Report from Dagstuhl Seminar 14381

Neural-Symbolic Learning and Reasoning

Edited by

Artur d'Avila Garcez[1], Marco Gori[2], Pascal Hitzler[3], and Luís C. Lamb[4]

1 **City University London, GB,** aag@soi.city.ac.uk
2 **University of Siena, IT,** marcoxgori@gmail.com
3 **Wright State University – Dayton, US,** pascal.hitzler@wright.edu
4 **Federal University of Rio Grande do Sul – Porto Alegre, BR,** LuisLamb@acm.org

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 14381 "Neural-Symbolic Learning and Reasoning", which was held from September 14th to 19th, 2014. This seminar brought together specialist in machine learning, knowledge representation and reasoning, computer vision and image understanding, natural language processing, and cognitive science. The aim of the seminar was to explore the interface among several fields that contribute to the effective integration of cognitive abilities such as learning, reasoning, vision and language understanding in intelligent and cognitive computational systems. The seminar consisted of contributed and invited talks, breakout and joint group discussion sessions.

Seminar September 14–19, 2014 – http://www.dagstuhl.de/14381

1998 ACM Subject Classification I.2 Artificial Intelligence, I.2.4 Knowledge Representation Formalisms and Methods, I.2.6 Learning, I.2.10 Vision and Scene Understanding, I.2.11 Distributed Artificial Intelligence

Keywords and phrases Neural-symbolic computation, deep learning, image understanding, lifelong machine learning, natural language understanding, ontology learning

Digital Object Identifier 10.4230/DagRep.4.9.50

1 Executive Summary

Artur S. d'Avila Garcez
Marco Gori
Pascal Hitzler
Luís C. Lamb

Neural-symbolic computation aims at building rich computational models and systems through the integration of connectionist learning and sound symbolic reasoning [1, 2]. Over the last three decades, neural networks were shown effective in the implementation of robust large-scale experimental learning applications. Logic-based, symbolic knowledge representation and reasoning have always been at the core of Artificial Intelligence (AI) research. More recently, the use of deep learning algorithms have led to notably efficient applications, with performance comparable to those of humans, in particular in computer image and vision understanding and natural language processing tasks [3, 4, 5]. Further, advances in fMRI allow scientists to grasp a better understanding of neural functions, leading to realistic neural-computational models. Therefore, the gathering of researchers from several communities seems fitting at this stage of the research in neural computation and machine learning, cognitive science, applied logic, and visual information processing. The seminar was an appropriate meeting for the discussion of relevant issues concerning the development of rich intelligent systems and models, which can, for instance integrate learning and reasoning or learning and vision. In addition to foundational methods, algorithms and methodologies for neural-symbolic integration, the seminar also showcase a number of applications of neural-symbolic computation.

The meeting also marked the 10th anniversary of the workshop series on neural-symbolic learning and reasoning (NeSy), held yearly since 2005 at IJCAI, AAAI or ECAI. The NeSy workshop typically took a day only at these major conferences, and it became then clear that given that the AI, cognitive science, machine learning, and applied logic communities share many common goals and aspirations it was necessary to provide an appropriately longer meeting, spanning over a week. The desire of many at NeSy to go deeper into the understanding of the main positions and issues, and to collaborate in a truly multidisciplinary way, using several applications (e. g. natural language processing, ontology reasoning, computer image and vision understanding, multimodal learning, knowledge representation and reasoning) towards achieving specific objectives, has prompted us to put together this Dagstuhl seminar marking the 10th anniversary of the workshop.

Further, neural-symbolic computation brings together an integrated methodological perspective, as it draws from both neuroscience and cognitive systems. In summary, neural-symbolic computation is a promising approach, both from a methodological and computational perspective to answer positively to the need for effective knowledge representation, reasoning and learning systems. The representational generality of neural-symbolic integration (the ability to represent, learn and reason about several symbolic systems) and its learning robustness provides interesting opportunities leading to adequate forms of knowledge representation, be they purely symbolic, or hybrid combinations involving probabilistic or numerical representations.

The seminar tackled diverse applications, in computer vision and image understanding, natural language processing, semantic web and big data. Novel approaches needed to tackle such problems, such as lifelong machine learning [6], connectionist applied logics [1, 2], deep learning [4], relational learning [7] and cognitive computation techniques have also been extensively analyzed during the seminar. The abstracts, discussions and open problems listed below briefly summarize a week of intense scientific debate, which illustrate the profitable atmosphere provided by the Dagstuhl scenery. Finally, a forthcoming article describing relevant challenges and open problems will be published at the Symposium on Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches at the AAAI Spring Symposium Series, to be held at Stanford in March 2015 [8]. This article also adds relevant content and a view of the area, illustrating its richness which may indeed lead to rich cognitive models integrating learning and reasoning effectively, as foreseen by Valiant [9].

Finally, we see neural-symbolic computation as a research area which reaches out to distinct communities: computer science, neuroscience, and cognitive science. By seeking to achieve the fusion of competing views it can benefit from interdisciplinary results. This contributes to novel ideas and collaboration, opening interesting research avenues which involve knowledge representation and reasoning, hybrid combinations of probabilistic and symbolic representations, and several topics in machine learning which can lead to both the construction of sound intelligent systems and to the understanding and modelling of cognitive and brain processes.

References

1 Artur S. d'Avila Garcez, Luis C. Lamb, and Dov M. Gabbay, Neural-Symbolic Cognitive Reasoning. Cognitive Technologies, Springer, 2009.

2 Barbara Hammer, Pascal Hitzler (Eds.): Perspectives of Neural-Symbolic Integration. Studies in Computational Intelligence 77, Springer 2007.

3 D. C. Ciresan, U. Meier, J. Schmidhuber. Multi-column Deep Neural Networks for Image Classification. IEEE Conf. on Computer Vision and Pattern Recognition CVPR 2012.

4 G. E. Hinton, S. Osindero, and Y. Teh, A fast learning algorithm for deep belief nets. Neural Computation, 18, 1527–1554, 2006..

5 Abdel-rahman Mohamed, George Dahl & Geoffrey Hinton. Acoustic Modeling Using Deep Belief Networks. IEEE Transactions on Audio, Speech, and Language Processing. 20(1):14–22, 2012.

6 D. Silver, Q. Yang, and L. Li, Lifelong machine learning systems: Beyond learning algorithms. Proceedings of the AAAI Spring Symposium on Lifelong Machine Learning, Stanford University, AAAI, March, 2013, pp. 49–55.

7 Stephen Muggleton, Luc De Raedt, David Poole, Ivan Bratko, Peter A. Flach, Katsumi Inoue, Ashwin Srinivasan: ILP turns 20 – Biography and future challenges. Machine Learning, 86(1):3–23, 2012.

8 Artur d'Avila Garcez, Tarek R. Besold, Luc de Raedt, Peter Foeldiak, Pascal Hitzler, Thomas Icard, Kai-Uwe Kuehnberger, Luis C. Lamb, Risto Miikkulainen, Daniel L. Silver. Neural-Symbolic Learning and Reasoning: Contributions and Challenges. Proceedings of the AAAI Spring Symposium on Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches, Stanford, March 2015.

9 L. G. Valiant, Knowledge Infusion: In Pursuit of Robustness in Artificial Intelligence. FSTTCS, pp. 415–422, 2008.

2 Table of Contents

3 Overview of Talks

3.1 Symbolic neural networks for cognitive capacities

Tsvi Achler (IBM Almaden Center, US)

Main reference T. Achler, "Symbolic neural networks for cognitive capacities," Biologically Inspired Cognitive Architectures, 9:71–81, 2014.
URL http://dx.doi.org/10.1016/j.bica.2014.07.001

Pattern recognition (identifying patterns from the environment using those stored in memory) and recall (describing or predicting the inputs associated with a stored pattern that can be recognized) are essential for neural-symbolic processing. Without them the brain cannot interact with the world e. g.: understand the environment, logic, and reason. Neural networks are efficient, biologically plausible algorithms that can perform large scale recognition. However, most neural network models of recognition perform recognition but not recall. It remains difficult to connect models of recognition with models of logic and emulate fundamental brain functions, because of the symbolic recall limitation. Before discussing symbolic networks further, one of the important realizations from the Dagstuhl seminar is that folks that focus on neural networks have a different definition of symbolic (and sub-symbolic) than folks that focus on logic. This matter was not fully solved. Subsequently I carefully define symbolic and note that in some literatures this term may be used differently. Here symbolic (call it "functional" symbolic?) is defined by the relation between input features and outputs (e. g. zebra has 4 legs). I assume that weights of neurons responsible for zebra demark mark this in connection weights that do not change. Let me clarify. There are two types of neural networks in the literature defined by how neurons learn for recognition processing: localist and globalist. In localist methods only neurons related to the information adjust their weights based learning on rules quantified within the neuron. Simple Hebbian learning is an example of this rule. Globalist methods in contrasts may require all neurons (including those that are not directly responsible) to change their weights to learn a new relation. PDP and feedforward models are examples of global learning. My symbolic definition is localist because I assumed the zebra neuron is independent of other neurons in that it does not change if another neuron is added with another symbolic relation (e. g. there exists another neuron representing another animal that has 0,4,6,8 or however many legs). Using this definition a neural network that is symbolic neural network cannot be globalist. A symbolic network also requires the ability to recall: to be able to derive from the symbol (e. g. zebra) what are the characteristic components (e. g. 4 legs, stripes etc). Thus the label (e. g. zebra) behaves as a symbol that encapsulates the components that are associated with it (legs, stripes, tail, hooves etc). Globalist networks cannot recall and subsequently in some literatures are called sub-symbolic (e. g. [2, 3]). Fortunately localist networks involve symmetrical top-down connections (from label to components) and the best example of such networks are auto-associative networks (e. g. Restricted Boltzmann Machines for Deep Learning). However auto-associative networks have self-excitatory symmetrical connections (positive feedback). A property of self-excitatory feedback is that iterative activation of even small values will lead to the maximal values regardless whether non-binary values are used. This degrades performance. I introduce a different localist model from auto-associative networks that uses are self- inhibitory symmetrical connections (negative feedback). The proposed model can converge to non-binary real-valued activations and is sensitive to real-valued weights. Moreover the network can be shown mathematically to obtain analogous solutions as standard feedforward (globalist) neural networks. Thus we have a model that can be as powerful as popular globalist neural networks, but is localist and symbolic. It can perform recall: retrieve the components involved in recognizing the label [1]. I hope to see more focus on these type of approaches within the neural symbolic community.

References

1 Achler T. Symbolic Neural Networks for Cognitive Capacities, Special issue of the Journal on Biologically Inspired Cognitive Architectures, 9, 71–81, 2014.
2 Fodor J.A., Pylyshyn Z.W. Connectionism and Cognitive Architecture: A Critical Analysis, Cognition 28, 3–71, 1988.
3 Sun R. Duality of the Mind: A Bottom-up Approach Toward Cognition. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

3.2 On extracting Rules for: enriching ontological knowledge bases, complementing heterogeneous sources of information, empowering the reasoning process

Claudia d'Amato (University of Bari, IT)

Joint work of d'Amato, Claudia; Staab, Steffen; Bryl, Volha; Serafini, Luciano

The Linked Open Data (LOD) cloud, which represents a significant example of Bid Data, could be seen as a huge portion of assertional knowledge whose intentional part is formally defined by existing OWL ontologies freely available on the Web. LOD constitutes a tremendous source of knowledge, that as such needs effective and efficient methods for its management. Data mining techniques could play a key role with this respect. The focus of the talk is on the discovery and extraction of knowledge patterns that are hidden in the (often noisy and inherently incomplete) data. Hidden knowledge patterns are extracted in the form of (relational) association rules by exploiting the evidence coming from the ontological knowledge bases [1] and/or from heterogeneous sources of information (i. e. an ontology and a relational databases referring to the same domain) [2] as well as by exploiting reasoning capabilities. While using methods at the state of the art, that as such necessarily need a further and deeper investigation for really scaling on very large data sets, the main focus will be on the potential that the extracted rules may have for: enriching existing ontological knowledge bases, for complementing heterogeneous sources of information, and for empowering the deductive reasoning process.

