans perform in terms of the complexity of the concept and the number of examples provided, and how much their behaviour diverge [3]. This first joint analysis of machine teaching and language models can address key questions for artificial intelligence and machine learning, such as whether strong priors, and Occam's razor in particular, can be distilled from data, making learning from a few examples possible without the need of providing domain knowledge or common sense knowledge about the world.

References

1 Telle, J. A., Hernández-Orallo, J., & Ferri, C. (2019). *The teaching size: computable teachers and learners for universal languages.* Machine Learning, 108(8), 1653-1675.
2 Hernández-Orallo, J., & Telle, J. A. (2020). *Finite and Confident Teaching in Expectation: Sampling from Infinite Concept Classes.* In 24th European Conference on Artificial Intelligence (ECAI2020)
3 Jaimovitch López, G. E. (2020). *Comparison between machine learning and human learning from examples generated with machine teaching* (Degree disertation). Universitat Politècnica de València

3.9 Analyzing massive biomedical datasets with graph-based rule mining for drug repurposing

Tomáš Kliegr (University of Economics – Prague, CZ)

Main reference Václav Zeman, Tomáš Kliegr, Vojtech Svátek: "RDFRules: Making RDF rule mining easier and even more efficient", Semantic Web, Vol. 12(4), pp. 569–602, 2021.
URL https://doi.org/10.3233/SW-200413

Large biomedical datasets containing structured data could contain hidden knowledge useful for the response to the pandemic. In this talk, we present our attempt to apply a graph-based rule mining system to the KG-Covid-19 dataset [2]. This RDF knowledge graph is a result of ingestion of multiple specialized knowledge sources, such as DrugBank, as well as extraction from COVID-19 research literature selected for their relevance for drug repurposing efforts. The published version of the dataset contains 377,482 nodes and 21,433,063 edges. Due to the large size of the dataset, we initially chose the AMIE+ rule mining algorithm, which was shown to be orders of magnitude faster than the previous approaches. Due to the need to search for highly specific patterns, our initial attempts to apply "vanilla" AMIE+ were not successful due to the combinatorial explosion associated with mining of low support rules and associated memory issues. We also experimented with AnyBURL [3], which had excellent performance, but also lacked the possibility to finely define the sought pattern. We finally settled on RDFRules [5], which is a comprehensive set of extensions for the AMIE+ that also includes the possibility to apply fine-grained patterns for constraining the search space. Using RDFRules, it was possible to find rules with low support even on the full KG-Covid-19 dataset (without metadata) on a single computer using less than 64 GB of RAM. As an example use case, we described a mining task performed in Summer 2020 to Fall 2020, using a Spring 2020 release of KG-Covid-19 as described in detail in [4]. This task aimed to find drugs that "molecularly interact with" the ACE 1 receptor (UniProtKB ID P12821) and at the same time they are connected through an arbitrary predicate to an intermediary resource, which is using the "interact with" predicate connected to the ACE 2 receptor (UniProtKB ID Q9BYF1). Note that "molecularly interact with" and "interact with" are Biolink Model predicates. The task was thus to find rules complying to the following RDFRules pattern: (<Any> <interacts_with> <Q9BYF1>) ∧ [(<Any>

<Any> <Any>)] ⇒ (<Any> <molecularly_interacts_with> <P12821>). This task executed with minsupp= 1 and mining with constants led to the discovery of a single logical rule (?b <interacts_with> <Q9BYF1>) ∧ (?a <molecularly_interacts_with> ?b) → (?a <molecularly_interacts_with> <P12821>), whose instantiation resulted in five drugs. These included a widely used antihypertensive drug "Telmisartan" for which a subsequent literature search showed that it is a viable drug repurposing target. A recent open multicenter randomized clinical trial has shown that Telmisartan could, through anti-inflammatory effects, reduce mortality and morbidity in hospitalized patients infected with SARS-CoV-2 [1]. We conclude that rule mining is a viable approach for finding "nuggets" in large knowledge graphs. Our work was limited in that we have not explored the possibility to use the embeddings-based approaches and we have not performed comparison with more direct methods of analyzing graph data.

