Cognitive Neuro-Symbolic Reasoning Systems

Alessandro Oltramari[1]

[1]Bosch Research and Technology Center & Bosch Center for Artificial Intelligence
2555 Smallman, 15222 Pittsburgh (USA)

Abstract
Knowledge-infusion methods are key to enhance neural models and improve their performance, but they are not sufficient to enable high-level reasoning, which is typically required by tasks such as natural language understanding, activity recognition, decision making in complex scenarios. Accordingly, we propose to use a cognitive architecture as orchestrator of the integration between symbolic knowledge and machine learning.

Keywords
neuro-symbolic AI, cognitive architecture, high-level reasoning

1. Introduction

A large part of neuro-symbolic systems is based on transforming symbolic knowledge into sub-symbolic representations that are suitable for learning algorithms. For instance, Knowledge Graph Embedding (KGE) is a prominent approach to reduce knowledge graph (KG) triples to latent vectors [1]. Such transformation is instrumental to efficient computability of KG properties, as well as to application in a variety of downstream tasks. Whether the KGE process is realized by geometric, tensor or deep learning models, the purpose is to *compress* KG structures into a low-dimensional space, where symbolic statements are replaced by dense, sub-symbolic expressions. Furthermore, concatenation, non-linear mapping, attention-like mechanisms, gating mechanisms, are additional methods used to adapt knowledge structures to neural computations (e.g., [2, 3, 4]).
In this position paper we claim that, while these knowledge-infusion methods are key to enhance neural models and improve their performance, they are not sufficient to enable high-level reasoning, which is typically required by tasks such as natural language understanding, activity recognition, decision making in complex scenarios. We propose, instead, to adopt a cognitive architecture [5, 6] as a framework to combine knowledge representation and reasoning with machine learning.

In A. Martin, K. Hinkelmann, H.-G. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), Hyatt Regency, San Francisco Airport, California, USA, March 27-29, 2023.
alessandro.oltramari@us.bosch.com (A. Oltramari)
https://www.bosch.com/research/know-how/research-experts/alessandro-oltramari-ph-d/ (A. Oltramari)

Figure 1: The *Elephant in the Room*: the probability that a label assigned by an object detection system is correct increases when the context is factored in: in this example, the label 'elephant' could plausibly denote a picture of the pachyderm, but not the pachyderm itself.

2. Motivations

2.1. Lack of Context and Reasoning

Over the last decade, the integration of deep learning in computer vision systems has yielded substantial advancements. For instance, neural models can achieve good performance in object detection when training and testing domains originate from the same data distribution. However, recent work shows that minimal/regional modifications implanted in the data at test time cause significant drop in accuracy (e.g., [7, 8]). The examples documented in [8] are of particular interest, as they indicate how common sense contextualization, by means of incorporating a priori structured knowledge into deep models, can mitigate the effect of those perturbations, resulting in more robust performance [9]. In general, a visual model suitably infused with knowledge extracted from semantic resources like ConceptNet [10] can strengthen the connections holding within instances of the same conceptual domain (e.g., *couch*, *television*, *table*, *lamp* are located in living rooms) and discard out-of-context interpretations (e.g., no real *elephants* are located in living rooms, but photographs of elephant may be – figure 1 depicts such case).

When shifting to the language domain, and to tasks like automated question answering, the key role played by knowledge-based contextualization remains evident. For instance, it has been demonstrated that using KG triples to disambiguate textual elements in a sentence, and embed the corresponding concepts and relations in large language models [11], significantly improves performance (e.g., [12]). In fact, despite of the impressive results that Neural Language Modeling (NLM) is producing in Natural Language Processing (NLP) [13, 14, 15], basic reasoning capabilities are still largely missing[1]. Let's expand on this argument, and consider some examples. In ProtoQA [16], GPT-3 [17] fails to select options like 'pumpkin', 'cauliflower', 'cabbage' as top candidates, for the question 'one vegetable that is about as big as your head is?': instead, 'broccoli', 'cucumber', 'beet', 'carrot' are predicted. In this case, the different models learn some essential properties of vegetables from the training data, but do not seem to acquire the capability of comparing their size to that of other types of objects,