Particularly, the talk is organized in two parts. In the first one, the focus is on extracting hidden knowledge patterns from purely ontological knowledge bases. In the second one, the focus is on extracting hidden knowledge patterns from heterogeneous source of information.

The key observation motivating the first part of the talk is given by the fact that ontological knowledge bases are often not complete in the sense that missing concept and role assertions, with respect to the reference domain, can be found, as well as missing disjointness axioms and/or relationships. In order to cope with this problem, a method for discovering DL-Safe [4, 5] *Relational Association rules*, represented with SWRL [3] language, is presented [1]. This method is intended to discover all possible hidden knowledge patters that may be used for: a) (semi-)automatizing the completion of the assertional knowledge (given the pattern in the left hand side of a discovered rule, a new concept/role assertion may be induced by the right hand side of the rule); b) straightforwardly extending and enriching the expressive power of existing ontologies with formal rules, while ensuring and maintaining the decidability of the reasoning operators (because DL-Safe SWRL rules are extracted [3, 5]); c) suggesting knew knowledge axioms (induced by the discovered association rules). Inspired to [11, 12], the proposed method implements a level-wise generate-and-test approach that, starting with an initial general pattern, i. e. a concept name (jointly with a variable name) or a role name (jointly with variable names) proceeds, at each level, with generating a number of specializations by the use of suitable operators defined for the purpose. Each specialized pattern is then evaluated, on the ground of formally defined conditions, for possible pruning. This process is iterated until a predefined stopping criterion met. Besides of developing a scalable algorithm, the experimental evaluation of the developed method represents one of the most challenging problem since it requires the availability of gold standards (currently not available) with respect to which assessing the validity of the induced new knowledge. A possible solution is presented in [6].

As regards the second part of the talk, the motivating observation is given by the fact that even if available domain ontologies are increasing over the time, there is still a huge amount of data stored and managed with RDBMS and referring to the same domain. The key idea is that this complementarity could be exploited for discovering knowledge patterns that are not formalized within the ontology (or the RDBMS) but that are learnable from the data. For the purpose, a framework for extracting hidden knowledge patterns across ontologies and relational DBMS, called *Semantically Enriched Association Rules*, is illustrated [2, 13]. It is grounded on building an integrated view of (a part of) the RDBM and the ontology in a tabular representation which allows the exploitation of well know state of the art algorithms, such as the Apriori algorithm [14], for extracting Association Rules. The extracted patterns can be used for enriching the available knowledge (in both format) and for refining existing ontologies. Additionally, the extracted semantically enriched association rules can be exploited when performing deductive reasoning on an ontological knowledge bases. Specifically, a modified Tableaux algorithm, that we call *Data Driven Tableaux algorithm* is introduced [15, 13]. It is intended as a method for performing automated reasoning on grounded knowledge bases (i. e. knowledge bases linked to RDBMS data) which combines logical reasoning and statistical inference (coming from the discovered semantically enriched association rules) thus making sense of the heterogeneous data sources. The goals of the Data Driven Tableaux algorithm are twofold. On one hand it aims at reducing the computational effort for finding a model for a given (satisfiable) concept. On the other hand it aims at suppling the "most plausible model", that is the one that best fits the available data, for a given concept description. The key point of the algorithm is a defined heuristic, exploiting the semantically enriched association rules, to be used when random choices (e. g. when processing a concepts disjunction) occur. The proposed framework has to be intended as the backbone of a mixed models representation and reasoning.

The exploitation of association rules is not new in the Semantic Web context. In [6], a framework for discovering association rules for predicting new role assertions from an RDF data source is proposed, but no reasoning capabilities and TBox information are exploited for the purpose. Additionally, the extracted patterns are not integrated in the considered source of knowledge. Heterogeneous sources of information have been considered in [7, 8], where frequent patterns are discovered, respectively in the form of DATALOG clauses, from an $\mathcal{AL}$-Log knowledge base at different granularity level, and in the form of conjunctive queries, given a specified objective. Additional usages of association rules have been proposed in [9], where association rules are learnt from RDF data for inducing a schema ontology, but without exploiting any reasoning capabilities and in [10] where association rules are exploited for performing RDF data compression.

References

1 C. d'Amato and S. Staab. Predicting Assertions in Ontological Knowledge Bases by Discovering Relational Association Rules. Technical Report. http://www.di.uniba.it/~cdamato/TechAssoRules.pdf (2013).
2 C. d'Amato and Volha Bryl and Luciano Serafini. Semantic Knowledge Discovery from Heterogeneous Data Sources. In Proc. of the Knowledge Engineering and Knowledge Management – 18th International Conference (EKAW'12), vol. 7603, pp. 26–31, Springer, LNCS (2012).
3 I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean. Swrl: A semantic web rule language combining owl and ruleml http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/, 2004.
4 B. Motik, U. Sattler, and R. Studer. Query answering for owl-dl with rules. In *Proceedings of the International Semantic Web Conceference*, volume 3298 of *LNCS*, pages 549–563. Springer, 2004.
5 B. Motik, U. Sattler, and R. Studer. Query answering for owl-dl with rules. *Web Semantics*, 3(1):41–60, 2005.
6 L. Galárraga, C. Teflioudi, F. Suchanek, and Katja Hose. Amie: Association rule mining under incomplete evidence in ontological knowledge bases. In *Proceedings of the 20th International World Wide Web Conference (WWW 2013)*. ACM, 2013.
7 F. A. Lisi. Al-quin: An onto-relational learning system for semantic web mining. *International Journal of Semantic Web and Information Systems*, 2011.
8 J. Józefowska, A. Lawrynowicz, and T. Lukaszewski. The role of semantics in mining frequent patterns from knowledge bases in description logics with rules. *Theory and Practice of Logic Programming*, 10(3):251–289, 2010.
9 J. Völker and M. Niepert. Statistical schema induction. In G. Antoniou et al., editors, *The Semantic Web: Research and Applications – 8th Extended Semantic Web Conference, (ESWC 2011), Proc., Part I*, volume 6643 of *LNCS*, pages 124–138. Springer, 2011.
10 A. K. Joshi and P. Hitzler and G. Dong Logical Linked Data Compression In *The Semantic Web: Research and Applications – 10th Extended Semantic Web Conference, (ESWC 2013), Proceedings*, volume 7882 of *LNCS*, pages 170–184. Springer, 2013.
11 L. Dehaspeand and H. Toironen. Discovery of frequent datalog patterns. *Data Mining and Knowledge Discovery*, 3(1):7–36, 1999.
12 B. Goethals and J. Van den Bussche. Relational association rules: Getting warmer. In *Proceedings of the International Workshop on Pattern Detection and Discovery*, volume 2447 of *LNCS*, pages 125–139. Springer, 2002.
13 C. d'Amato and Volha Bryl and Luciano Serafini. Semantic Knowledge Discovery and Data-Driven Logical Reasoning from Heterogeneous Data Sources. In F. Bobillo et al.(Eds.), ISWC International Workshops, URSW 2011-2013, Revised Selected Papers, vol. 8816 Springer, LNCS/LNAI (2014).
14 R. Agrawal, T. Imielinski, and A. N. Swami. Mining association rules between sets of items in large databases. In P. Buneman and S. Jajodia, editors, *Proc. of the 1993 ACM SIGMOD Int. Conf. on Management of Data*, pages 207–216. ACM Press, 1993.
15 C. d'Amato and Volha Bryl and Luciano Serafini. Data-Driven Logical Reasoning. In Proc. of the 8th Int. Workshop on Uncertainty Reasoning for the Semantic Web (URSW'12), vol. 900, CEUR (2012).

3.3 Neural-Symbolic Computing, Deep Logic Networks and Applications

Artur d'Avila Garcez (City University London, GB)

Joint work of cf. references
Main reference A. S. d'Avila Garcez, L. C. Lamb, D. M. Gabbay, "Neural-Symbolic Cognitive Reasoning," Springer, ISBN 978-3-642-09229-9, 2009.
URL http://www.springer.com/computer/ai/book/978-3-540-73245-7

In this talk I reviewed the work carried out with many collaborators over the past 15 years in the area of neural-symbolic computing, starting with the CILP system for integrating logic programming and recurrent neural networks trained with backpropagation [1]. CILP networks take advantage of background knowledge during learning, which can improve training performance as shown in power systems and bioinformatics applications [2]. Knowledge extraction allows CILP networks to be described in symbolic form for the sake of transfer learning and explanation [3]. Extensions of CILP, including the use of feedback, network ensembles and nested networks, allows the representation and learning of various forms of nonclassical reasoning, including modal, temporal and epistemic reasoning [4, 5], as well as abduction [6]. This has led to a full solution in connectionist form of the so-called muddy children puzzle in logic [7]. Fibring of CILP networks offers further expressive power by combining networks of networks for simultaneous learning and reasoning [8]. Applications have included training and assessment in simulators, normative reasoning and rule learning, integration of run-time verification and adaptation, action learning and description in videos [9, 10, 13]. Current developments and efforts have been focused on: fast relational learning using neural networks (the CILP++ system) [11] and effective knowledge extraction from large networks, including deep networks and the use of knowledge extraction for transfer learning [12]. Future applications include the analysis of complex networks, social robotics and health informatics, and multimodal learning and reasoning combining video and audio data with metadata.

References

1 A. S. d'Avila Garcez, K. Broda and D. M. Gabbay, *Neural-Symbolic Learning Systems: Foundations and Applications*, Springer, 2002.
2 A. S. d'Avila Garcez and G. Zaverucha. The Connectionist Inductive Learning and Logic Programming System. Applied Intelligence 11(1):59–77, 1999.
3 A. S. d'Avila Garcez, K. Broda and D. M. Gabbay. Symbolic Knowledge Extraction from Trained Neural Networks: A Sound Approach. Artificial Intelligence, 125(1–2):153–205, January 2001.
4 A. S. d'Avila Garcez and L. C. Lamb. A Connectionist Computational Model for Epistemic and Temporal Reasoning. Neural Computation, 18(7):1711–1738, July 2006.
5 A. S. d'Avila Garcez, L. C. Lamb and D. M. Gabbay. Connectionist Modal Logic: Representing Modalities in Neural Networks. Theoretical Computer Science, 371(1–2):34–53, February 2007.
6 A. S. d'Avila Garcez, D.M. Gabbay, O. Ray and J. Woods. Abductive Reasoning in Neural-Symbolic Learning Systems. Topoi: An International Review of Philosophy, 26:37–49, March 2007.
7 A. S. d'Avila Garcez, L. C. Lamb and D. M. Gabbay. *Neural-Symbolic Cognitive Reasoning*, Springer, 2009.

8 A. S. d'Avila Garcez and D. M. Gabbay. Fibring Neural Networks. In Proc. 19th National Conference on Artificial Intelligence AAAI 2004. San Jose, California, USA, AAAI Press, July 2004.

9 R. V. Borges, A. S. d'Avila Garcez and L. C. Lamb. Learning and Representing Temporal Knowledge in Recurrent Networks. IEEE Transactions on Neural Networks, 22(12):2409–2421, December 2011.

