

Reasoning with Deep Learning: an Open Challenge

Marco Lippi

DISMI – Università degli Studi di Modena e Reggio Emilia
marco.lippi@unimore.it

Abstract. Building machines capable of performing automated reasoning is one of the most complex but fascinating challenges in AI. In particular, providing an effective integration of learning and reasoning mechanisms is a long-standing research problem at the intersection of many different areas, such as machine learning, cognitive neuroscience, psychology, linguistic, and logic. The recent breakthrough achieved by deep learning methods in a variety of AI-related domains has opened novel research lines attempting to solve this complex and challenging task.

1 Motivation

In the last decade, deep learning has brought a real revolution in the area of artificial intelligence (AI) and in many of its related fields, producing stunning results in a variety of different application domains. In computer vision, image classification and object detection systems can now be trained to recognize thousands of different semantic categories [1], that sometimes are difficult to distinguish even for humans. Speech recognition and music retrieval can be performed with an accuracy that was hard to imagine only one decade ago [2]. For natural language processing and understanding, tasks such as machine translation or sentiment analysis have moved huge steps forward with respect to earlier state-of-the-art systems [3]. In addition, in many of such contexts, these successful applications have produced a tremendous impact also from a technological point of view, with all major ICT companies in the world (Google, Facebook, Microsoft, IBM, etc.) now actively working in the field of AI more than ever before, aiming to continuously develop more efficient and more accurate systems. Whereas these are indeed impressive advancements, there is no doubt that many of the problems that are really at the core of AI are far from being solved. This is particularly true for those tasks that have to deal with reasoning operations, such as induction, deduction, abduction, probabilistic inference, spatial or temporal reasoning, and especially combinations of those. We can now build machines that can easily and accurately translate a text between languages, that can spot whether an object appears in an image or in a video, that are capable of recognizing spoken language at very high accuracy levels, but which cannot yet answer higher-level questions related to the content they have just processed. Building a machine that can read any kind of short story, or watch a movie of any genre, and that

Marco Lippi

can answer simple questions about the plot and the characters, questions that a child would certainly be able to answer, still remains a dream. Clearly, these are extremely complex tasks, that humans learn to perform during the first years of life, and that involve learning to analyze large amounts of information, to extract and somehow store some form of knowledge from such information, and finally to digest this information and reason about it.

2 Methods

Historically, there has always been a dichotomy between symbolic and sub-symbolic (often named connectionist) frameworks to model reasoning [4]. The symbolic approach has its roots in the study of logic and philosophy, and it sees reasoning as the capability of deriving additional information from that already encoded in a collection of given symbols, by performing elaboration and manipulation on the given structured representations. From the perspective of connectionism, reasoning is instead the result or derivation of multiple, interconnected, simple processing devices, one major example being neural networks. The main motivation behind connectionism comes from cognitive neuroscience, since the human neural circuitry is clearly capable of storing and retrieving knowledge organized in short- and long-term memory, by continuously analyzing and processing new, complex information, and reasoning upon it.

2.1 Pioneering approaches

Throughout the years, there have been many attempts to combine learning and reasoning processes by integrating connectionist and symbolic paradigms. Between the 80s and the 90s, a significant number of pioneering works started to circulate, such as connectionist approaches to encode semantic networks [5], or knowledge-based artificial neural networks, named KBANNs [6]. Within this context, research has been mainly directed along distinct but strongly intertwined directions: (i) inserting background knowledge into the structure of neural networks, (ii) refining sets of rules via neural networks, (iii) extracting rules or classification patterns from trained neural networks.

The main idea behind KBANNs is that of considering input-to-output paths in a neural network as sub-symbolic realizations of some symbolic rules given in advance: output units can be thought of as the final conclusions of the rules, input units are supporting facts and hidden units represent intermediate conclusions. Standard backpropagation can be applied to tune the weights of the network, by employing a training set, as for standard neural networks. This framework can be adopted both for initializing the structure of a neural network with background knowledge, and to extract a set of refined rules from the final learned network, thus addressing the long-standing problem of neural network interpretability. Similar approaches had been proposed for Recurrent Neural Networks to handle sequential data [7]. Despite having shown promising results in computational biology tasks [6], KBANNs have found applications only in small-sized domains,

and encoding simple rules. One of the main limitations of this model was in fact due to the difficulty, in the 90s, of training deep neural networks, whose structure was induced by complex rules. In this direction, the recent advancements of deep learning could certainly offer a valuable contribution.

2.2 Combining symbolic and sub-symbolic methods

More recent attempts to combine symbolic and sub-symbolic techniques for reasoning include the research lines carried out by the so-called neural-symbolic community [8]. Several theoretical results have been successfully achieved in this area. Many studies have been conducted on the neural binding problem [9], that aims to explain how connections between different brain regions are coordinated, so as to retrieve and manipulate information, activate distant neural circuits, and finally perform reasoning. Other research has focused on the analysis of the capability of neural networks to represent modal and temporal logics [10] as well as fragments of first-order logic [11, 12]. Despite being successfully applied in some proof-of-concept settings, the existing neural-symbolic approaches still lack a thorough application on large-scale, real-world problems.