References

1 Duarte, M., Pelorosso, F., Nicolosi, L. N., Salgado, M. V., Vetulli, H., Aquieri, A., ... & Rothlin, R. P. (2021). *Telmisartan for treatment of Covid-19 patients: An open multicenter randomized clinical trial.* EClinicalMedicine, 37, 100962.
2 Reese, J. T., Unni, D., Callahan, T. J., Cappelletti, L., Ravanmehr, V., Carbon, S., ... & Mungall, C. J. (2021). *KG-COVID-19: A framework to produce customized knowledge graphs for COVID-19 response.* Patterns, 2(1), 100155.
3 Meilicke, C., Chekol, M. W., Ruffinelli, D., & Stuckenschmidt, H. (2019). *Anytime Bottom-Up Rule Learning for Knowledge Graph Completion.* In IJCAI (pp. 3137-3143).
4 Šimečková, J. *Extrakce pravidel ze znalostních grafů.* Bachelor thesis. University of Economics, Prague, 2020
5 Zeman, Václav, Tomáš Kliegr, and Vojtěch Svátek. *RDFRules: Making RDF rule mining easier and even more efficient.* Semantic Web. 2021

3.10 Towards Robust, Data-Efficient, and Explainable Deep Learning

Pasquale Minervini (University College London, GB)

Deep Learning models are a class of Machine Learning models that use multiple processing layers to progressively extract higher-level features from raw inputs. Over the past decade, it has become one of the most impactful research areas in Artificial Intelligence, with many notable commercially important applications. However, Deep Learning models still fall short in terms of data efficiency, out-of-distribution generalisation, interpretability, and complexity. We discuss several ways of overcoming such limitations, by increasing their statistical robustness [1, 2, 3], incorporating prior knowledge [4, 5, 6], combining symbolic and sub-symbolic computation models [7, 8, 9, 10], and developing more computationally efficient neural models [11, 12, 13].

References

1 Joe Stacey, Pasquale Minervini, Haim Dubossarsky, Sebastian Riedel, Tim Rocktäschel: Avoiding the Hypothesis-Only Bias in Natural Language Inference via Ensemble Adversarial Training. EMNLP (1) 2020: 8281-8291
2 Johannes Welbl, Pasquale Minervini, Max Bartolo, Pontus Stenetorp, Sebastian Riedel: Undersensitivity in Neural Reading Comprehension. EMNLP (Findings) 2020: 1152-1165

3 Oana-Maria Camburu, Brendan Shillingford, Pasquale Minervini, Thomas Lukasiewicz, Phil Blunsom: Make Up Your Mind! Adversarial Generation of Inconsistent Natural Language Explanations. ACL 2020: 4157-4165
4 Pasquale Minervini, Thomas Demeester, Tim Rocktäschel, Sebastian Riedel: Adversarial Sets for Regularising Neural Link Predictors. UAI 2017
5 Pasquale Minervini, Luca Costabello, Emir Muñoz, Vít Novácek, Pierre-Yves Vandenbussche: Regularizing Knowledge Graph Embeddings via Equivalence and Inversion Axioms. ECML/PKDD (1) 2017: 668-683
6 Pasquale Minervini, Sebastian Riedel: Adversarially Regularising Neural NLI Models to Integrate Logical Background Knowledge. CoNLL 2018: 65-74
7 Leon Weber, Pasquale Minervini, Jannes Münchmeyer, Ulf Leser, Tim Rocktäschel: NLProlog: Reasoning with Weak Unification for Question Answering in Natural Language. ACL (1) 2019: 6151-6161
8 Pasquale Minervini, Matko Bosnjak, Tim Rocktäschel, Sebastian Riedel, Edward Grefenstette: Differentiable Reasoning on Large Knowledge Bases and Natural Language. AAAI 2020: 5182-5190
9 Pasquale Minervini, Sebastian Riedel, Pontus Stenetorp, Edward Grefenstette, Tim Rocktäschel: Learning Reasoning Strategies in End-to-End Differentiable Proving. ICML 2020: 6938-6949
10 Erik Arakelyan, Daniel Daza, Pasquale Minervini, Michael Cochez: Complex Query Answering with Neural Link Predictors. ICLR 2021
11 Yuxiang Wu, Sebastian Riedel, Pasquale Minervini, Pontus Stenetorp: Don't Read Too Much Into It: Adaptive Computation for Open-Domain Question Answering. EMNLP (1) 2020: 3029-3039
12 Yuxiang Wu, Pasquale Minervini, Pontus Stenetorp, Sebastian Riedel: Training Adaptive Computation for Open-Domain Question Answeringwith Computational Constraints. ACL 2021
13 Patrick S. H. Lewis, Yuxiang Wu, Linqing Liu, Pasquale Minervini, Heinrich Küttler, Aleksandra Piktus, Pontus Stenetorp, Sebastian Riedel: PAQ: 65 Million Probably-Asked Questions and What You Can Do With Them. CoRR abs/2102.07033 (2021)

3.11 Advances in Meta-Interpretive Learning/ILP and Cognitive Artificial Intelligence

Stephen H. Muggleton (Imperial College London, GB)

Joint work of Stephen Muggleton, Andrew Cropper, Sebastijan Dumancic, Richard Evans
Main reference Andrew Cropper, Sebastijan Dumancic, Stephen H. Muggleton: "Turning 30: New Ideas in Inductive Logic Programming", in Proc. of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020, pp. 4833–4839, ijcai.org, 2020.
URL https://doi.org/10.24963/ijcai.2020/673

Inductive logic programming (ILP) is a form of logic-based machine learning. The goal of ILP is to induce a hypothesis (a logic program) that generalises given training examples and background knowledge. As ILP turns 30, we survey recent work in the field. In this survey, we focus on (i) new meta-level search methods, (ii) techniques for learning recursive programs that generalise from few examples, (iii) new approaches for predicate invention, and (iv) the use of different technologies, notably answer set programming and neural networks. We conclude by discussing some of the current limitations of ILP and discuss directions for future research.

