Neural-Symbolic Learning and Reasoning: Contributions and Challenges

Conference Paper · March 2015

DOI: 10.13140/2.1.1779.4243

10 authors, including:

Luc De Raedt
University of Leuven

Peter Földiák

Luís da Cunha Lamb
Universidade Federal do Rio Grande do Sul

Daniel L. Silver
Acadia University

Neural-Symbolic Learning and Reasoning: Contributions and Challenges

Artur d'Avila Garcez[1], Tarek R. Besold[2], Luc de Raedt[3], Peter Földiak[4], Pascal Hitzler[5], Thomas Icard[6], Kai-Uwe Kühnberger[2], Luis C. Lamb[7], Risto Miikkulainen[8], Daniel L. Silver[9]

[1]City University London, UK; [2]Universität Osnabrück, Germany; [3]KU Leuven, Belgium; [4]Univ. of St. Andrews, UK; [5]Wright State University, OH; [6]Stanford University, CA; [7]UFRGS, Brazil; [8]UT Austin, TX; [9]Acadia University, Canada

Abstract

Neural-symbolic computation aims at integrating robust connectionist learning algorithms with sound symbolic reasoning. The recent impact of neural learning, in particular of deep networks, has led to the creation of new representations that have, so far, not really been used for reasoning. Results on neural-symbolic computation have shown to offer powerful alternatives for knowledge representation, learning and inference in neural computation. This paper presents key challenges and contributions of neural-symbolic computation to this area.

1. Introduction

In order to respond to one of the key challenges of AI - the effective and transparent integration of learning and reasoning (Valiant 2008) - both symbolic inference and statistical learning should be integrated in an effective way. But over the last three decades, statistical learning and symbolic reasoning have been largely developed by distinct communities (but see below). Recent developments in deep neural networks are strongly connected to and have contributed novel insights into representational issues. But so far these representations are still rather low level, and have not really been integrated with the high-level symbolic representations used in knowledge representation. It is exactly in this niche that the area of neural symbolic learning and reasoning has been active for over two decades and has contributed many relevant representations and reasoning techniques that are very relevant to deep learning. Neural-Symbolic Learning and Reasoning integrates principles of neural network learning and logical reasoning. It is an interdisciplinary field involving components of knowledge representation, neuroscience, machine learning and cognitive science. Another area that is very relevant to Valiant's challenge is that of statistical relational learning and learning and probabilistic logic learning (Getoor et al., 2007; De Raedt et al., 2007), which aim at integrating probabilistic graphical models rather than connectionist methods with logical and relational reasoning. This note provides a brief overview of some of the achievements in neural-symbolic computation and also outlines some of the key challenges and opportunities. These key challenges were identified at a recent Dagstuhl seminar on Neural-Symbolic Learning and Reasoning, in Wadern, Germany (2014), which was the tenth meeting in a Series that started at IJCAI-05. This paper presents a summary of the main issues discussed, and outlines the challenges and opportunities identified collectively at the seminar. For details about the seminar's presentations, please visit: http://www.dagstuhl.de/14381.

The integration of the symbolic and connectionist paradigms of AI has been pursued by a relatively small research community over the last two decades and has yielded several significant results. Over the last decade, neural-symbolic systems have been shown capable of overcoming the *propositional fixation* of neural networks pointed out by McCarthy (1988) in response to Smolensky's *on the proper treatment of connectionism* (1988); see also (Hinton, 1990). Neural networks were shown capable of representing modal and temporal logics (d'Avila Garcez and Lamb, 2006; d'Avila Garcez, Lamb, and Gabbay, 2007) and fragments of first-order logic (Bader, Hitzler, Hölldobler, 2008; d'Avila Garcez, Lamb, Gabbay, 2009). Further, neural-symbolic systems have been applied to a number of real-world problems in the areas of bioinformatics, control engineering, software verification and adaptation, visual intelligence, ontology learning, and computer games (Borges, d'Avila Garcez, and Lamb, 2011; de Penning et al., 2011; Hitzler, Bader and d'Avila Garcez, 2005). Most of the work on knowledge representation and learning in neural networks has focused on variable-free logic fragments. However, one should note that several approaches have dealt with alternative formalizations of variable binding, and the representation of relations (Bader, Hitzler, and Hölldobler, 2008; d'Avila Garcez, Lamb, and Gabbay, 2009; Pinkas, Lima, and Cohen, 2012; Franca,,d'Avila Garcez and Zaverucha 2014).