[1]This is also the reason why it's more appropriate to refer to these tasks as NLP, and not NLU (Natural Language Understanding), which would entail that robust and comprehensive reasoning capabilities are present.

revealing a substantial lack of *analogical reasoning* [18]. The same issues are observed when ChatGPT-3, a recent popular version of GPT-3 optimized for conversations, is considered: the main difference is that ChatGPT-3 is capable of generating plausible answers only when the question is submitted literally, but fails to do so when the question is paraphrased by using synonyms of the verbal form 'about as big as', e.g., 'about the same size', 'about the same shape', 'comparable to', etc. This hypersensitivity to surface-level linguistic features (vocabulary, syntax, etc.) – a proxy of the model's incapability to generalize over textual variations of the same conceptual content – seem to indicate that the model cannot perform the necessary (analogical) reasoning steps needed to answer to the question correctly. Along these lines, recent work [19] has shown that lack of complex inferences, role-based event prediction, and understanding the conceptual impact of negation, are some of the weaknesses diagnosed when BERT [11], one of prominent open source language models, is applied to benchmark datasets. ProtoQA again provides good examples of these deficiencies: in general, neural models struggle to correctly interpret the scope of modifiers like 'not' (*reasoning under negation*), 'often' and 'seldom' (*temporal reasoning*). Regarding the latter, in task 14 of bAbI [20], a comprehensive benchmark challenge designed by Facebook Research, NLM systems exhibit variable accuracy in grasping temporal ordering entailed by prepositions like 'before' and 'after'. Similarly, in bAbI task 17, which concerns *spatial reasoning*, NLM systems fail to infer basic positional information that require interpreting the semantics of 'to the left/right of', 'above/below', etc. If NLM systems are inaccurate when dealing with common characteristics of the physical world, their performance doesn't improve when sentiments are considered: for instance, in SocialIQA [21], given a context like 'in the school play, Robin played a hero in the struggle to death with the angry villain', models are unable to consistently select 'hopeful that Robin will succeed' over 'sorry for the villain' when required to pick the correct answer to 'how would others feel afterwards?'. It's not surprising that *reasoning about emotional reactions* represents a difficult task for pure learning systems, when we consider that such form of inference is deeply rooted in the sphere of human experiences and social life, which involves a 'layered' understanding of mental attitudes, intentions, motivations, empathy.

2.2. The Cognitive Factor

The anecdotal errors presented above are representative of a widespread phenomenon: i.e., large language models are currently not suited for human-like reasoning. But, are mainstream neuro-symbolic approaches sufficient to guarantee it? Limitations emerge in this case too: latent, sub-symbolic expressions can only augment training signals with features derived from explicit semantic content, but this infusion process doesn't carry any information about the inferential mechanisms needed to process the learned knowledge.[2] Such mechanisms are based on general logic-based reasoning, e.g. Region-Connection-Calculus for spatial reasoning [24], Allen's temporal axioms [25], and on domain/task-dependent reasoning, which is often associated with decision making. Implementing these mechanisms, and integrating them with neuro-symbolic

[2]Relevant work exists showing how deep learning models can replicate logical reasoning (e.g., [22, 23]), but it doesn't follow that any form of reasoning should be reduced to sub-symbolic learning (or at least this is an assumption only for *some* closely-paired neuro-symbolic systems).