10 L. de Penning, A. S. d'Avila Garcez, L. C. Lamb and J. J. Meyer. A Neural-Symbolic Cognitive Agent for Online Learning and Reasoning. In Proc. IJCAI'11, Barcelona, Spain, July 2011.

11 M. Franca, G. Zaverucha and A. S. d'Avila Garcez. Fast Relational Learning using Bottom Clause Propositionalization with Artificial Neural Networks, Machine Learning, 94(1):81–104, 2014.

12 S. Tran and A. S. d'Avila Garcez. Logic Extraction from Deep Belief Networks. In Proc. ICML Workshop on Representation Learning, Edinburgh, Scotland, July 2012.

13 G. Boella, S. Colombo-Tosatto, A. S. d'Avila Garcez, V. Genovese, A. Perotti and L. van der Torre. Learning and Reasoning about Norms using Neural-Symbolic Systems. In Proc. 11th International Conference on Autonomous Agents and Multiagent Systems, AAMAS'12, Valencia, Spain, June 2012.

3.4 Dreaming and Consciousness in Deep Neural-Symbolic Cognitive Agents

Leo de Penning (TNO Behaviour and Societal Sciences – Soesterberg, NL)

Joint work of de Penning, Leo; Garcez, Artur d'Avilla; Meyer, John-Jules
Main reference L. de Penning, A. d'Avilla Garcez, J.-J. C. Meyer, "Dreaming Machines: On multimodal fusion and information retrieval using Neural-Symbolic Cognitive Agents," in Proc. of the 2013 Imperial College Computing Student Workshop (ICCSW'13), OASIcs, Vol. 35, pp. 89–94, 2013.
URL http://dx.doi.org/10.4230/OASIcs.ICCSW.2013.89

Deep Boltzmann Machines (DBM) have been used as a computational cognitive model in various AI-related research and applications, notably in computational vision and multimodal fusion. Being regarded as a biological plausible model of the human brain, the DBM is also becoming a popular instrument to investigate various cortical processes in neuroscience. In this paper, we describe how a multimodal DBM is implemented as part of a Neural-Symbolic Cognitive Agent (NSCA) for real-time multimodal fusion and inference of streaming audio and video data. We describe how this agent can be used to simulate certain neurological mechanisms related to hallucinations and dreaming and how these mechanisms are beneficial to the integrity of the DBM. Also we will explain how the NSCA is used to extract multimodal information from the DBM and provide a compact and practical iconographic temporal logic formula for complex relations between visual and auditory patterns. Finally we will discuss the implications of the work in relation to Machine Consciousness.

3.5 Progress in Probabilistic Logic Programming

Luc De Raedt (KU Leuven, BE)

Probabilistic logic programs combine the power of a programming language with a possible world semantics, typically based on Sato's distribution semantics and they have been studied for over twenty years. In this talk, I introduced the concepts underlying probabilistic programming, their semantics, different inference and learning mechanisms. I then reported on recent progress within this paradigm. This was concerned with an extension towards dealing with continuous distributions as well as coping with dynamics. This is the framework of distributional clauses that has been applied to several applications in robotics, for tracking relational worlds in which objects or their properties are occluded in real time. Finally, some remaining open challenges were discussed.

See also the websites http://dtai.cs.kuleuven.be/problog/ and http://dtai.cs.kuleuven.be/ml/systems/DC/ for more details and an interactive tutorial on ProbLog.

References

1 Nitti, D., De Laet, T., & De Raedt, L. (2013). A particle filter for hybrid relational domains. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2013.
2 De Raedt, L., & Kimmig, A. Probabilistic Programming Concepts, arXiv:1312.4328, 2013.

3.6 Semantic and Fuzzy Modelling and Recognition of Human Activities in Smart Spaces. A case study on Ambient Assisted Living

Natalia Díaz-Rodríguez (Turku Centre for Computer Science, FI)

Joint work of Díaz-Rodríguez, Natalia; León Cadahía, Olmo; Cuéllar, M. P.; Lilius, Johan; Delgado Calvo-Flores, Miguel
Main reference N. Díaz-Rodríguez, O. León Cadahía, M. P. Cuéllar, J. Lilius, M. Delgado Calvo-Flores, "Handling real-world context-awareness, uncertainty and vagueness in real-time human activity tracking and recognition with a fuzzy ontology-based hybrid method," MDPI Physical Sensors, 14(10):18131–18171, 2014.
URL http://dx.doi.org/10.3390/s141018131

Human activity recognition in everyday environments is a critical task in Ambient Intelligence applications to achieve proper Ambient Assisted Living. Key challenges still remain to be tackled to achieve robust methods. Our hybrid system allows to model and recognize a set of complex scenarios where vagueness and uncertainty is inherent to the human nature of the users that perform it. We provide context meaning to perform sub- activity tracking and recognition from depth video data. To achieve a more loosely coupled model that lets flexibility to be part of the recognition process, we validate the advantages of a hybrid data-driven and knowledge-driven system with a challenging public dataset and achieve an accuracy of 90.1% and 91.1% respectively for low and high-level activities. The handling of uncertain, incomplete and vague data (i. e., missing sensor readings or execution variations) is tackled for first time with a public depth-video dataset taking into account the semantics of activities, sub-activities and real-time object interaction. This entails an improvement over both entirely data-driven approaches and merely ontology- based approaches.

3.7 Making the latent category structure of fMRI data explicit with Formal Concept Analysis

Dominik Endres (Universität Marburg, DE)

Joint work of Endres, Dominik; Adam, Ruth; Giese, Martin A.; Noppeney, Uta
Main reference D. Endres, R. Adam, M. A. Giese, U. Noppeney, "Understanding the Semantic Structure of Human fMRI Brain Recordings With Formal Concept Analysis," in Proc. of the 10th Int'l Conf. on Formal Concept Analysis (ICFCA'12), LNCS, Vol. 7278, pp. 96-111, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-29892-9_13

Understanding how semantic information is represented in the brain has been an important research focus of Neuroscience in the past few years. The work I presented in this talk is aimed at extracting concepts and their relationships from brain activity, and to correlate these concept with behavioral measures. We showed previously (Endres et al 2010) that Formal Concept Analysis (FCA) can reveal interpretable semantic information (e. g. specialization hierarchies, or feature-based representations) from electrophysiological data. Unlike other analysis methods (e. g. hierarchical clustering), FCA does not impose inappropriate structure on the data. FCA is a mathematical formulation of the explicit coding hypothesis (Foldiak, 2009) Furthermore we (Endres et al 2012) investigated whether similar findings can be obtained from fMRI BOLD responses recorded from human subjects. While the BOLD response provides only an indirect measure of neural activity on a much coarser spatio-temporal scale than electrophysiological recordings, it has the advantage that it can be recorded from humans, which can be questioned about their perceptions during the experiment. Furthermore, the BOLD signal can be recorded from the whole brain simultaneously. In our experiment, a single human subject was scanned while viewing 72 grayscale pictures of animate and inanimate objects in a target detection task. These pictures comprise the formal objects for FCA. We computed formal attributes by learning a hierarchical Bayesian classifier, which maps BOLD responses onto binary features, and these features onto object labels. The connectivity matrix between the binary features and the object labels can then serve as the formal context. In a high-level visual cortical area (IT), we found a clear dissociation between animate and inanimate objects with the inanimate category subdivided between animals and plants when we increased the number of attributes extracted from the fMRI signal. The inanimate objects were hierarchically organized into furniture and other common items, including vehicles. We also used FCA to display organizational differences between high-level and low-level visual processing areas. For a quantitative validation of that observation, we show that the attribute structure computed from the IT fMRI signal is highly predictive of subjective similarity ratings, but we found no such relationship to responses from early visual cortex. Collaborators: Peter Foldiak, Uta Priss, Ruth Adam, Uta Noppeney

References

1. D. Endres, P. Földiak and U. Priss (2010) An Application of Formal Concept Analysis to Semantic Neural Decoding, Annals of Mathematics and Artificial Intelligence, 57(3–4), pp. 233–248.
2. D. Endres, R. Adam, M. A. Giese and U. Noppeney (2012) Understanding the Semantic Structure of Human fMRI Brain Recordings With Formal Concept Analysis., In ICFCA 2012, 10th Inte'l Con. on Formal Concept Analysis, LNCS 7278, Springer, 96–111.
3. P. Földiak (2009) Neural Coding: Non-Local but Explicit and Conceptual, Current Biology, 19(19):R904–R906, 2009.
4. B. Ganter and R. Wille (1999). Formal concept analysis: Mathematical foundations, Springer.

3.8 Symbolic Data Mining Methods Applied to Human Sleep Records

Jacqueline Fairley (Emory University – Atlanta, US)

Joint work of Fairley, Jacqueline; Georgoulas, George; Karvelis, Petros; Stylios, Chrysostomos; Rye, David; Bliwise, Donald
Main reference J. A. Fairley, G. Georgoulas, P. Karvelis, C. D. Stylios, D. B. Rye, D. L. Bliwise, "Symbolic Representation of Human Electromyograms for Automated Detection of Phasic Activity during Sleep," 2014.
URL https://www.academia.edu/7659893/SYMBOLIC_REPRESENTATION_OF_HUMAN_ELECTROMYOGRAMS_FOR_AUTOMATED_DETECTION_OF_PHASIC_ACTIVITY_DURING_SLEEP

Background: Phasic electromyographic (EMG)/muscle activity in human overnight polysomnograms (PSGs) represent a potential indicator/quantitative metric for identifying various neurodegenerative disorder populations and age-matched controls [1].

Unfortunately, visual labeling of phasic EMG activity is time consuming making this method unscalable for clinical implementation. Therefore, we propose computerized labeling of EMG activity in a detection scheme utilizing k-Nearest Neighbor classification and Symbolic Aggregate approXimation (SAX), a novel algorithm from the field of time series data mining that transforms a time series, such as EMG, into a string of arbitrary symbols [2]. A primary advantage of SAX analysis includes access to robust symbolic based data mining algorithms viable for scalable computing.

Methods: Six male subjects (S001:S006) polysomnograms (PSGs)/sleep data sets were visually scored, using one second epochs, for phasic and non-phasic left and right leg EMG activity (sampling rate 200Hz), by the same trained visual scorer. Phasic muscle activity epochs were characterized by amplitudes visually exceeding four times the surrounding background activity and having time durations between 100 to 500 msec. SAX was applied to all EMG records using a one second non-overlapping moving window, four symbol alphabet, and $\frac{1}{2}$ sec frames, followed by translation of the SAX string into an intelligent icon, a color mapped image representing the frequency of each word in the SAX string. Results: SAX based classification scheme results, using 10-fold cross validation and k- Nearest Neighbor Classification (best of k:1:1:7; minimum value:increment value:maximum value), were compared to visual labeling [3]. Detection of non-phasic EMG activity exceeded 90% for all six subjects: S001 (98.4), S002 (97.8), S003 (98.1), S004 (93.6), S005 (95.2), and S006 (95.8). Phasic EMG activity detection surpassed 80% for three subjects: S001 (90.5), S004 (81.8), and S006 (87.1). However, phasic EMG activity detection decreased in performance for S002 (61.0), S003 (53.6) and S005 (68.0).

Conclusions: Detection rates for half of the subjects indicate feasibility of replacing tedious expert visual scoring with the proposed computational scheme. However, this scheme lacks robustness across all subjects, and requires refinement of SAX alphabet size and frame length along with comparison with other classification algorithms such as Support Vector Machines and Random Forest. Most importantly, efficient fine-tuning of this computational scheme promises to hasten computerized EMG activity scoring for neurodegenerative disorder tracking in clinical settings.