In deep learning (Hinton, Osindero, and The, 2006), the generalization of symbolic rules may be a crucial process, but it is not fully understood yet (S. Tran and A. d'Avila Garcez, 2013). Deep architectures have to manage issues such as representation abstraction levels, modularity, parallel and distributed representations. Several techniques developed under the umbrella of neural-symbolic computation can be useful towards this goal. For instance, fibring neural networks may offer the expression of levels of symbolic abstraction. Connectionist modal logics are modular by construction; temporal and modal logic representations are distributed (d'Avila Garcez, Lamb and Gabbay, 2009).

This paper outlines the main challenges and opportunities toward this aim, according to the main discussions held at the Dagstuhl seminar on Neural-Symbolic Learning and Reasoning: (i) Structure learning remains to be fully understood, whether it consists of hypothesis search at the concept level, including (probabilistic) Inductive Logic Programming (ILP) and statistical AI approaches, or iterative adaptation processes such as Hebbian learning or contrastive divergence. (ii) The learning of generalizations of symbolic rules is a crucial process and not well understood; the adoption of neural networks that can offer some degree of modularity, such as deep networks, and the neural-symbolic methods for knowledge insertion into neural networks may help shed light into this process. (iii) Effective knowledge extraction from large-scale networks remains a challenge; computational complexity issues and the provision of compact, expressive descriptions continue to be a barrier for explanation, lifelong learning and transfer. Topics (i) to (iii) above open up a number of research opportunities, to be discussed in what follows.

2. State-of-The-Art Results and Challenges

Representation: Most of the work on neural-symbolic learning and reasoning has been focused on propositional logics, essentially adhering to what John McCarthy called neural networks' propositional fixation (McCarthy, 1988). Early approaches were based essentially on the connectionist representation of propositional logic, a line of research which has since been substantially extended to other finitary logics (d'Avila Garcez, Lamb, Gabbay 2009).

Some primary proposals for overcoming the propositional fixation of neural networks include (Gust, Kühnberger, and Geibel, 2007) which leverages variable-free representations of predicate logic using category-theoretic Topoi, (Bader, Hitzler, and Hölldobler, 2008) for predicate Horn logic programs (with function symbols) which utilizes an encoding of logic as vectors of real numbers mediated by the Cantor set, and (Guillame-Bert, Broda, and d'Avila Garcez, 2010) for learning first-order rules based on term encoding also as vectors. These systems have been shown to work in limited proof-of-concept settings or small examples, and attempts to achieve useful performance in practice have so far not been successful.

To advance, it may well be necessary to consider logics of intermediate expressiveness, such as description logics (DL), in particular logics in the Horn DL family (Krötzsch, S. Rudolph, and P. Hitzler, 2013), so-called propositionalization methods, as used by ILP (Blockeel et. al, 2011; França, Zaverucha, and d'Avila Garcez, 2014) and answer-set programming (Lifschitz, 2002), and propositional modal logic (d'Avila Garcez, Lamb and, Gabbay, 2009), known to be strictly more expressive than propositional logic and yet decidable. More recent results regarding the integration of description logics and rules (Krisnadhi, Maier, and Hitzler, 2011, Krötzsch et al., 2011) indicate the feasibility of the approach w.r.t. neural-symbolic integration (Hitzler, Bader and d'Avila Garcez, 2005). The variable binding problem, though, and the question of how neural networks reason with variables remain central to this enterprise (d'Avila Garcez, Broda, Gabbay 2002; Feldman, 2006; Pinkas, Lima, and Cohen, 2012).

There has been much successful work in the neural-symbolic computation community on extracting logical expressions such as logic programs from trained neural networks, and using this extracted knowledge to seed learning in further tasks (see d'Avila Garcez, Lamb, and Gabbay (2009) for an overview). Meanwhile, there has been some suggestive recent work showing that neural networks can learn entire sequences of actions, thus amounting to "mental simulation" of some concrete, temporally extended activity (many references). There is also a very well developed logical theory of action, for instance related to the basic propositional logic of programs (Propositional Dynamic Logic; see Harel, Kozen, and Tiuryn 2001), capturing what holds true after various combinations of actions. A natural place to extend the aforementioned work would be to explore extraction from a trained network exhibiting this kind of simulation behavior, a PDL expression capturing a high-level description of that sequence of actions. As argued by Feldman (2006), if the brain is not a network of neurons that represent things, but a network of neurons that do things, action models should be playing a central role.