3.12 Generating Contrastive Explanations for Inductive Logic Programming Based on a Near Miss Approach

Johannes Rabold (Universität Bamberg, DE), Ute Schmid (Universität Bamberg, DE)

Joint work of Johannes Rabold, Ute Schmid, Michael Siebers

In recent research, human-understandable explanations of machine learning models have received a lot of attention. Often explanations are given in the form of model simplifications or visualizations. However, as shown in cognitive science as well as in early AI research, concept understanding can also be improved by aligning a given instance for a concept with a similar counterexample. Contrasting a given instance with a structurally similar example which does not belong to the concept highlights what characteristics are necessary for concept membership. Such near misses have been proposed by Winston (1970) as efficient guidance for learning in relational domains. We introduce an explanation generation algorithm for relational concepts learned with Inductive Logic Programming (GeNME). The algorithm identifies near miss examples from a given set of instances and ranks them by their degree of closeness to a specific positive instance. A modified rule which covers the near miss but not the original instance is given as an explanation. We illustrate GeNME with the well known family domain consisting of kinship relations, the visual relational Winston arches domain and a real-world domain dealing with file management. We also present a psychological experiment comparing human preferences of rule-based, example-based, and near miss explanations in the family and the arches domains.

3.13 Learning Episodic Memory Retrieval Procedures Using First-Order Ripple-Down Rules

Claude Sammut (UNSW – Sydney, AU)

Joint work of Claude Sammut, Colm Flanagan, Eric Martin, Michael Bain

The long term memory of an agent can be separated into three categories: procedural, declarative and episodic [4]. The distinguishing features of episodic memories is that they store knowledge of specific events, including contextual information such, time, location and participating agents. It is useful for an agent, such as a robot, to be able to recall past events that are similar to one that has just been observed. The problem here is to define similarity and how to retrieve similar events. A typical approach taken in many case-based reasoning systems [3, 6, 5] is to create a geometric distance measure, where the dimensions in the event space are the observed features. Devising such as measure becomes difficult when the observations are complex, as is the case for a robot operating a complex environment such as a real home or work place or in search and rescue operations. Flanagan [2] describes a system that learns as matching procedure that is customised for each type of event and is capable of matching structured object descriptions. It associates a Ripple-Down Rule [1]

References

1 Compton, P. and Kang, B. (2021). *RIPPLE-DOWN RULES: The Alternative to Machine Learning*. Chapman and Hall/CRC.
2 Flanagan, C. (2021). *Episodic and Semantic Memory in Cognitive Robots*. Phd thesis, School of Computer Science and Engineering, The University of New South Wales.
3 Kolodner, J. (1993). *Case-Based Reasoning*. Morgan Kaufmann, San Mateo.
4 Laird, J. E., Kinkade, K. R., Mohan, S., and Xu, J. Z. (2012). Cognitive Robotics Using the Soar Cognitive Architecture. In *8th International Conference on Cognitive Robotics, (Cognitive Robotics Workshop, Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*, pages 46–54, Toronto. ACM Press.
5 Riesbeck, C. and Schank, R. (1989). *Inside Case-based Reasoning*. Erlbaum, Northvale, NJ.
6 Veloso, M. and Aamodt, A., editors (1995). *Case-Based Reasoning Research and Development: Proceedings of the First International Conference on Case-Based Reasoning*, Berlin. Springer Verlag.

3.14 Ultra-strong machine learning with explanatory dialogs

Ute Schmid (Universität Bamberg, DE)

Joint work of Bettina Finzel, Ute Schmid

At Dagstuhl AAIP 2017, Stephen Muggelton reminded us about Donald Michie's criteria which require a machine learning system (1) to show improved predictive performance with increasing amounts of data (weak criterion), (2) additionally to present the learned model in symbolic –human understandable– form (strong criterion), which (3) teach the model to a human in such a way that the performance is increased to a level beyond that of the human studying the training data alone (ultra-strong criterion). Michie's ultra-strong criterion corresponds to the main claim of current research on explainable AI (XAI). Our discussion at Dagstuhl inspired us to empirically research whether models learned with inductive logic programming (ILP) fulfill the ultra-strong criterion [1, 2]. We extended this work to applying ILP to teaching best moves in game playing be generating verbal explanations from the learned Prolog rules [3]. However, in this work, explanations are given once and in one specific way. In contrast, when one human teaches another, explanations are often a process based on a dialog between generator and receiver of the explanation. Currently, we realize an approach to generate such explanatory dialogs from reasoning traces from Prolog and combine such verbal explanations with prototypes and near miss examples.