References

1 Bliwise, D. L., He, L., Ansari, F. P., and Rye, D. B. (2006) Quantification of electromyographic activity during sleep: a phasic electromyographic metric. Journal of Clinical Neurophysiology, Vol. 23, pp. 59–67.

2 Lin, J., Keogh, E., Lonardi, S., and Chiu, B. (2003) A Symbolic Representation of Time Series, with Implications for Streaming Algorithms, In Proc. of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. San Diego, CA.
3 Fairley, J. A., Georgoulas, G., Karvelis, P., Stylios, C. D., Rye, D. B., Bliwise, D. L. (2014). Symbolic Representation of Human Electromyograms for Automated Detection of Phasic Activity during Sleep, Statistical Signal Processing Workshop (SSP) Proceedings, 2014 IEEE, pp. 185–188, Gold Coast, Australia.

3.9 Affordances, Actionability, and Simulation

Jerry A. Feldman (ICSI – Berkeley, US)

The notion of affordances depends crucially on the actions available to an agent in context. When we add the expected utility of these actions in context, the result has been called actionability. There is increasing evidence that AI and Cognitive Science would benefit from shifting from a focus on abstract "truth" to treating actionability as the core issue for agents. Actionability also somewhat changes the traditional concerns of affordances to suggest a greater emphasis on active perception. An agent should also simulate (compute) the likely consequences of actions by itself or other agents. In a social situation, communication and language are important affordances.

3.10 Simulation Semantics and the Rebirth of NLU

Jerry A. Feldman (ICSI – Berkeley, US)

Natural Language Understanding (NLU) was one of the main original goals of artificial intelligence and cognitive science. This has proven to be extremely challenging and was nearly abandoned for decades. We describe an implemented system that supports full NLU for tasks of moderate complexity. The natural language interface is based on Embodied Construction Grammar and simulation semantics. The system described here supports dialog with an agent controlling a simulated robot, but is flexible with respect to both input language and output task.

3.11 The Neural Binding Problem(s)

Jerry A. Feldman (ICSI – Berkeley, US)

As with many other "problems" in vision and cognitive science, "the binding problem" has been used to label a wide range of tasks of radically different behavioral and computational structure. These include a "hard" version that is currently intractable, a feature-binding variant that is productive routine science and a variable-binding case that is unsolved, but should be solvable. The talk will cover all these and some related problems that seem intractably hard as well as some that are unsolved, but are being approached with current and planned experiments.

References

1 Jerome A. Feldman. *The Neural Binding Problem(s)*. Cognitive Neurodynamics, Vol. 7, Issue 1, pp. 1–11, February 2013

3.12 Fast Relational Learning using Neural Nets

Manoel Franca (City University London, GB)

Joint work of Franca, Manoel; Garcez, Artur; Zaverucha, Gerson
URL http://www.dagstuhl.de/mat/Files/14/14381/14381.FrancaManoel.ExtAbstract.pdf

Relational learning can be described as the task of learning first-order logic rules from examples. It has enabled a number of new machine learning applications, e. g. graph mining and link analysis. We introduce a fast method and system for relational learning, called CILP++, which handles first-order logic knowledge and have been on several ILP datasets, comparing results with Aleph. The results show that CILP++ can achieve accuracy comparable to Aleph, while being generally faster. Several alternative approaches, both for BCP propositionalization and for CILP++ learning, are also investigated.

3.13 Evolutionary and Swarm Computing for the Semantic Web

Christophe D. M. Gueret (DANS – Den Hague, NL)

Joint work of Gueret, Christophe D. M.; Schlobach, S. ; Dentler, K. ; Schut, M. ; Eiben, G.
Main reference C. Gueret, S. Schlobach, K. Dentler, M. Schut, G. Eiben, "Evolutionary and Swarm Computing for the Semantic Web," IEEE Computational Intelligence Magazine, 7(2):16–31, 2012.
URL http://dx.doi.org/10.1109/MCI.2012.2188583

The Semantic Web has become a dynamic and enormous network of typed links between data sets stored on different machines. These data sets are machine readable and unambiguously interpretable, thanks to their underlying standard representation languages. The expressiveness and flexibility of the publication model of Linked Data has led to its widespread adoption and an ever increasing publication of semantically rich data on the Web. This success however has started to create serious problems as the scale and complexity of information outgrows the current methods in use, which are mostly based on database technology, expressive knowledge representation formalism and high-performance computing. We argue that methods from computational intelligence can play an important role in solving these problems. In this paper we introduce and systemically discuss the typical application problems on the Semantic Web and argue that the existing approaches to address their underlying reasoning tasks consistently fail because of the increasing size, dynamicity and complexity of the data. For each of these primitive reasoning tasks we will discuss possible problem solving methods grounded in Evolutionary and Swarm computing, with short descriptions of existing approaches. Finally, we will discuss two case studies in which we successfully applied soft computing methods to two of the main reasoning tasks; an evolutionary approach to querying, and a swarm algorithm for entailment.

3.14 Computer Science for Development

Christophe D. M. Gueret (DANS – Den Hague, NL)

Joint work of Gueret, Christophe D. M.; Schlobach, S. ; de Boer V.
Main reference World Wide Semantic Web community, "Let's build a world wide Semantic Web!," 2011.
URL http://worldwidesemanticweb.org

Data sharing usually focuses on centralized and very powerful solutions centred around Web hosted servers and (mobile) clients accessing it. As a direct consequence, the usage of Linked Data technology depends on the availability of a Web infrastructure compassing data-centres, high speed reliable Internet connection and modern client devices. If any of this is missing, our community is not able, yet, to provide any Linked Data enabled data management solution. Still, the digital divide that is currently widely recognized separates the world into those who have access to Web-based platforms and those who don't. When designing Linked Data platforms we tend to forget those 4 Billion persons who don't have access to Internet but would benefit from being able to share structured data. We should keep everyone in mind when we design Linked Data platforms and aim at helping to reduce this digital divide. We believe that achieving this goal implies working on three aspects (Infrastructure, Interfaces and Relevancy) around open data.

This problem the Semantic Web community faces doing knowledge representation in developing countries is only one facet of Computer Science. Many other aspects of it are also concerned. For instance, Human-Computer Interaction (HCI) need to account for users that don't read or write or don't speak any "common" language, Engineering need to be performed on smaller scale devices with sparse networkings and Information retrieval need to be done with a focus on locally relevant information. These many aspects of Computer Sciences affected by the specific challenges posed by using ICT in the developing world call for a global study over CS4D where researchers would join in ensuring the technology they work on is inclusive and usable by everyone world wide.

3.15 Combining Learning and Reasoning for Big Data

Pascal Hitzler (Wright State University – Dayton, US)

Reasoning and learning are natural allies. The former provides deductive expert system-like capabilities for dealing with interpretation of data, while the latter focuses on finding patterns in data. This perspective suggests a rather obvious workflow in which inductive and statistical methods analyze data, resulting in metadata which describes higher-level conceptualizations (metadata) of the data, which in turn enables the use of the data and metadata in deduction-based systems. However, this apparently obvious pipeline is broken since the current state of the art leaves gaps which need to be bridged by new innovations.

In this presentation, we discuss some of the recent work which addresses these gaps, with the goal of stimulating further research on the interplay between learning and reasoning.

References

1 P. Hitzler and F. van Harmelen. A reasonable semantic web. *Semantic Web*, 1(1–2):39–44, 2010.

2 P. Jain, P. Hitzler, P. Z. Yeh, K. Verma, and A. P. Sheth. Linked Data is Merely More Data. In D. Brickley, V. K. Chaudhri, H. Halpin, and D. McGuinness, editors, *Linked Data Meets Artificial Intelligence*, pages 82–86. AAAI Press, Menlo Park, CA, 2010.

3 P. Hitzler, M. Krötzsch, and S. Rudolph. *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC, 2010.

4 P. Hitzler and K. Janowicz. Linked Data, Big Data, and the 4th Paradigm. *Semantic Web*, 4(3):233–235, 2013.

5 K. Janowicz and P. Hitzler. The Digital Earth as knowledge engine. *Semantic Web*, 3(3):213–221, 2012.

6 P. Jain, P. Hitzler, A. P. Sheth, K. Verma, and P. Z. Yeh. Ontology alignment for linked open data. In P. F. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, L. Zhang, J. Z. Pan, I. Horrocks, and B. Glimm, editors, *The Semantic Web – ISWC 2010 – 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7–11, 2010, Revised Selected Papers, Part I*, volume 6496 of *Lecture Notes in Computer Science*, pages 402–417. Springer, 2010.

7 A. Joshi, P. Jain, P. Hitzler, P. Yeh, K. Verma, A. Sheth, and M. Damova. Alignment-based querying of Linked Open Data. In R. Meersman, H. Panetto, T. Dillon, S. Rinderle-Ma, P. Dadam, X. Zhou, S. Pearson, A. Ferscha, S. Bergamaschi, and I. F. Cruz, editors, *On the Move to Meaningful Internet Systems: OTM 2012, Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012, Rome, Italy, September 10–14, 2012, Proceedings, Part II*, volume 7566 of *Lecture Notes in Computer Science*, pages 807–824. Springer, Heidelberg, 2012.

3.16 From Human Reasoning Episodes to Connectionist Models

Steffen Hölldobler (TU Dresden, DE)

URL http://www.wv.inf.tu-dresden.de/~sh/

I present a new approach to model human reasoning based on reasoning towards an appropriate logical form, weak completion semantics, three-valued Lukasiewicz logic, and an appropriate semantic operator. The approach admits least models and, hence, reasoning is performed with respect to least models. After adding abduction the approach can adequately handle human reasoning episodes like the suppression and the selection task. Moreover, it can be mapped into a connectionist model using the core method.

3.17 On Concept Learning as Constructive Reasoning

Francesca Alessandra Lisi (University of Bari, IT)

Main reference F. A. Lisi, "A Declarative Modeling Language for Concept Learning in Description Logics," in Proc. of the 22nd Int'l Conf. on Inductive Logic Programming (ILP'12), LNCS, Vol. 7842, pp. 151–165, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-38812-5_11

In this talk I provided a novel perspective on Concept Learning, which relies on recent results in the fields of Machine Learning (ML)/Data Mining (DM) and Knowledge Representation (KR), notably De Raedt *et al.*'s work on declarative modeling of ML/DM problems [2] and Colucci *et al.*'s work on non-standard reasoning in the KR framework of Description Logics (DLs) [1]. In particular, I provided a formal characterization of Concept Learning which arises from the observation that the inductive inference deals with finding – or constructing – a concept. More precisely, non-standard reasoning services which support the inductive inference can be modeled as constructive reasoning tasks where the solution construction may be subject to optimality criteria. Under this assumption, I defined a declarative language – based on second-order DLs – for modeling different variants of the Concept Learning problem (namely, Concept Induction, Concept Refinement and Concept Formation) [3]. The language abstracts from the specific algorithms used to solve the Concept Learning problem in hand. However, as future work, I emphasized the need for an efficient and/or effective solver to make the proposed language more attractive from a practical point of view.