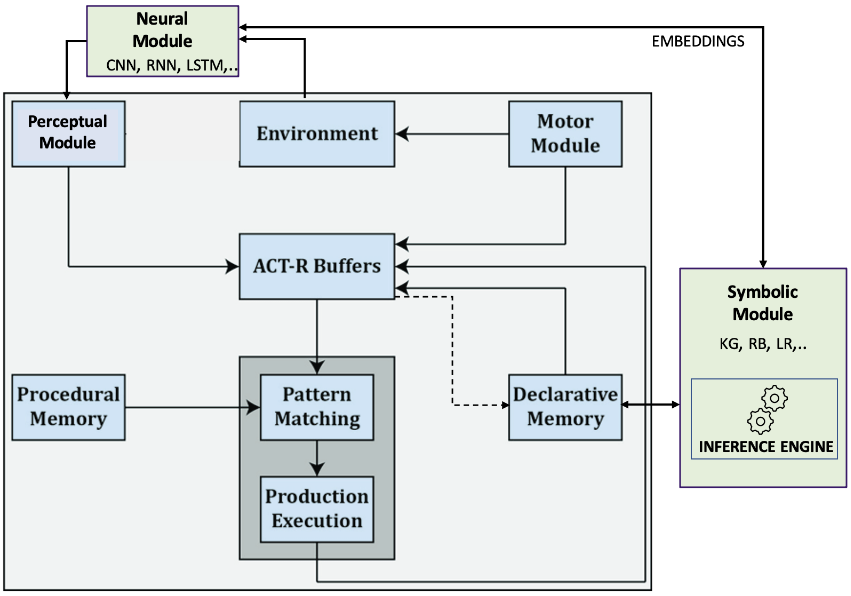

Figure 2: ACT-R integrated with independent neural and symbolic modules.

approaches, is what we advocate for in this position paper: in particular, in the next section we make the case for developing a **cognitive neuro-symbolic reasoning** framework, namely a framework where a cognitive architecture is integrated with a neural and a symbolic module.[3]

3. Method

Cognitive architectures attempt to capture at the computational level the invariant mechanisms of human cognition, including those underlying the functions of control, learning, memory, adaptivity, perception and action. ACT-R (Adaptive Control of Thought, Rational), in particular, [27], is designed as a modular framework including perceptual, motor and memory components, synchronized by a procedural module through limited capacity buffers. Over the years, ACT-R has accounted for a broad range of tasks at a high level of fidelity, reproducing aspects of complex human behavior, from everyday activities like event planning [28] and car driving [29], to highly technical tasks such as piloting an airplane [30], and monitoring a network to prevent cyber-attacks [31]. In previous work, ACT-R has been used as a component in pipelines that include either learning algorithms (e.g., biologically-inspired neural networks [32]) or external knowledge (e.g., [33, 34]): no effort exists, however, to integrate the cognitive architecture with neuro-symbolic methods and structures. We claim that such extension would be instrumental to enhance AI-systems and enable high-level reasoning.
Figure 2 outlines our proposed framework: the boxes in blue, enclosed in the grey rectangle,

[3] Our approach is complemental to the body of work on *neuro-symbolic cognitive reasoning* (see for instance [26], which investigates how neuro-symbolic approaches can be used to realize human-like cognitive reasoning.

represent the default components of ACT-R, those in green the proposed extensions. The integration occurs along three main directions:

- **knowledge** $\leftrightsquigarrow$ **memory**: the external symbolic module, which can include background/-domain knowledge graphs (KG), lexical resources (LR), rule bases (RB), and a suitable inference engine, is linked to the declarative memory. This is a two-way integration: the symbolic module can be *read* or *written* by ACT-R, where the latter operation is triggered when populating or pruning world knowledge is needed as part of task-execution.
- **neural** $\rightsquigarrow$ **perception**: the neural module, which can include convolutional, recurrent, long-short-term memory networks etc., is trained and tested with raw data processed from the environment, providing relevant patterns of information to the perceptual module. This integration bypasses the direct connection holding – in standard ACT-R – between the perceptual module and the environment.[4]
- **knowledge** $\rightsquigarrow$ **neural**: embedding mechanisms govern knowledge-infusion in the neural module, enabling knowledge-based contextualization of patterns of information distilled from the environment, and used as input for the ACT-R's perceptual module.