As regards knowledge representation in the brain, one of the key challenges is to understand how neural activations, which are widely distributed and sub-symbolic, give rise to behavior that is symbolic, such as language and logical reasoning. Recent advances in fMRI and MEG analysis make it possible to develop and test such theories. For instance, formal concept analysis (Ganter and Wille, 1999; Endres and Foldiak, 2009) leads to characterization of semantic structures in the brain, and conceptual attribute representations (Binder and Desai 2011) make it possible to model how semantics concepts map to areas of the brain. A major challenge for the future is to understand how such semantics are constructed and affected by context, such as a sequence of words in a sentence.

Consolidation: Khardon and Roth introduced Learning to Reason (L2R) as a framework that makes learning an integral part of the reasoning process (Khardon and Roth, 1997). L2R studies the entire process of learning a knowledge base representation from examples of the truth-table of a logical expression, and then reasoning with that knowledge base by querying it with similar examples. Learning is done specifically for the purpose of reasoning with the learned knowledge in the Probably Approximately Correct (PAC) sense. This work has close connections to the *neuroidal* model of reasoning developed by Valiant

(2000) that examines computationally tractable learning and reasoning given PAC constraints. These constraints consider limiting the agent's environment via a probability distribution over the input space and relaxing performance bounds on learning and reasoning (Khardon and Roth, 1999). Consequently, this approach exhibits a non-monotonic reasoning behavior because error correction based on reasoning mistakes can be used, incrementally, to improve the knowledge base.

Despite a number of interesting early findings (Valiant, 2008, Juba, 2013), there is much work to be done to make this a practical approach to the integration of learning and reasoning. One major question 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. This suggests that a Lifelong Machine Learning (LML) approach is needed that can integrate, or consolidate, the knowledge of individual examples over many learning episodes (Silver, 2013a; Fowler, 2011). The consolidation of learned knowledge is seen here as a necessary requirement as it facilitates the efficient and effective retention and transfer of knowledge (e.g. rapid and beneficial inductive bias) when learning a new task (Silver, 2013b). 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, however, represent knowledge at different levels of abstraction in a modular way. This may be related to the representation of modal logics, which are intrinsically modular (d'Avila Garcez, Lamb, and Gabbay, 2007) and decidable, offering a *sweet spot* in the complexity-expressiveness landscape (Vardi, 1996). Modularity of deep networks seem suitable to (relational) knowledge extraction, which may help reduce the computational complexity of extraction algorithms (A. d'Avila Garcez, Broda, and Gabbay, 2001; Tran and d'Avila Garcez, 2013).

Transfer: Knowledge transfer between, at first site, unrelated domains is a characteristic feature and crucial cornerstone of human learning: As, for example, evidenced by the work in (Gentner, Holyoak, and Kokinov, 2001), analogy (the human ability of perceiving and operating on dissimilar domains as being similar with respect to certain aspects based on shared commonalities) is considered essential for learning abstract concepts or procedures and for adapting existing knowledge to newly encountered scenarios and contexts. Whilst most of the prominent computational models of analogy, such as, for instance, the Structure Mapping Engine (Falkenhainer, Forbus, and Gentner 1989) or Heuristic-Driven Theory Projection (Schmidt, Krumnack, Gust, Kühnberger, 2014) are logic-based and thus symbolic in nature, recent developments in structure learning in a neural-symbolic paradigm may open the way for a meaningful application of analogy also on a sub-symbolic level. The expected gain is enormous: instead of having to retrain a network model on a new domain, already obtained insights could meaningfully be transferred between different networks, giving subsequent models a *head start*. Also, by mutually informing learning processes between networks being trained in parallel, a speed-up in training could be expected. Still, many questions have to be answered with two of the most foundational ones being: How can the knowledge-level notion of analogical transfer practically be implemented in connectionist architectures? How can possible analogies between different domains be discovered on a sub-symbolic level in the first place? Some work such heterogeneous transfer learning (Yang et al, 2009) has been directed at these questions but there is much to do.