References

1 Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F. M., Ragone, A.: A unified framework for non-standard reasoning services in description logics. In: Coelho, H., Studer, R., Wooldridge, M. (eds.) ECAI 2010 – 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16–20, 2010, Proceedings. Frontiers in Artificial Intelligence and Applications, vol. 215, pp. 479–484. IOS Press (2010)
2 De Raedt, L., Guns, T., Nijssen, S.: Constraint programming for data mining and machine learning. In: Fox, M., Poole, D. (eds.) Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11–15, 2010. AAAI Press (2010)
3 Lisi, F. A.: A declarative modeling language for concept learning in description logics. In: Riguzzi, F., Zelezny, F. (eds.) Inductive Logic Programming, 22nd International Conference, ILP 2012, Dubrovnik, Croatia, September 17–19, 2012, Revised Selected Papers. Lecture Notes in Computer Science, vol. 7842. Springer Berlin Heidelberg (2013)

3.18 Interactive Intelligent Systems: Scaling Learning with the Expert in the Loop

Dragos Margineantu (Boeing Research & Technology, US)

URL http://www.dmargineantu.net/

Research in intelligent neural and symbolic systems has made significant advances with respect to the accuracy of predictions, detections, classifications. However in order to deploy these algorithms and tools, to execute or assist the execution of real world tasks, in most of the cases, these methods require the assistance of an AI expert. A suite of practical tasks can be addressed optimally at this point in time by a team that combines the expertise of the user with the strength of automated intelligent systems. Can we develop (or adapt our) existing algorithms for such tasks? We believe so! By formulating our research questions to capture the expert-intelligent system goals. This presentation will show how we formulated the research questions and adapted techniques such as inverse reinforcement learning (IRL) or active learning for assisting experts in tasks such as detecting abnormal agent behavior, scene analysis, and estimating intent. We will also outline some open research questions for usable expert-interactive learning.

3.19 Concepts, Goals and Communication

Vivien Mast (Universität Bremen, DE)

Joint work of Mast, Vivien;Couto Vale, Daniel; Falomir, Zoe; Elahi, Fazleh Mohammad
Main reference V. Mast, V., D. Couto Vale, Z. Falomir, F. M. Elahi, "Referential Grounding for Situated Human-Robot Communication," in Proc. of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SemDial'14), pp. 223–225, 2014.
URL http://www.macs.hw.ac.uk/InteractionLab/Semdial/semdial14.pdf

Much work in computational linguistics and cognitive science implicitly rests on the idea, dating back to Plato and Aristotle, that there are rational categories which are sets of entities in the real world, defined by necessary and sufficient properties, and that the power of linguistic expressions and mental concepts stems from their correspondence to such rational categories. I will discuss some limitations of a rational notion of concepts and meaning in the domain of reference, and argue that human concepts should be viewed from the perspective of actionability, as suggested by Jerry Feldman at this seminar. In particular, I will argue that concept assignment depends on context and the goals of the conceptualizing agent.

In the standard paradigm of REG (Krahmer & van Deemter, 2012), objects are represented by attribute-value pairs. The task of REG is defined as finding, for a given target object, a distinguishing description – a set of attribute-value pairs whose conjunction is true of the target but not of any of the other objects in the domain. However, research on collaborative reference has shown that reference ultimately does not rely on truth, but on common ground and efficient grounding mechanisms (Clark & Bangerter, 2004). I will argue that meta-knowledge about the potential of conceptual mismatch and miscommunication guide concept assignment in reference, and I will present the Probabilistic Reference And GRounding mechanism PRAGR for generating and interpreting referring expressions (Mast et al., 2014; Mast & Wolter, 2013). PRAGR is geared towards maximizing mutual understanding by flexibly assigning linguistic concepts to objects, depending on context.

References

1 Clark, H. H. & Bangerter, A. *Changing Ideas about Reference*. In: Sperber, Dan and Noveck, Ira A. (Ed.), Experimental Pragmatics. Hampshire, NY: Palgrave Macmillan, 2004.
2 Krahmer, E. & van Deemter, K. *Computational Generation of Referring Expressions: A Survey*. Computational Linguistics, 38(1), pp. 173–218, 2012.
3 Mast, V., Couto Vale, D., Falomir, Z. & Elahi, F. M. *Referential Grounding for Situated Human-Robot Communication.*. Proceedings of DialWatt/SemDial. 2014.

4 Mast, V. & Wolter, D. *A Probabilistic Framework for Object Descriptions in Indoor Route Instructions.*. In Tenbrink, T. and Stell, J. and Galton, A. and Wood, Z. (Ed.), Spatial Information Theory (Vol. 8116, pp. 185–204). Springer International Publishing, 2013

3.20 Grounding Meaning in Perceptual Representations

Risto Miikkulainen (University of Texas – Austin, US)

Joint work of Miikkulainen, Risto; Aguilar, Mario; Aguirre-Celis, Nora; Binder, Jeffrey R.; Connolly, Patrick; Fernandino, Leo; Morales, Isaiah; Williams, Paul;

How word meaning may be grounded in perceptual experience is a fundamental problem in neural-symbolic learning. I will describe an artificial neural network model that shows how this process may take place through learned associations between visual scenes and linguistic phrases. I will then describe ongoing work on identifying such associations from fMRI images of sentence comprehension.

3.21 Mining Graphs from Event Logs

Andrey Mokhov (Newcastle University, GB)

Joint work of Mokhov, Andrey; Carmona, Josep
Main reference A. Mokhov, J. Carmona, "Process Mining using Parameterised Graphs," Technical memo, Newcastle University, August 2014.
URL http://async.org.uk/tech-memos/NCL-EEE-MICRO-MEMO-2014-009.pdf

We introduce a mathematical model for compact representation of large families of (related) graphs [1], detecting patterns in graphs, and using such compact representations for process mining [2]. By process mining we mean understanding or explanation of behaviour of complex systems by observing events occurring in them. These events come in the form of event logs that record event types, time stamps and other associated metadata. The task of process mining is to extract useful knowledge from such logs, for example, to explain, predict or diagnose complex systems. We present graph-theoretic methods that extract information about concurrency and causality from such logs, and then attempt to represent the result in the most compact/simple form hopefully amenable to human understanding [3].

References

1 Andrey Mokhov, Alex Yakovlev. *Conditional partial order graphs: Model, synthesis, and application*, IEEE Transactions on Computers 59 (11), 1480–1493, 2010.
2 Wil van der Aalst. *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Springer, 2011.
3 Andrey Mokhov, Victor Khomenko. *Algebra of Parameterised Graphs*, ACM Transactions on Embedded Computing, 2014.

3.22 Learning Compositional Robot Activities from Examples

Bernd Neumann (Universität Hamburg, DE)

Joint work of Neumann, Bernd; Hotz, Lothar; Rost, Pascal; Lehmann, Jos
Main reference B. Neumann, L. Hotz, P. Rost, J. Lehmann, "A Robot Waiter Learning from Experiences," in Proc. of the 10th Int'l Conf. on Machine Learning and Data Mining in Pattern Recognition (MLDM'14), LNCS, Vol. 8556, pp. 286–299, Springer, 2014.
URL http://dx.doi.org/10.1007/978-3-319-08979-9_22

In the KR framework of the EU project RACE, as in many other systems, robot activities are described by compositional hierarchies connecting activity concepts at higher abstraction levels with components at lower levels, down to action primitives of the robot platform with quantitative parameters, and down to percepts at neural level. One way for a service robot to increase its competence is to learn new activities based on known subactivities and coarse instructions. Given an initial repertoire of basic operations, such a process can establish compositional structures at increasingly high levels of abstraction and complexity. In this talk I describe recent advances in learning compositional structures using a Description Logic (DL) extended by semantic attachments as formal knowledge representation framework. A learning curriculum, based on positive examples, is presented where the robot has to determine autonomously which spatiotemporal conditions must be satisfied for a newly learnt activity. It is shown that the robot can construct conceptual descriptions from the examples in such a way that the intended target description is approached with monotonously increasing generality. The generalization process is realized by aligning concept graphs obtained from DL representations and merging corresponding nodes by a Good Common Subsumer (GCS). It is shown that this process can also be used for adapting an existing concept to a new situation. Examples are presented for a service robot learning waiter activities in a restaurant domain.

3.23 Neural-Symbolic Runtime Verification

Alan Perotti (University of Turin, IT)

Joint work of Perotti, Alan; d'Avila Garcez, Artur; Boella, Guido
Main reference A. Perotti, A. S. d'Avila Garcez, G. Boella, "Neural Networks for Runtime Verification," in Proc. of the 2014 Int'l Joint Conf. on Neural Networks (IJCNN'14), pp. 2637–2644, IEEE, 2014.
URL http://dx.doi.org/10.1109/IJCNN.2014.6889961

I introduced RuleRunner, a novel Runtime Verification system for monitoring LTL properties over finite traces. By exploiting results from the Neural-Symbolic Integration area, a RuleRunner monitor can be encoded in a recurrent neural network. The results show that neural networks can perform real-time runtime verification and techniques of parallel computing can be applied to improve the performance in terms of scalability. Furthermore, our framework allows for property adaptation by using a standard neural network learning algorithm.

3.24 Symbolic Computation, Binding and Constraint Learning in Bolzmann Machines

Gadi Pinkas (Center for Academic Studies, Or-Yehudah and Bar-Ilan University, IL)

Joint work of Pinkas, G, Cohen S., Lima P.
Main reference G. Pinkas, P. Lima, S. Cohen, "Representing, binding, retrieving and unifying relational knowledge using pools of neural binders," Journal of Biologically Inspired Cognitive Architectures, 6(October 2013):87–95, 2013.
URL http://dx.doi.org/10.1016/j.bica.2013.07.005

For a long time, connectionist architectures have been criticized for having propositional fixation, lack of compositionality and, in general, for their weakness in representing sophisticated symbolic information, learning it and processing it. This work offers an approach that allows full integration of symbolic AI with the connectionist paradigm. We show how to encode, learn and process relational knowledge using attractor based artificial neural networks, such as Boltzmann Machines. The neural architecture uses a working memory (WM), consisting of pools of "binders", and a long-term synaptic-memory (LTM) that can store a large relational knowledge-base (KB). A compact variable binding mechanism is proposed which dynamically allocates ensembles of neurons when a query is clamped; retrieving KB items till a solution emerges in the WM. A general form of the Hebbian learning rule is shown that learns from constraint violations. The learning rule is applied to High-Order Boltzmann machines (with sigma-pi connections) and is shown to learn networks with attractors (energy minima) representing correct symbolic inferences. We illustrate the mechanism using predicate logic inference problems and planning in block-world.

The mechanism uses a biologically inspired cognitive architecture, which is based on relatively compact Working Memory and larger synaptic Long-Term-Memory which stores knowledge that constrains the neural activation of the WM and forms attractors in its dynamics. In this architecture, knowledge items are retrieved from LTM into the WM only upon need, and, graph-like structures, that represent solution inferences, emerge at thermal equilibrium as an activation pattern of the neural units. Our architecture is based on the fact that Boltzmann Machines may be viewed as performing constraint satisfaction, where, at equilibrium, fixed-points maximally satisfy a set of weighted constraints. We show how to encode and bind arbitrary complex graphs as neural activation in WM and how a supervised learner may use miscalculations to adjust synapses so that constraints are better enforced, in order to correctly retrieve and process such complex structures. The architecture allows learning representations as expressive as First-Order-Logic (with bounded proof length), has no central control and is inherently robust to unit failures. The mechanism is goal directed in the sense, that the query may drive the processing, as well as the current activation pattern in the WM. It is universal and has a simple underlying computational principle. As such, it may be further adapted for applications that combine the advantages of both connectionist and traditional symbolic AI and may be used in modeling aspects of human' cognition.