References

1 Ute Schmid, Christina Zeller, Tarek R. Besold, Alireza Tamaddoni-Nezhad, Stephen Muggleton: How Does Predicate Invention Affect Human Comprehensibility? ILP 2016: 52-67
2 Stephen H. Muggleton, Ute Schmid, Christina Zeller, Alireza Tamaddoni-Nezhad, Tarek R. Besold: Ultra-Strong Machine Learning: comprehensibility of programs learned with ILP. Mach. Learn. 107(7): 1119-1140 (2018)
3 Lun Ai, Stephen H. Muggleton, Céline Hocquette, Mark Gromowski, Ute Schmid: Beneficial and harmful explanatory machine learning. Mach. Learn. 110(4): 695-721 (2021)

3.15 Generalized Planning as Heuristic Search

Javier Segovia-Aguas (UPF – Barcelona, ES)

Joint work of Javier Segovia-Aguas, Sergio Jiménez, Anders Jonsson
Main reference Javier Segovia Aguas, Sergio Jiménez Celorrio, Anders Jonsson: "Generalized Planning as Heuristic Search", CoRR, Vol. abs/2103.14434, 2021.
URL https://arxiv.org/abs/2103.14434

Generalized planning is a booming research topic in the automated planning community, which aims at computing algorithm-like plans, e.g. plans that branch and loop, that solve a possibly infinite set of planning instances of a given domain. Generalized planning is then a fascinating meeting point for automated planning and program synthesis since both pursue a well-founded integration of (i), knowledge representation with human-comprehensible models (ii), model-based reasoning and (iii), the learning of such models from examples. In this work we show that the heuristic search paradigm, that has traditionally shown successful for classical planning, applies also to the computation of generalized plans using a Random Access Machine, a best-first search algorithm, and different evaluation/heuristic functions for guiding the search in a tractable (though combinatorial) solution space. We believe this is a promising research direction for achieving a tighter integration of the representation, reasoning and learning facets in Artificial Intelligence.

Remote Participants

- Lun Ai
Imperial College London, GB
- Nada Amin
Harvard University – Allston, US
- Martin Atzmüller
Universität Osnabrück, DE
- Feryal Behbahani
DeepMind – London, GB
- Thea Behrens
TU Darmstadt, DE
- Andrew Cropper
University of Oxford, GB
- James Cussens
University of Bristol, GB
- Wang-Zhou Dai
Imperial College London, GB
- Luc De Raedt
KU Leuven, BE
- Thomas Demeester
Ghent University, BE
- Amit Dhurandhar
IBM TJ Watson Research Center – Yorktown Heights, US
- Sebastijan Dumancic
KU Leuven, BE
- Kevin Ellis
Cornell University – Ithaca, US
- Richard Evans
DeepMind – London, GB
- Cesar Ferri Ramirez
Technical University of Valencia, ES
- Bettina Finzel
Universität Bamberg, DE
- Peter Flach
University of Bristol, GB
- Johannes Fürnkranz
Johannes Kepler Universität Linz, AT
- Artur Garcez
City – University of London, GB
- Manuel Garcia-Piqueras
University of Castilla-La Mancha, ES
- José Hernández-Orallo
Technical University of Valencia, ES
- Céline Hocquette
Imperial College London, GB
- Frank Jäkel
TU Darmstadt, DE
- Gonzalo Jaimovitch López
Technical University of Valencia, ES
- Susumu Katayama
University of Miyazaki, JP
- Tomáš Kliegr
University of Economics – Prague, CZ
- Stefan Kramer
Universität Mainz, DE
- Maithilee Kunda
Vanderbilt University, US
- Sara Magliacane
University of Amsterdam, NL
- Roman Manevich
Facebook – London, GB
- Fernando Martinez-Plumed
European Commission – Sevilla, ES
- Pasquale Minervini
University College London, GB
- Stephen H. Muggleton
Imperial College London, GB
- Stassa Patsantzis
Imperial College London, GB
- Johannes Rabold
Universität Bamberg, DE
- Claude Sammut
UNSW – Sydney, AU
- Stephan Scheele
Universität Bamberg, DE
- Ute Schmid
Universität Bamberg, DE
- Javier Segovia-Aguas
UPF – Barcelona, ES
- Gustavo Soares
Microsoft Corporation – Redmond, US
- Stefano Teso
University of Trento, IT
- Jan Tinapp
bidt – München, DE