Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 47 (2018) 42–67

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Learning Łukasiewicz logic

Action editor: Peter Erdi

Frederik Harder [a], Tarek R. Besold [b,*]

[a] Faculty of Science, University of Amsterdam, The Netherlands

[b] Digital Media Lab, Center for Computing and Communication Technologies, University of Bremen, Bilbiothekstraße 5, 28359 Bremen, Germany

Received 20 February 2017; received in revised form 25 June 2017; accepted 26 July 2017
Available online 9 August 2017

Abstract

The integration between connectionist learning and logic-based reasoning is a longstanding foundational question in artificial intelligence, cognitive systems, and computer science in general. Research into neural-symbolic integration aims to tackle this challenge, developing approaches bridging the gap between sub-symbolic and symbolic representation and computation. In this line of work the *core method* has been suggested as a way of translating logic programs into a multilayer perceptron computing least models of the programs. In particular, a variant of the core method for three valued Łukasiewicz logic has proven to be applicable to cognitive modelling among others in the context of Byrne's suppression task. Building on the underlying formal results and the corresponding computational framework, the present article provides a modified core method suitable for the supervised learning of Łukasiewicz logic (and of a closely-related variant thereof), implements and executes the corresponding supervised learning with the backpropagation algorithm and, finally, constructs a rule extraction method in order to close the neural-symbolic cycle. The resulting system is then evaluated in several empirical test cases, and recommendations for future developments are derived.

Keywords: Neural networks; Logic programs; Neural-symbolic integration; Cognitive modelling; Reasoning

1. Introduction

Neural-symbolic integration attempts to bridge the gap between two prominent paradigms in artificial intelligence. Symbolic AI, the first of the two, encompasses explicit knowledge representation, logic programming and search-based problem solving techniques which have been responsible for many of the early successes in artificial intelligence such as game playing, automated theorem proving and natural language processing (Hsu, 2002; Robinson & Voronkov, 2001; Winograd, 1972). While the paradigm is still very much alive in expert systems managing and reasoning over vast quantities of symbolic data, it is also at times referred to as "good old-fashioned AI" or GOFAI (Haugeland, 1985), having lost some of its appeal as its limitations have become apparent. Learning from, and finding structure in sets of noisy data is something symbolic AI largely fails at. Unfortunately this means that whole classes of problems which are integral to a common conception of intelligence, such as image and voice recognition, on a general scale currently can hardly be addressed using symbolic AI.[1] Also, while (mostly non-monotonic) logic-based cognitive modelling is still being pursued with valuable results,

* Corresponding author.
E-mail addresses: frederikharder@gmail.com (F. Harder), tbesold@uni-bremen.de (T.R. Besold).

[1] Recent logic-based approaches such as, for instance, Meta-Interpretive Learning for Logical Vision Dai, Muggleton, and Zhou (2015) might in the future help to mitigate this problem, but currently have only reached proof of concept state and still have to confirm their generalizability across tasks and domains, and their scalability to real-world problem sizes.

http://dx.doi.org/10.1016/j.cogsys.2017.07.004
1389-0417/

the brittleness of the corresponding models together with their necessary restriction to high-level cognition (leaving out the bigger part of the actual representation and processing apparatus of human cognizers), are clear drawbacks when compared to connectionist or statistical approaches.

The second paradigm is that of machine learning. As the name suggests, it refers to a variety of methods for learning from data, artificial neural networks (ANN) being one prominent group of these methods. Aided by a leap in processing power and available data, machine learning has been credited with most of the more recent accomplishments in AI, from the now commonplace feat of handwriting recognition to self-driving cars and the fully autonomous learning of computer games (Berger & Rumpe, 2012; Mnih et al., 2015; Plamondon & Srihari, 2000). Promising as the paradigm may be, there are areas in which, on its own, it performs very poorly. While the learning of simple logical dependencies from data is achieved with relative ease, the process becomes increasingly difficult when higher order concepts are involved (Garcez et al., 2015). Examples for the latter impasse are numerous, including connectionist systems' problems with high-level visual analysis taking into account partial occlusion, light source identification, or shadow prediction, or with higher-level inference such as the recognition of intentions of depicted agents. Also, as knowledge is represented in connectionist systems in a distributed fashion that is hard to interpret from an outside perspective, it is usually difficult to provide background knowledge in a format which the machine learning algorithm can use, or to extract learned features from a network for instance for verification purposes. All of these are problems that often become trivial when tackled with a symbolic system.