3.25 Learning Action-oriented Symbols: Abstractions over Decision Processes

Subramanian Ramamoorthy (University of Edinburgh, GB)

Joint work of Ramamoorthy, Subramanian; Mahmud, M. M. Hassan; Hawasly, Majd; Rosman, Benjamin; Pokorny, Florian T.

A key question at the interface between sub-symbolic and symbolic learning and reasoning is that of how symbols can be acquired from experience, and grounded. Can such symbols be action-oriented in that they consistently abstract the underlying process?

I discuss two approaches we have recently developed for categorising policies obtained through processes such as reinforcement learning or motion planning in robots. The goal of categorisation is to arrive at a set of action- relevant symbols that better enable reasoning about changes associated with dynamic environments; taking a transfer/lifelong learning perspective.

The first approach is to cluster decision processes in terms of similarities in the effects of the actions. We define a novel distance and a clustering algorithm that yields a smaller set of decision processes that make continual transfer algorithms more effective.

The second approach draws on new mathematical tools from computational topology to abstract a set of trajectories associated with motion plans, yielding entirely qualitative descriptions of the underlying domain – which can again be used to separate quantitative detail from other global structural aspects of the tasks. I end by asking how these principles can be incorporated with a variety of models being studies by the NeSy community, including in particular deep networks.

References

1 M. M. H. Mahmud, M. Hawasly, B. Rosman, S. Ramamoorthy, Clustering Markov Decision Processes for continual transfer, arXiv:1311.3959 [cs.AI].
2 F. T. Pokorny, M. Hawasly, S. Ramamoorthy, Multiscale topological trajectory classification with persistent homology, In Proc. *Robotics: Science and Systems* (R:SS), 2014.

3.26 Mixing Low-Level and Semantic Features for Image Interpretation: A framework and a simple case study

Luciano Serafini (Fondazione Bruno Kessler and University of Trento, IT)

Joint work of Donadello, Ivan; Serafini, Luciano
Main reference I. Donadello, L. Serafini, "Mixing low-level and semantic features for image interpretation," in Proc. of the 1st Int'l Workshop on Computer Vision and Ontologies, to appear.

In recent years internet has seen a terrific increase of digital images. Thus the need of searching for images on the basis of human understandable descriptions, as in the case of textual documents, is emerging. For this reason, sites as YouTube, Facebook, Flickr, Grooveshark allow the tagging of the media and support search by keywords and by examples. Tagging activity is very stressful and often is not well done by users. For this reason automatic methods able to automatically generate a description of the image content, as in textual documents, become a real necessity. There are many approaches to image understanding which try to generate a high level description of an image by analysing low-level information (or features), such as colours, texture and contours, thus providing such a high level description in terms of semantic concepts, or high-level information. This would allow a person to search, for instance, for an image containing "a man is riding an horse". The difficulty to find the correspondence between the low-level features and the human concepts is the main problem in content-based image retrieval. It is the so-called *semantic gap* [2]. It's widely recognised that, to understand the content of an image, contextual information (aka background knowledge) is necessary [3]. Background knowledge, relevant to the context of an image, can be expressed in terms of logical languages in an ontology [4]. In image interpretation ontologies can be used for two main purposes. First, ontologies allow the expression of a set of constraints on the possible interpretations which can be constructed by considering only low-level features of an image. The satisfaction of such constraints can be checked via logical reasoning. Second, the terminology introduced in the ontology can be used as formal language to describe the content of the images. This will enable semantic image retrieval using queries expressed in the language introduced by the ontology. The background knowledge formalizes the semantics of the human understandable concepts and will provide the set of types of objects that can be found in a picture (e. g., horse, human, etc.) and the set of relations that can exist between depicted objects (e. g., rides is a relation between a human and an animal, part-of is a general relation between physical objects, etc.). Furthermore, the background knowledge provides constraints on types of objects and relations, e. g. a vehicle has at least two wheels or horses are animals that can be ridden by men. The advantage of having the tags as concepts coming from a background knowledge allows to reason over the image. For example the tag "horse" enables to infer the presence of an animal.

In the present work we adopt the natural idea that, already introduced for instance in [5, 6, 7] where an interpretation of a picture, in the context of an ontology, is a (partial) model of the ontology itself that expresses the state of affairs of the world in the precise moment in which the picture has been taken. We propose to formalize the notion of image interpretation, w.r.t. an ontology, as *a segmented image, where each segment is aligned with an object of a partial model of the reference ontology.* To cope with the fact that a picture reports only partial information on the state of affairs we use the notion of partial model of a logical theory [8]; to cope with the possibility of having multiple alternative interpretations of a picture we introduce the notion of *most plausible interpretation* an image, which is the interpretation that maximises some scoring function.

In order to have a preliminary evaluation of our idea, we implemented this framework, for a specific and limited case. We developed a fully unsupervised method to generate image interpretations able to infer the presence of complex objects from the parts present in the picture, thus inferring the relative "part-whole" structure. The method jointly exploits the constraints on the part-whole relation given by the ontology, and the low-level features of the objects available in the image. From a preliminary evaluation the presented approach shows promising results.

References

1 Donadello, I., Serafini, L.: Mixing low-level and semantic features for image interpretation. In: 1st International Workshop on Computer Vision and Ontologies at ECCV 2014, Zurich, Switzerland, Sept, 2014. (2014) To appear.
2 Liu, Y., Zhang, D., Lu, G., Ma, W.Y.: A survey of content-based image retrieval with high-level semantics. Pattern Recognition **40**(1) (January 2007) 262–282
3 Oliva, A., Torralba, A.: The role of context in object recognition. Trends in cognitive sciences **11**(12) (2007) 520–527

4 Bannour, H., Hudelot, C.: Towards ontologies for image interpretation and annotation. In Martinez, J.M., ed.: 9th International Workshop on Content-Based Multimedia Indexing, CBMI 2011, Madrid, Spain, June 13–15, 2011, IEEE (2011) 211–216
5 Reiter, R., Mackworth, A.K.: A logical framework for depiction and image interpretation. Artificial Intelligence **41**(2) (1989) 125–155
6 Neumann, B., Möller, R.: On scene interpretation with description logics. Image and Vision Computing **26**(1) (2008) 82–101 Cognitive Vision-Special Issue.
7 Straccia, U.: Reasoning within fuzzy description logics. J. Artif. Intell. Res. (JAIR) **14** (2001) 137–166
8 Staruch, B., Staruch, B.: First order theories for partial models. Studia Logica **80**(1) (2005) 105–120

3.27 Lifelong Machine Learning and Reasoning

Daniel L. Silver (Acadia University – Wolfville, CA)

Lifelong Machine Learning (LML) considers systems that learn many tasks over a lifetime, accurately and efficiently retaining and consolidating the knowledge they have learned and using that knowledge to more quickly and accurately learn new tasks [2, 1]. Since 1999, I have investigated aspects of LML for Learning to Classify (L2C) problem domains. In [3] I provide an overview of prior work in LML, present a framework for LML, and discuss its two essential ingredients – knowledge retention [4] and transfer learning [1]. Transfer learning is about using prior knowledge to more accurately develop models for a new task, from fewer training examples and in shorter periods of time. Knowledge retention is about efficient and effective methods of storing learned models for use in transfer learning and potentially reasoning. The proposed research program extends my prior work on LML to the learning of knowledge for purposes of reasoning. I am motivated by the belief that intelligent agents, like humans, should develop in their abilities as a function of their experience.

My previous research has focused on the theory and application of transfer learning and knowledge consolidation. We have published results on functional and representational knowledge transfer using multiple task learning (MTL), task rehearsal using synthesized training examples, and selective transfer for classification and regression problems [2]. Most significantly, we have developed context-sensitive MTL (csMTL); a transfer learning method that uses an additional context input, rather than an additional output for each new task [Silver09]. This approach overcomes a number of significant problems of standard MTL when applied to a LML

Our research has shown that knowledge of a new task can be integrated, or consolidated, with that of prior tasks in order for a LML solution to overcome the stability-plasticity problem and scale for practical use [4]. The stability-plasticity problem is the loss of prior task knowledge in a neural network when learning the examples of a new task [5]. Our work has demonstrated that MTL and csMTL networks can mitigate this problem by maintaining functional accuracy of prior tasks (stability) through the relearning, or rehearsal, of prior task examples while modifying the representation of the network (plasticity) through the learning new task examples. This can be accomplished using the back-propagation (BP) algorithm under the conditions described in [4, 5]. Recently, we have shown that a mix of proper selection of task rehearsal examples and more advanced methods of regularization can improve consolidation in csMTL networks.

In 2013, Qiang Yang and I encouraged the machine learning community to move beyond learning algorithms to systems that are capable of learning, retaining and using knowledge over a lifetime [3]. ML now has many practical applications of L2C; the next generation of ML needs to consider the acquisition of knowledge in a form that can be used for more general AI, such as Learning to Reason (L2R). We argue that opportunities for advances in AI lie at the locus of machine learning and knowledge representation; specifically, that methods of knowledge consolidation will provide insights into how to best represent knowledge for use in future learning and reasoning.

A survey of ML methods that create knowledge representations that can be used for learning and reasoning revealed three major bodies of work. The first is Neural-Symbolic Integration (NSI) [6, 8]. NSI research considers the benefits of integrating robust neural network learning with expressive symbolic reasoning capabilities. Much of NSI work focuses on the extraction of symbolic rules from trained network weights and the transfer of knowledge from logical expressions to network weights prior to training. Since the early 2000s, members of this community have called for a joint treatment of learning and reasoning [7]. At the IJCAI 2013 NeSy'13 workshop I presented an invited talk on the common ground shared by LML and NSI. I proposed an integrated framework for NSI and LML and discussed how the requirement of reasoning with learned knowledge places an additional constraint on the representational language and search methods used by LML systems. Learning is necessary to acquire knowledge for reasoning, however, reasoning informs us about the best ways to store and access knowledge. Thus, learning and reasoning are complimentary and should be studied together. Recent work at CMU on the NELL system agrees with this combined view [9]. The second major body of work is Learning to Reason (L2R) [10, 11], also referred to as Knowledge Infusion [12, 14, 15]. L2R work is not as abundant as that of NSI; however, it suggests a promising approach to developing is most promising in terms of our proposed research. The L2R framework is concerned with both learning a knowledge representation and with it doing deductive reasoning. The perspective is that an agent only needs to learn the knowledge required to reason in his environment, and to the level of performance demanded by that environment. Unlike prior approaches to engineering common knowledge, such as Cyc [16], L2R takes a probabilistic perspective on learning and reasoning. An L2R agent need not answer all possible knowledge queries, but only those that are relevant to the environment of the agent in a probably approximately correct (PAC) sense; that is, assertions can be learned to a desired level of accuracy with a desired level of confidence [12]. In [10] and [12] both authors show that a L2R framework allows for efficient learning of Boolean logical assertions in the PAC-sense (polynomial in the number of variables and training examples). Further to this, they prove that the knowledge learned can be used to efficiently reason with a similar level of accuracy and confidence. In this way, L2R agents are robust learners, acquiring most accurately the common knowledge that they need to reason in accord with their environment [12]. The authors make the point that traditional artificial intelligence has chosen knowledge representations for their transparency (e. g. preferring CNF over DNF representations) whereas the L2R framework chooses knowledge representations because they are learnable and facilitate reasoning. The third body of work is Deep Learning Architectures (DLA) and includes recent publications on Semi-supervised Learning [17], Co-training [18], Self-taught Learning [24], Representation Learning [20, 21], and Deep Learning [25, 26, 28, 22]. All share a common interest with LML in that they develop knowledge representations of the world from examples that can be used for future learning. This fall we are finalizing a survey of transfer learning and consolidation methods using DLAs.