Application: From a more practical perspective, neural-symbolic integration has been applied to training and assessment in simulators, normative reasoning, rule learning, integration of run-time verification and adaptation, action learning and description in videos (Borges, d'Avila Garcez, Lamb, 2011; de Penning et al., 2011). Future application areas that seem promising include the analysis of complex networks, social robotics and health informatics, and multimodal learning and reasoning combining video and audio data with metadata. Overall, neural-symbolic integration seems suitable in application areas where large amounts of heterogeneous data are available and knowledge descriptions are needed, such as e.g. when video and audio data are tagged with ontological metadata (de Penning, 2011, Tran and d'Avila Garcez, 2013), including robot navigation and communication, health, genomics, hardware/software specification, multimodal data fusion for information retrieval, big data understanding and, ultimately, language understanding.

Several features illustrate the advantages of neural-symbolic computation when it comes to specific applications: its explanation capacity, no a-priori assumptions, its comprehensive cognitive models integrating symbolic and statistical learning with sound logical reasoning. Ultimately, however, in each of the above application areas, measurable criteria for comparison should include: accuracy and efficiency measures, knowledge readability.

3. Conclusions

Neural-symbolic computing as an area of research reaches out to two communities and seeks to achieve the fusion of competing views, when such fusion can be beneficial. In doing so, it sparks new ideas and promotes coopetition.

Further, neural-symbolic computation brings together an integrated methodological perspective, as it draws from both neuroscience and cognitive systems. Methodologically, it bridges several gaps, as new ideas can emerge from such a methodology through *changes of representation*. 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. Both its representational generality (the ability to represent, learn and reason about several symbolic systems) and its learning robustness can open several interesting opportunities leading to adequate forms of knowledge representation, be they purely symbolic, or hybrid combinations involving probabilistic or numerical representations.

References

S. Bader, P. Hitzler, and S. Hölldobler, 2008. Connectionist model generation: A first-order approach. Neurocomputing, 71:2420-2432, 2008.

J. R. Binder, R. Desai, 2011. The neurobiology of semantic memory. Trends in Cognitive Sciences, 15(11), 527-536.

H. Blockeel, K.M. Borgwardt, L. De Raedt, P. Domingos, K. Kersting, X. Yan, 2011: Guest editorial to the special issue on inductive logic programming, mining and learning in graphs and statistical relational learning. Machine Learning 83(2):133-135.

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

A. d'Avila Garcez, K. Broda, and D. M. Gabbay, 2001. Symbolic knowledge extraction from trained neural networks: A sound approach. Artificial Intelligence, 125:155- 207.

A. d'Avila Garcez, K. Broda, and D. M. Gabbay, 2002. Neural-Symbolic Learning Systems: Foundations and Applications. Perspectives in Neural Computing. Springer, Berlin.

A. d'Avila Garcez, L. Lamb, 2006. A Connectionist Computational Model for Epistemic and Temporal Reasoning. Neural Computation 18(7):1711-1738.

A. d'Avila Garcez, L. C. Lamb and D. M. Gabbay, 2007. Connectionist modal logic: Representing modalities in neural networks. Theoretical Computer Science 371(1-2):34-53.

A. d'Avila Garcez, L. C. Lamb, and D. M. Gabbay. 2009. Neural-Symbolic Cognitive Reasoning. Springer-Verlag.

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

De Raedt, Luc, et al., 2008. Probabilistic Inductive Logic Programming-Theory and Applications, vol. 4911, Lecture Notes in Computer Science.

D. M. Endres and P. Foldiak, 2009. An application of formal concept analysis to semantic neural decoding. Ann. Math. Artif. Intell.57:233-248.

B. Falkenhainer, K. Forbus, and D. Gentner, 1989. The Structure Mapping Engine: Algorithm and Examples. Artif. Intell. 41:1-63.

J. A. Feldman, 2006. From Molecule to Metaphor: A Neural Theory of Language, Bradford Books, Cambridge, MA: MIT Press.

B. Fowler, D. Silver 2011, Consolidation using Context-Sensitive Multiple Task Learning. 24th Conf. Canadian AI Assoc. Springer, LNAI 6657.

M.V. França, G. Zaverucha and A.S. d'Avila Garcez, 2014. Fast relational learning using bottom clause propositionalization with artificial neural networks. Machine Learning 94(1):81-104.