Much stands to be gained from a unification of the two paradigms that could cancel out their respective weak spots and highlight their strengths. Neural-symbolic integration (Garcez, Broda, & Gabbay, 2002) offers some ideas in how this may be achieved, centering around the concept of the neural-symbolic cycle (see Fig. 1). The cycle contains two reasoning systems. One is symbol-based, utilizing available expert knowledge, and the other is a connectionist system or ANN, which learns from data. The challenge of interfacing these systems is twofold. Coming from the symbolic side, the first task is to find a way of translating the existing symbolic knowledge into the connectionist system, finding a representation that is appropriate for the network. Secondly, one needs to devise methods for extracting the information gained by the connectionist system through learning and convert it back into a clean symbolic format. Equipped with these processes of representation and extraction the system as a whole is capable of incorporating both background knowledge and training data as either become available.

When asked about the feasibility of integrating both paradigms, the human brain and mind serve as prime examples and proof of concept. The brain has a neural structure which operates on the basis of low-level processing of perceptual signals, but cognition also exhibits the capability to efficiently perform abstract reasoning and symbol processing; in fact, processes of the latter type are taken to provide the foundations for thinking, decision-making, and other mental activities (Fodor & Pylyshyn, 1988). It is precisely this seamless coupling between learning and reasoning which is commonly considered the basis for intelligence in humans—see also, e.g., Valiant, 2013, p. 163: "While I do not regard intelligence as a unitary phenomenon, I do believe that the problem of reasoning from learned data is a central aspect of it."—and, in close analogy, quite plausibly also for the (re-) creation of cognitive capacities up to human-level intelligence in artificial systems.

Returning to the neural-symbolic cycle discussed above, it should be made clear, that the task of constructing such a cycle rapidly increases in difficulty when raising the expressive capacities of the involved systems. There are approaches for fragments of first order logic (Bader, Hitzler, Hölldobler, & Witzel, 2007; Gust, Kühnberger, & Geibel, 2007), but most results focus on various propositional logics. Furthermore, extraction algorithms for connectionist systems tend to be intractable. So while the general method of the field can be described in a few pages, the underlying problems are hard and there is still a long way to go before neural-symbolic integration may be applied to state-of-the-art methods of either paradigm.

As one of the currently most prominent and best understood methods, Hölldobler's and Kalinke's *core method* (Hölldobler & Kalinke, 1994) has since been developed as a neural-symbolic system for, among others, propositional modal (d'Avila Garcez, Lamb, & Gabbay, 2007) and cov-

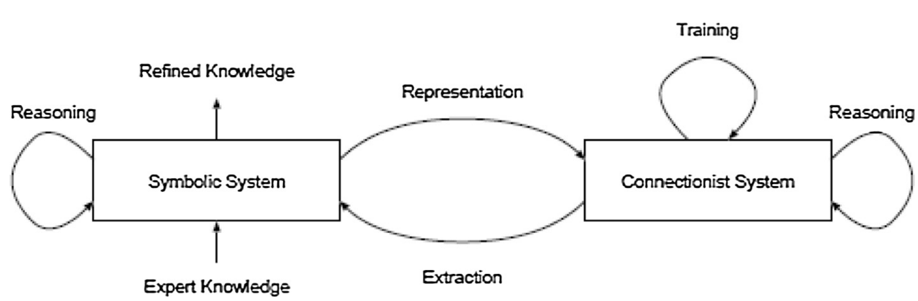

Fig. 1. A conceptual overview of the neural-symbolic cycle (as introduced in Bader and Hitzler (2005)).