My future research goals are to (1) develop and test Lifelong Machine Learning and Reasoning (LMLR) systems that can retain learned knowledge in a form that can be used for reasoning as well as future learning; and to (2) study the practical benefits and limitations of a prototype LMLR system applied to real-world problems in data mining and intelligent agents. To advance on the first goal, we will develop a system that can learn a series of logic assertions, such as $A|B \Rightarrow C$ and $C \Rightarrow D$, from examples of those expressions. The resulting knowledge base model can then be used to reason that $A \Rightarrow D$ by testing the model with examples. To advance on the second goal, I will scale the system up such that it can learn to reason from images that encode similar assertions. Such a system could be used by an intelligent agent to provide recommendations on next best action.

This work will create new theory on the learning and representation of knowledge from examples acquired from the learner's environment and methods by which to reason using that learned knowledge. Finding solutions to consolidating new with prior knowledge from examples that contain only part of the input space will be a major challenge. The methods and findings will be of interest to researchers working on machine learning, knowledge representation, reasoning, and applied areas such as data mining and intelligent agents.

References

1 Silver, D. and Mercer, R. 2009. Life-long Learning Through Task Rehearsal and Selective Knowledge Transfer. Theory and Novel Applications of Machine Learning, Editor: Meng Joo Er and Yi Zhou, IN-Tech Education and Pub, pp. 335–356.

2 Silver, D. and Poirier, R. and Currie, D. 2008. Inductive Transfer with Context-Sensitive Neural Networks. Machine Learning – Special Issue on Inductive Transfer, Springer, 73(3):313–336.

3 Silver, D., Yang, Q. and Li, L. Lifelong machine learning systems: Beyond learning algorithms. Proceedings of the AAAI Spring Symposium on Lifelong Machine Learning, Stanford University, AAAI, March, 2013, pp. 49–55.

4 Fowler, B. and Silver, D. 2011, Consolidation using Context-Sensitive Multiple Task Learning, Advances in Artificial Intelligence, 24th Conference of the Canadian Artificial Intelligence Association (AI 2011), St. John's, Nfld, May, 2011, P. Lingras (Ed.), Springer, LNAI 6657, pp. 128–139.

5 Silver, D. The Consolidation of Task Knowledge for Lifelong Machine Learning. Proceedings of the AAAI Spring Symposium on Lifelong Machine Learning, Stanford University, CA, AAAI, March, 2013, pp. 46–48.

6 Artur S. d'Avila Garcez, Luis C. Lamb and Dov M. Gabbay (2009). Neural-Symbolic Cognitive Reasoning. Cognitive Technologies, Springer, 2009.

7 Sebastian Bader, Pascal Hitzler, Steffen Hölldobler (2006). The Integration of Connectionism and First-Order Knowledge Representation and Reasoning as a Challenge for Artificial Intelligence. Information 9 (1).

8 Luis C. Lamb (2008). The Grand Challenges and Myths of Neural- Symbolic Computation. Dagstuhl Seminar on Recurrent Neural Networks, 2008.

9 A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E.R. Hruschka Jr. and T.M. Mitchell (2010). Toward an Architecture for Never-Ending Language Learning. In Proceedings of the 2010 Conference on Artificial Intelligence (AAAI-2010).

10 R. Khardon and D. Roth (1997), Learning to Reason, Journal of the ACM , vol. 44, 5, pp. 697–725.

11 R. Khardon and D. Roth (1999), Learning to Reason with a Restricted View, Machine Learning , vol. 35, 2, pp. 95–117.

12 Valiant, L. G. (2008). Knowledge Infusion: In Pursuit of Robustness in Artificial Intelligence. In FSTTCS (pp. 415–422).
13 Leslie G. Valiant (2012). A neuroidal architecture for cognitive computation. J. ACM 47(5):854–882.
14 Juba, Brendan (2012). Learning implicitly in reasoning in PAC- Semantics. arXiv preprint arXiv:1209.0056.
15 Juba, Brendan (2013). Implicit learning of common sense for reasoning. IJCAI- 2013.
16 Douglas Lenat and R. V. Guha (1990). Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project. Addison-Wesley.
17 Chapelle, Olivier; Schölkopf, Bernhard; Zien, Alexander (2006). Semi-supervised learning. Cambridge, Mass, MIT Press.
18 Blum, A., Mitchell, T (1998). Combining labeled and unlabeled data with co-training. COLT: Proceedings of the Workshop on Computational Learning Theory, Morgan Kaufmann, pp. 92–100.
19 Bengio, Y. (2011). Deep learning of representations for unsupervised and transfer learning. In JMLR W&CP: Proc. Unsupervised and Transfer Learning.
20 Yoshua Bengio, Aaron Courville and Pascal Vincent (2013). Representation Learning: A Review and New Perspectives in: Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35:8(1798–1828)
21 Yoshua Bengio (2013), Deep learning of representations: looking forward, in: Statistical Language and Speech Processing, pages 1–37, Springer.
22 Adam Coates, Andrej Karpathy, and Andrew Y. Ng (2012). Emergence of Object-Selective Features in Unsupervised Feature Learning. In NIPS 2012.
23 Raina, R., Madhavan, A., and Ng, A. Y. (2009). Large-scale deep unsupervised learning using graphics processors. In L. Bottou and M. Littman, editors, ICML 2009, pages 873-880, New York, NY, USA.
24 Raina, R., Battle, A., Lee, H., Packer, B., and Ng, A. Y. (2007). Self- taught learning: transfer learning from unlabeled data. In ICML-2007.
25 Hinton, G. E., Osindero, S., and Teh, Y. (2006). A fast learning algorithm for deep belief nets. Neural Computation, 18, 1527–1554.
26 Hinton, G. E. and Salakhutdinov, R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504–507.
27 Hinton, G., Krizhevsky, A., and Wang, S. (2011). Transforming auto- encoders. In ICANN-2011.
28 Le, Q.; Ranzato, M.; Monga, R.; Devin, M.; Chen, K.; Corrado, G.; Dean, J.; and Ng, A. (2012) Building high-level features using large scale unsupervised learning. In Proc. of International Conference in Machine Learning.

3.28 Representation Reuse for Transfer Learning

Son Tran (City University London, GB)

Joint work of Tran, Son; Artur d'Avila Garcez

The recent success of representation learning is built upon the learning of relevant features, in particular from unlabelled data available in different domains. This raises the question of how to transfer and reuse such knowledge effectively so that the learning of a new task can be made easier or be improved. This poses a difficult challenge for the area of transfer learning where there is no label in the source data, and no source data is ever transferred to the target domain. In previous work, the most capable approach has been self- taught learning which, however, relies heavily upon the compatibility across the domains. In this talk, I propose a novel transfer learning framework called Adaptive Transferred-profile Likelihood Learning (aTPL), which performs transformations on the representations to be transferred, so that they become more compatible with the target domain. At the same time, it learns supplementary knowledge about the target domain. Experiments on five datasets demonstrate the effectiveness of the approach in comparison with self- taught learning and other common feature extraction methods. The results also indicate that the new transfer method is less reliant on source and target domain similarity, and show how the proposed form of adaptation can be useful in the case of negative transfer.

3.29 Decoding the Symbols of Life: Learning Cell Types and Properties from RNA Sequencing Data

Joshua Welch (University of North Carolina – Chapel Hill, US)

Recent breakthroughs in biochemical techniques for low-input sequencing have enabled whole-transcriptome quantification (RNA-seq) of single cells. This technology enables molecular biologists to dissect the relationship between gene expression and cellular function at unprecedented resolution. In particular, the cell type composition of tissues is now open to investigation. There is a need for unsupervised learning approaches that can identify cell types and properties from single-cell RNA sequencing data in a purely unbiased manner, rather than relying on previously known cell type markers. This task of identifying cell types and the relationships between them is not unlike recognizing symbols in text or images. An overview of the relevant biological questions, single-cell RNA sequencing technology, and existing approaches to solving this problem are presented. The goal of the talk is to initiate discussion about how neural- symbolic approaches can be used to identify cell types and their properties from single-cell RNA sequencing data.

4 Working Groups

During the workshop several working groups were formed, and lively discussions on relevant research challenges took place. Next, we briefly summarize the results and questions raised during the breakout sessions.

4.1 Consolidation of learned knowledge

This discussion session was related to the concept of Learning to Reason, as investigated by Valiant, Khardon, Roth and many others. Significant advances in AI lie at the locus of machine learning and knowledge representation; specifically, methods of knowledge consolidation will provide insights into how to best represent common knowledge for use in future learning and reasoning. Knowledge consolidation is about efficient and effective methods of sequentially storing knowledge as it is learned. Overcoming the stability-plasticity problem is the main challenge here. Consolidation in neural networks can occur through a slow process of interleaved learning of a new and old task examples within a large multiple task learning (MTL) network. Task rehearsal is one approach to overcoming the stability-plasticity problem of forgetting of previously learned tasks stored in a MTL network by relearning synthesized examples of those tasks while simultaneously learning a new task. However this method faces scaling problems as the number of prior task examples increases. This session discussed new approaches to overcoming the stability plasticity problem so that knowledge consolidation is tractable over long sequences of learning.

4.2 Affordabilities and actionability

Inspired by points raised by Feldman, this session started from the ancient idea that the goal of thought is "truth", which has been productive, but it is also limiting. There are multiple reasons to believe that replacing "truth" with "actionability" will be more fruitful and that this move is necessary for a unified cognitive science. For more on this topic the reader is invited to the work on affordances by Feldman: ftp://ftp.icsi.berkeley.edu/pub/feldman/affordances.jf.pdf

4.3 Closing the gap in the pipeline: how to use learned knowledge for reasoning

Deductive and inductive approaches are natural allies. The former uses high-level conceptualizations to logically reason over data, while the latter focuses on finding higher-level patterns in data. This perspective suggests a rather obvious workflow in which inductive and statistical methods analyze data, resulting in metadata which describes higher level features of the data, which in turn enables the use of the data in intelligent systems. However, this apparently obvious pipeline is broken since the current state of the art leaves gaps which need to be bridged by new innovations. It would be helpful to start establishing the exact nature of these gaps, and to brainstorm about ways how to address these. Advances on this topic should provide added value for large-scale data management and analysis.

4.4 What is symbolic, what is sub-symbolic?

An old debate took place: what is the meaning of the terms symbolic and sub-symbolic in neural computation? Several questions were raised and analyzed. Certain neural networks are symbolic while others are not. What are factors that determine this? How can recognition be performed with symbolic networks? How can recall necessary for reasoning be performed with non-symbolic networks? How can both recognition and recall be achieved with the same networks?

4.5 How far can nature inspire us?

Among all the nature-inspired computation techniques, neural networks are about the only ones to have made their way into knowledge representation and reasoning so far. What about swarm computing or evolutionary computing? Could they also have a role and a use for learning and reasoning problems? The conclusion is that this is a discussion we can have looking at existing prototypes and ongoing research, including recent progress in the area of autonomous agents and multi-agent systems.

4.6 Demos & Implementation Fest

Finally, a hands on session took place. A lively display of neural-symbolic tools was presented by a number of the seminar’s participants. The participants had the opportunity to showcase their NeSy related software and get others to try, evaluate and discuss their work. Future extensions and integrations of the showcased work were proposed. Most participants had the opportunity to experiment with existing tools and prototypes that use state-of-the-art neural-symbolic computation techniques for image, audio, video and multimodal learning and reasoning.