Ganter B., Wille R., 1999. Formal Concept Analysis: Mathematical Foundations. Springer, New York.

D. Gentner, K. Holyoak and B. Kokinov, eds. 2001. The Analogical Mind: Perspectives from Cognitive Science. MIT Press.

Getoor, Lise, and Ben Taskar, eds, 2007. Introduction to statistical relational learning. MIT press.

M. Guillame-Bert, K. Broda, and A. d'Avila Garcez, 2010. First-order logic learning in artificial neural networks. Intl. Joint Conference on Neural Networks, IJCNN 2010, Barcelona, Spain.

H. Gust, K.-U. Kühnberger, and P. Geibel, 2007. Learning models of predicate logical theories with neural networks based on topos theory. In B. Hammer and P. Hitzler, editors, Perspectives of Neural-Symbolic Integration, pages 233-264. Springer.

D. Harel, D. Kozen, J. Tiuryn, 2001: Dynamic logic. SIGACT News 32(1): 66-69.

G. E. Hinton, S. Osindero, Y. Teh, 2006: A fast learning algorithm for deep belief nets. Neural Computation 18, pp 1527-1554.

G. E. Hinton, 1990: Preface to the special issue on connectionist symbol processing. Artificial Intelligence 46, 1-4.

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.

B. Juba, 2013. Implicit learning of common sense for reasoning. IJCAI.

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

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

A. Krisnadhi, F. Maier, and P. Hitzler, 2011. OWL and Rules. In A. Pollereset et. al. Eds., Reasoning Web: Semantic Technologies for the Web of Data. 7th Summer School Galway, vol. 6848. LNCS. Springer.

M. Krötzsch, F. Maier, A. A. Krisnadhi, and P. Hitzler, 2011. A better uncle for OWL: Nominal schemas for integrating rules and ontologies. In S. Sadagopan, K. Ramamritham, A. Kumar, M. Ravindra, E. Bertino, and R. Kumar, eds. Proc. 20th Intl. World Wide Web Conference, Hyderabad, India, pages 645-654. ACM.

M. Krötzsch, S. Rudolph, and P. Hitzler, 2013. Complexity of Horn description logics. ACM Trans. Comput. Logic, 14(1).

J. Lehmann, S. Bader, and P. Hitzler, 2010. Extracting reduced logic programs from artificial neural networks. Appl. Intell. 32(3):249-266.

V. Lifschitz, 2002. Answer set programming and plan generation. Artif. Intell. 138(1-2): 39-54.

J. McCarthy, 1988. Epistemological problems for connectionism. Behavioral and Brain Sciences, 11:44.

W. McCulloch and W. Pitts, 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5:115-133.

G. Pinkas, P.V. Lima, S. Cohen, 2012. A Dynamic Binding Mechanism for Retrieving and Unifying Complex Predicate-Logic Knowledge. ICANN (1) 2012: 482-490.

M. Schmidt, U. Krumnack, H. Gust, and K.-U. Kühnberger, 2014. Heuristic-Driven Theory Projection: An overview. Computational Approaches to Analogical Reasoning: Current Trends. Springer. 163-194.

D. Silver, 2013a. The Consolidation of Task Knowledge for Lifelong Machine Learning. Proc. AAAI Spring Symposium on Lifelong Machine Learning, Stanford University, pp. 46-48.

D. Silver, 2013b. On Common Ground: Neural-Symbolic Integration and Lifelong Machine Learning. 9th Workshop on Neural-Symbolic Learning and Reasoning, NeSy'13 at IJCAI, Beijing, China.

P. Smolensky, 1988: On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11 (1):1-23.

S. Tran and A. d'Avila Garcez, 2013. Knowledge Extraction from Deep Belief Networks for Images. Proc. IJCAI Workshop on Neural-Symbolic Learning and Reasoning, NeSy'13, Beijing.

L.G. Valiant, 2000. A neuroidal architecture for cognitive computation. J. ACM 47(5): 854-882.

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

M. Y. Vardi, 1996: Why is Modal Logic So Robustly Decidable? Descriptive Complexity and Finite Models, 149-184.

Yang, Q.; Chen, Y.; Xue, G.-R.; Dai, W.; and Yu, Y. 2009. Heterogeneous transfer learning for image clustering via the social web. In *Proc. of the Joint Conf. of the 47th Annual Meeting of the ACL.*