If the mutual connections between the two proposed modules and ACT-R provide comprehensive knowledge structures along with scalable learning functionalities, they don't – per se – bring about high-level reasoning: this capability emerges from two features of the integrated framework, namely the cognitive architecture's own procedural module and the inference engine in the external symbolic module.
The procedural module matches the content of the other module buffers and coordinates their activity using production rules, which are 'condition-action' pairs tied to the task at hand. Productions use an utility-based computation to select, from a set of task-specific plausible rules, the single rule that is executed at any point in time. For instance, when building a recommendation system to support a mechanic in troubleshooting a car engine, a relevant scenario that needs to be covered is a vehicle that doesn't start but has power; in this example, a high-utility production rule should capture the following heuristic: `if the engine holds compression well, and the fuel system is working correctly, then check the spark plugs`. The conditions in this rule clearly require empirical evidence, as it is often the case when cognitive architectures are applied to real-world problems: in our scenario, such evidence could be actually gathered by a real technician using the recommendation system in a human-machine-teaming fashion, a type of application that would fall under the 'cognitive model as oracle' paradigm [35].
The inference engine in the symbolic module is used to derive knowledge from assertions in the semantic resource of reference, a well-known feature of symbolic AI systems. What is important to stress here, is that – in our proposal - this form of logic-based reasoning has two functions: 1) providing a combination of asserted and inferred knowledge that ACT-R declarative memory can process and pass to the production system; 2) supporting knowledge-infusion into neural modules. In particular, the first functionality helps to decouple basic forms of reasoning, e.g. temporal and spatial, from cognitive assessments performed by the production system on

[4]Such connection assumes symbolic representations of visual and auditory signals being available to the architecture through pre-processing.

conditional actions. Such feature makes our proposed system efficient, as ACT-R productions are not well-suited to logical reasoning.

4. Conclusion

In the current debate on the limits of AI, the split is oftentimes between those who think that "more data" is the panacea, and those that support designing systems that integrate knowledge representation and reasoning with learning algorithms. In this position paper, which can be conceived as a product of the second category, we made the case for adopting a cognitive approach to perform that integration, inspired by the results that architectures like ACT-R have produced, over the last decades, in modeling complex human tasks and high-level reasoning. We described the main components of a *cognitive neuro-symbolic reasoning system*, and outlined their intrinsic characteristics and functionalities. At present, we are working on a first proof-of-concept of such system, focused on use cases from Industry 4.0.
To paraphrase Yoshua Bengio [36], we don't think this is the only possibility to reach human-level reasoning in AI, but through a diversity of explorations, we'll increase our chances to find the ingredients we are missing.

References

[1] Q. Wang, Z. Mao, B. Wang, L. Guo, Knowledge graph embedding: A survey of approaches and applications, IEEE Transactions on Knowledge and Data Engineering 29 (2017) 2724–2743. doi:10.1109/TKDE.2017.2754499.

[2] M. E. Peters, W. Ammar, C. Bhagavatula, R. Power, Semi-supervised sequence tagging with bidirectional language models, arXiv preprint arXiv:1705.00108 (2017).

[3] F. Strub, M. Seurin, E. Perez, H. De Vries, J. Mary, P. Preux, A. C. Pietquin, Visual reasoning with multi-hop feature modulation, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 784–800.

[4] K. Margatina, C. Baziotis, A. Potamianos, Attention-based conditioning methods for external knowledge integration, arXiv preprint arXiv:1906.03674 (2019).

[5] I. Kotseruba, J. K. Tsotsos, 40 years of cognitive architectures: core cognitive abilities and practical applications, Artificial Intelligence Review 53 (2020) 17–94.

[6] P. Langley, J. E. Laird, S. Rogers, Cognitive architectures: Research issues and challenges, Cognitive Systems Research 10 (2009) 141–160.

[7] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song, Robust physical-world attacks on deep learning visual classification, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 1625–1634.

[8] A. Rosenfeld, R. Zemel, J. K. Tsotsos, The elephant in the room, arXiv preprint arXiv:1808.03305 (2018).

[9] K. Marino, R. Salakhutdinov, A. Gupta, The more you know: Using knowledge graphs for image classification, arXiv preprint arXiv:1612.04844 (2016).