5 Open Problems

After each discussion session, challenges and open problems were identified. It is clear that a number of research avenues lay ahead of the communities that participated in the seminar. The list below reflects, in part, the interdisciplinary nature of the research presented and the open problems identified at the seminar, leading to interesting future developments and applications. A companion paper [9] complements the list below and also identifies several opportunities and challenges for the neural-symbolic community.

- Over the last decades, most of the work has been focused on propositional approaches, which was seen as *propositional fixation* by McCarthy [1]. However, novel approaches have significantly contributed to the representation of other logical systems in neural networks, leading to successful application in temporal specification and synchronization [2, 3], distributed knowledge representation [4, 5] and even fragments of first-order logic inference [6]. In order to make progress in this open problem, perhaps one should consider logics of intermediate expressiveness such as description logics of the Horn family [7]. There remains a number of open issues in knowledge representation and reasoning in neural networks, in particular with regard to learning. The integration of neural-symbolic systems and inductive logic programming [8] may also lead to relevant developments. The companion paper [9] also identifies challenges in this area.
- Recently, it has been shown that neural networks are able to learn sequences of actions, a point raised by Icard during the discussions. Thus, it may well be possible that a “mental simulation” of some concrete, temporally extended activity can be effected by connectionist models. Theories of action, based on propositional dynamic logic can thus be useful. Feldman in [10] has argued that if the brain is not a network of neurons that represent things, but a network of neurons that do things, action models would probably be central in this endeavour.
- With respect to how the brain actually represents knowledge, perhaps one can draw inspiration from advances in fMRI. The work of Endres and Foldiak [11] may lead to a biologically sound model of the brain’s semantic structures. It can also contribute to the construction of new learning algorithms, by contributing to identifying the functioning of the brain’s learning mechanisms.

- There is much work to be done with respect to learning to reason (L2R) in neural networks [12, 13]. A question raised by Silver is how a L2R agent can develop a complete knowledge-base over time when examples of the logical expressions arrive with values for only part of the input space. Perhaps a Lifelong Machine Learning (LML) approach is needed. Such an approach can integrate, or consolidate, the knowledge of individual examples over many learning episodes [14]. Consolidation of learned knowledge is a necessary requirement as it facilitates the efficient and effective retention and transfer of knowledge when learning a new task. It is also a challenge for neural-symbolic integration because of the computational complexity of knowledge extraction, in general, and the need for compact representations that would enable efficient reasoning about what has been learned.
- Deep networks represent knowledge at different levels of abstraction in a modular way. This may be related to the fibring of neural networks and the representation of modal logics in neural networks, which are intrinsically modular [4, 5] and decidable, offering a sweet spot in the complexity-expressiveness landscape [15]. Modularity of deep networks seem suitable to knowledge extraction, which may help reduce the computational complexity of extraction algorithms [16], contributing to *close the gap in the pipeline* and leading to potential advances in lifelong learning, transfer learning, and applications.
- Applications: neural-symbolic computation techniques and tools have been applied effectively to action learning and knowledge description in videos [17, 18], argumentation learning in AI [19, 20], intelligent transportation systems to reduce CO_2 emissions [21], run-time verification and adaptation [22, 23], hardware/software requirements specification and verification [3, 22], normative reasoning [24], concept and ontology learning, in particular considering description logics and the semantic web [25, 26, 27], training and assessment in driving simulators, action learning and the extraction of descriptions from videos [17]. The lively demo fest organized at the seminar showed the reach of the field where promising prototypes and tools were demonstrated.

References

1 J. McCarthy. Epistemological problems for connectionism. Behav-ioral and Brain Sciences, 11:44, 1988.
2 Luís C. Lamb, Rafael V. Borges, Artur S. d'Avila Garcez: A Connectionist Cognitive Model for Temporal Synchronisation and Learning. AAAI Conference on Artificial Intelligence, AAAI-07:827–832, 2007.
3 Rafael V. Borges, Artur S. d'Avila Garcez, Luís C. Lamb: Learning and Representing Temporal Knowledge in Recurrent Networks. IEEE Transactions on Neural Networks 22(12): 2409–2421, 2011.
4 A. S. d'Avila Garcez, Luís C. Lamb, Dov M. Gabbay: Connectionist computations of intuitionistic reasoning. Theor. Comput. Sci. 358(1):34–55, 2006.
5 A. S. d'Avila Garcez, L. C. Lamb and D.M. Gabbay. Connectionist Modal Logic: Representing Modalities in Neural Networks. Theoretical Computer Science, 371(1–2):34–53, 2007.
6 S. Bader, P. Hitzler, and S. Hoelldobler, 2008. Connectionist model generation: A first-order approach. Neurocomputing, 71:2420–2432, 2008.
7 M. Kroetzsch, S. Rudolph, and P. Hitzler. Complexity of Horn de-scription logics. ACM Trans. Comput. Logic, 14(1), 2013.
8 Stephen Muggleton, Luc De Raedt, David Poole, Ivan Bratko, Peter A. Flach, Katsumi Inoue, Ashwin Srinivasan: ILP turns 20 – Biography and future challenges. Machine Learning 86(1):3–23, 2012.

9 Artur d'Avila Garcez, Tarek R. Besold, Luc de Raedt, Peter Foeldiak, Pascal Hitzler, Thomas Icard, Kai-Uwe Kuehnberger, Luis C. Lamb, Risto Miikkulainen, Daniel L. Silver. Neural-Symbolic Learning and Reasoning: Contributions and Challenges. Proceedings of the AAAI Spring Symposium on Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches, Stanford, March 2015.
10 J. A. Feldman, 2006. From Molecule to Metaphor: A Neural Theory of Language, Bradford Books, Cambridge, MA: MIT Press
11 D. M. Endres and P. Foldiak. An application of formal concept analysis to semantic neural decoding. Ann. Math. Artif. Intell.57:233–248, 2009.
12 R. Khardon and D. Roth, Learning to Reason, Journal of the ACM, vol. 44, 5, pp. 697–725, 1997.
13 Valiant, L. G. Knowledge Infusion: In Pursuit of Robustness in Artificial Intelligence. FSTTCS, pp. 415–422, 2008.
14 D. Silver. The Consolidation of Task Knowledge for Lifelong Machine Learning. Proc. AAAI Spring Symposium on Lifelong Machine Learning, Stanford University, pp. 46–48, 2013.
15 M. Y. Vardi: Why is Modal Logic So Robustly Decidable? Descriptive Complexity and Finite Models, 149–184, 1996.
16 S. Tran and A. d'Avila Garcez. Knowledge Extraction from Deep Belief Networks for Images. Proc. IJCAI Workshop on Neural-Symbolic Learning and Reasoning, NeSy'13, Beijing, 2013.
17 Leo de Penning, Artur S. d'Avila Garcez, Luís C. Lamb, John-Jules Ch. Meyer: A Neural-Symbolic Cognitive Agent for Online Learning and Reasoning. IJCAI 2011:1653–1658, 2011
18 Leo de Penning, Artur S. d'Avila Garcez, John-Jules Ch. Meyer: Dreaming Machines: On multimodal fusion and information retrieval using neural-symbolic cognitive agents. ICCSW 2013:89–94, 2013
19 Artur S. d'Avila Garcez, Dov M. Gabbay, Luís C. Lamb: Value-based Argumentation Frameworks as Neural-symbolic Learning Systems. J. Log. Comput. 15(6):1041–1058, 2005.
20 A. S. d'Avila Garcez, Dov M. Gabbay, Luís C. Lamb: A neural cognitive model of argumentation with application to legal inference and decision making. J. Applied Logic 12(2):109–127, 2014.
21 Leo de Penning, Artur S. d'Avila Garcez, Luís C. Lamb, Arjan Stuiver, John-Jules Ch. Meyer: Applying Neural-Symbolic Cognitive Agents in Intelligent Transport Systems to reduce CO2 emissions. IJCNN 2014:55–62.
22 Rafael V. Borges, Artur S. d'Avila Garcez, Luís C. Lamb, Bashar Nuseibeh: Learning to adapt requirements specifications of evolving systems. ICSE 2011:856–859, 2011.
23 Alan Perotti, Artur S. d'Avila Garcez, Guido Boella: Neural Networks for Runtime Verification. IJCNN 2014:2637–2644.
24 Guido Boella, Silvano Colombo, Artur S. d'Avila Garcez, Valerio Genovese, Alan Perotti, Leendert van der Torre: Learning and reasoning about norms using neural-symbolic systems. AAMAS 2012:1023–1030.
25 C. d'Amato, Volha Bryl and Luciano Serafini. Semantic Knowledge Discovery from Heterogeneous Data Sources. In Proc. of the Knowledge Engineering and Knowledge Management – 18th International Conference (EKAW'12), vol. 7603 pp. 26–31, Springer, LNCS (2012).
26 P. Hitzler, S. Bader, and A. d'Avila Garcez, 2005. Ontology learning as a use case for neural-symbolic integration. In Proc. Workshop on Neural-Symbolic Learning and Reasoning, NeSy'05 at IJCAI-05.
27 F. A. Lisi. A declarative modeling language for concept learning in description logics. In: Riguzzi, F., Zelezny, F. (eds.) Inductive Logic Programming, 22nd International Conference, ILP 2012, Dubrovnik, Croatia, September 17–19, 2012, Revised Selected Papers. Lecture Notes in Computer Science, vol. 7842. Springer, 2013.

Participants

- Tsvi Achler
IBM Almaden Center, US
- Jim Benvenuto
HRL Labs – Malibu, US
- Tarek R. Besold
Universität Osnabrück, DE
- Srikanth Cherla
City University – London, GB
- Claudia d'Amato
University of Bari, IT
- Artur d'Avila Garcez
City University – London, GB
- James Christopher Davidson
Google Inc. – Mountain View, US
- Leo de Penning
TNO Behaviour and Societal Sciences – Soesterberg, NL
- Luc De Raedt
KU Leuven, BE
- Natalia Díaz-Rodríguez
Turku Centre for Computer Science, FI
- Dominik Endres
Universität Marburg, DE
- Jacqueline Fairley
Emory University – Atlanta, US
- Jerry A. Feldman
University of California – Berkeley, US
- Peter Földiak
University of St Andrews, GB
- Manoel Franca
City University – London, GB
- Christophe D. M. Gueret
DANS – Den Hague, NL
- Biao Han
NUDT – Hunan, CN
- Pascal Hitzler
Wright State University – Dayton, US
- Steffen Hölldobler
TU Dresden, DE
- Thomas Icard
Stanford University, US
- Randal A. Koene
Carboncopies – San Francisco, US
- Kai-Uwe Kühnberger
Universität Osnabrück, DE
- Luis Lamb
Federal University of Rio Grande do Sul, BR
- Francesca Alessandra Lisi
University of Bari, IT
- Dragos Margineantu
Boeing Research & Technology – Seattle, US
- Vivien Mast
Universität Bremen, DE
- Risto Miikkulainen
University of Texas – Austin, US
- Andrey Mokhov
Newcastle University, GB
- Bernd Neumann
Universität Hamburg, DE
- Günther Palm
Universität Ulm, DE
- Alan Perotti
University of Turin, IT
- Gadi Pinkas
Or Yehudah & Bar Ilan University, IL
- Subramanian Ramamoorthy
University of Edinburgh, GB
- Luciano Serafini
Bruno Kessler Foundation – Trento, IT
- Daniel L. Silver
Acadia Univ. – Wolfville, CA
- Son Tran
City University – London, GB
- Joshua Welch
University of North Carolina – Chapel Hill, US
- Mark Wernsdorfer
Universität Bamberg, DE
- Thomas Wischgoll
Wright State University – Dayton, US