Starting from a slightly different perspective, the area of statistical relational learning [13] (also known as probabilistic inductive logic programming) was born at the end of the 90s with similar goals. Statistical relational learning aims to combine the expressive power of logic representations with models handling uncertainty in data, such as statistical learning approaches and graphical models. Few attempts have been made in the direction of employing neural networks within the context. An example is given by ground-specific Markov logic networks [14], that allow to embed neural networks within the Markov logic framework, by learning the weights of the probabilistic logic clauses. The method has been successfully applied to bioinformatics and time-series forecasting, for problems where there is a crucial need to model background knowledge, handle structured data, and perform probabilistic inference. Yet, it was never used to handle reasoning tasks.

2.3 Recent advances: deep learning

In the last years, the task of reasoning with (deep) connectionist models has captured an enormous interest, that is evidenced by the approaches that have been proposed by some of the big companies that are currently investing in deep learning. This is the case of Neural Turing Machines by Google DeepMind [15], Memory Networks developed at Facebook AI Research [16], Dynamic Memory Networks proposed by MetaMind [17], the Neural Reasoner [18] by Huawei Technologies, and the Watson system developed by IBM [19]. Additional methods that are worth mentioning in this context are Wolfram Alpha, a computational knowledge engine which is capable of handling and manipulating encyclopedic knowledge to perform question-answering, and the GeoS system developed by the Allen Institute [20] which can solve geometry Scholastic Aptitude Tests at the level of the average US students.

Marco Lippi

Many of such methods employ a purely sub-symbolic framework, relying on supervised datasets to train a deep architecture from collections of examples. Most of these approaches share the common idea that a connectionist model aiming to perform reasoning has to maintain some memory that has to be efficiently organized and queried in order to retrieve the information necessary to provide solutions for the desired tasks. Memory Networks, for example, use a dedicated neural network for each step in the process of retrieving the correct answer to a given question: (i) computing feature representations for the input, (ii) updating memory, (iii) combining input and memory to compute the output, (iv) translate the output into an interpretable answer. In their original implementation, such model is presented in a purely supervised fashion, but extensions to semi-supervised settings are considered as well.

3 Discussion

Although producing remarkable advancements, recent approaches to reasoning with deep networks do not properly address the task of symbolic reasoning, thus leaving the problem of neural network interpretability unsolved. Most of the effort is in fact demanded to an efficient management of the memory of the network, and to fast matching and retrieval algorithms. Some of the existing approaches have been compared on a collection of benchmarks, called bAbI tasks [21], developed at Facebook AI Research. Such tasks include simple question answering problems, that typically require to perform some kind of reasoning and answer with a single word. The following is an example:

*In the afternoon Julie went to the park. Yesterday Julie was at school. Julie went to the cinema this evening. Where did Julie go after the park ? **Cinema***

To answer such questions, the system needs to perform many, advanced operations. First, it has to process the text and store the information in some form of memory, since even a short story like the one in the above example contains plenty of information. Then, it has to understand which are pieces of knowledge that are relevant to a given question, in order to finally formulate some hypothesis and provide the correct answer. These final steps include complex reasoning mechanisms, such as deductions and uncertainty handling, as well as temporal reasoning. Such skills are completely different from the technology that is present in existing sophisticated question answering systems, that mainly exploit encyclopedic background knowledge and answer highly specific questions.

Big data. The recent, impressive success of deep learning across several, different areas of AI is certainly strongly related to the availability of huge datasets, that nowadays can be easily collected from various and heterogeneous data sources over the Web, and also to the advancements in computer hardware performance, that have dramatically reduced computational requirements. From a theoretical point of view, models that are currently employed in many systems were already known decades ago, but efficient techniques for training them

successfully have been proposed only in recent years [22]. This is the case, for example, of Convolutional and Recurrent Neural Networks, now representing the state-of-the-art in a wide number of tasks. The injection of background knowledge in the structure of such networks is yet to be investigated.

Unsupervised learning. Among the open challenges, a crucial point is to automatically extract knowledge from data, and to encode it into a neural network model, rather than employing expert-given knowledge. Clearly, most of the existing methods for information extraction and knowledge representation employ supervised or at least semi-supervised data. But, in the future we expect that a key contribution will come from unsupervised learning approaches, also to extract commonsense knowledge. The advantages of using unsupervised data are undeniable: generating labeled corpora is in fact an extremely complex, time-consuming and costly operation, whereas unsupervised data are everywhere, available in a variety of different domains (text, video, audio, etc.). Unsupervised learning algorithms could be employed to extract relevant features and patterns from data. Although some algorithms for unsupervised learning have played a crucial role for the development of the whole deep learning area, it is widely recognized that a proper use of unsupervised data is still missing [22].

Incremental learning. Humans naturally implement a lifelong learning scheme, continuously acquiring knowledge. Such a feature seems to be a crucial element for the development of reasoning skills and thus it is likely that future attempts to this task will need to implement a dynamic, on-line mechanism that incrementally acquires knowledge, possibly by also changing the network topology.

Beyond the Turing test ? Reasoning tasks could certainly be employed in an advanced version of the Turing test. Recently, the computer vision community has proposed the Visual Turing Challenge [25] where automated vision systems have to answer questions regarding the content of some images or videos, thus requiring both visual and linguistic skills. Also the bAbI tasks [21] already mentioned represent another example of benchmark that in future could be integrated with an advanced Turing test.

References

1. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in NIPS. (2012) 1097–1105
2. Lee, H., Pham., P.T., Largman, Y., Ng, A.Y.: Unsupervised feature learning for audio classification using convolutional deep belief networks. In: Advances in NIPS. (2009) 1096–1104
3. Cho, K., van Merriënboer B., Gülcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder-decoder for statistical machine translation. In: Proceedings of EMNLP. (2014) 1724–1734

4. Dinsmore, J.: The symbolic and connectionist paradigms: closing the gap. Lawrence Erlbaum (2014)
5. Shastri, L.: A connectionist approach to knowledge representation and limited inference. *Cognitive Science* **12** (1988) 331–392
6. Towell, G.G., Shavlik, J.W.: Knowledge-based artificial neural networks. *Artificial intelligence* **70** (1994) 119–165
7. Frasconi, P., Gori, M., Maggini, M., Soda, G.: Unified integration of explicit knowledge and learning by example in recurrent networks. *IEEE Transactions on Knowledge and Data Engineering* **7** (1995) 340–346
8. d’Avila Garcez, A., Gori, M., Hitzler, P., Lamb, L.C.: Neural-symbolic learning and reasoning (dagstuhl seminar 14381). *Dagstuhl Reports* **4** (2015)
9. Feldman, J.: The neural binding problem (s). *Cognitive neurodynamics* **7** (2013) 1–11
10. Garcez, A.S.d., Lamb, L.C.: A connectionist computational model for epistemic and temporal reasoning. *Neural Computation* **18** (2006) 1711–1738
11. Bader, S., Hitzler, P., Hölldobler, S.: Connectionist model generation: A first-order approach. *Neurocomputing* **71** (2008) 2420–2432
12. Garcez, A.S.d., Lamb, L.C., Gabbay, D.M.: *Neural-symbolic cognitive reasoning*. Springer Science & Business Media (2008)
13. Getoor, L., Taskar, B.: *Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning)*. The MIT Press (2007)
14. Lippi, M., Frasconi, P.: Prediction of protein β -residue contacts by markov logic networks with grounding-specific weights. *Bioinformatics* **25** (2009) 2326–2333
15. Graves, A., Wayne, G., Danihelka, I.: Neural Turing machines. *arXiv preprint arXiv:1410.5401* (2014)
16. Sukhbaatar, S., Weston, J., Fergus, R., et al.: End-to-end memory networks. In: *Advances in neural information processing systems*. (2015) 2440–2448
17. Socher, R., Chen, D., Manning, C.D., Ng, A.: Reasoning with neural tensor networks for knowledge base completion. In: *NIPS*. (2013) 926–934
18. Peng, B., Lu, Z., Li, H., Wong, K.F.: Towards neural network-based reasoning. *arXiv preprint arXiv:1508.05508* (2015)
19. Gliozzo, A., Biran, O., Patwardhan, S., McKeown, K.: Semantic technologies in ibm watson. *ACL 2013* (2013) 85
20. Seo, M., Hajishirzi, H., Farhadi, A., Etzioni, O., Malcolm, C.: Solving geometry problems: Combining text and diagram interpretation. In: *Proceedings of EMNLP*. (2015) 17–21
21. Weston, J., Bordes, A., Chopra, S., Rush, A.M., van Merriënboer, B., Joulin, A., Mikolov, T.: Towards ai-complete question answering: A set of prerequisite toy tasks. *arXiv preprint arXiv:1502.05698* (2015)
22. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521** (2015) 436–444
23. Serre, T., Kouh, M., Cadieu, C., Knoblich, U., Kreiman, G., Poggio, T.: A theory of object recognition: computations and circuits in the feedforward path of the ventral stream in primate visual cortex. Technical report, MIT (2005)
24. Krawczyk, D.C., McClelland, M.M., Donovan, C.M.: A hierarchy for relational reasoning in the prefrontal cortex. *Cortex* **47** (2011) 588–597
25. Shan, Q., Adams, R., Curless, B., Furukawa, Y., Seitz, S.M.: The visual turing test for scene reconstruction. In: *2013 International Conference on 3D Vision-3DV 2013, IEEE* (2013) 25–32