[10] R. Speer, J. Chin, C. Havasi, Conceptnet 5.5: An open multilingual graph of general knowledge, in: Thirty-first AAAI conference on artificial intelligence, 2017.
[11] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).
[12] K. Ma, F. Ilievski, J. Francis, Y. Bisk, E. Nyberg, A. Oltramari, Knowledge-driven data construction for zero-shot evaluation in commonsense question answering, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, 2021, pp. 13507–13515.
[13] K. Ma, J. Francis, Q. Lu, E. Nyberg, A. Oltramari, Towards generalizable neuro-symbolic systems for commonsense question answering, in: Proc. of the First Workshop on Commonsense Inference in Natural Language Processing, 2019, pp. 22–32.
[14] L. Bauer, M. Bansal, Identify, align, and integrate: Matching knowledge graphs to commonsense reasoning tasks, in: Proc. of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, Association for Computational Linguistics, Online, 2021, pp. 2259–2272. URL: https://www.aclweb.org/anthology/2021.eacl-main.192.
[15] V. Shwartz, P. West, R. Le Bras, C. Bhagavatula, Y. Choi, Unsupervised commonsense question answering with self-talk, in: Proc. of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Online, 2020, pp. 4615–4629. URL: https://www.aclweb.org/anthology/2020.emnlp-main.373. doi:10.18653/v1/2020.emnlp-main.373.
[16] M. Boratko, X. L. Li, R. Das, T. O'Gorman, D. Le, A. McCallum, Protoqa: A question answering dataset for prototypical common-sense reasoning, arXiv preprint arXiv:2005.00771 (2020).
[17] R. Dale, Gpt-3: What's it good for?, Natural Language Engineering 27 (2021) 113–118.
[18] A. Ushio, L. Espinosa-Anke, S. Schockaert, J. Camacho-Collados, Bert is to nlp what alexnet is to cv: Can pre-trained language models identify analogies?, arXiv preprint arXiv:2105.04949 (2021).
[19] A. Ettinger, What bert is not: Lessons from a new suite of psycholinguistic diagnostics for language models, Transactions of the Association for Computational Linguistics 8 (2020) 34–48.
[20] J. Weston, A. Bordes, S. Chopra, A. M. Rush, B. van Merriënboer, A. Joulin, T. Mikolov, Towards ai-complete question answering: A set of prerequisite toy tasks, arXiv preprint arXiv:1502.05698 (2015).
[21] M. Sap, H. Rashkin, D. Chen, R. Le Bras, Y. Choi, Social IQa: Commonsense reasoning about social interactions, in: Proc. of EMNLP-IJCNLP, 2019, pp. 4463–4473.
[22] M. Ebrahimi, A. Eberhart, P. Hitzler, On the capabilities of pointer networks for deep deductive reasoning, arXiv preprint arXiv:2106.09225 (2021).
[23] A. d. Garcez, S. Bader, H. Bowman, L. C. Lamb, L. de Penning, B. Illuminoo, H. Poon, C. Gerson Zaverucha, Neural-symbolic learning and reasoning: A survey and interpretation, Neuro-Symbolic Artificial Intelligence: The State of the Art 342 (2022) 1.
[24] A. G. Cohn, B. Bennett, J. Gooday, N. M. Gotts, Qualitative spatial representation and reasoning with the region connection calculus, geoinformatica 1 (1997) 275–316.
[25] J. F. Allen, G. Ferguson, Actions and events in interval temporal logic, Journal of logic and computation 4 (1994) 531–579.

[26] A. S. Garcez, L. C. Lamb, D. M. Gabbay, Neural-symbolic cognitive reasoning, Springer Science & Business Media, 2008.
[27] J. R. Anderson, M. Matessa, C. Lebiere, Act-r: A theory of higher level cognition and its relation to visual attention, Human–Computer Interaction 12 (1997) 439–462.
[28] S. Somers, A. Oltramari, C. Lebiere, Cognitive twin: A cognitive approach to personalized assistants., in: AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering (1), 2020.
[29] M. Cina, A. B. Rad, Categorized review of drive simulators and driver behavior analysis focusing on act-r architecture in autonomous vehicles, Sustainable Energy Technologies and Assessments 56 (2023) 103044.
[30] H. Chen, S. Liu, L. Pang, X. Wanyan, Y. Fang, Developing an improved act-r model for pilot situation awareness measurement, IEEE Access 9 (2021) 122113–122124.
[31] N. Ben-Asher, A. Oltramari, R. F. Erbacher, C. Gonzalez, Ontology-based adaptive systems of cyber defense, in: Conference on Semantic Technology for Intelligence, Defense, and Security, CEUR-WS, 2015.
[32] D. J. Jilk, C. Lebiere, R. C. O'Reilly, J. R. Anderson, Sal: An explicitly pluralistic cognitive architecture, Journal of Experimental and Theoretical Artificial Intelligence 20 (2008) 197–218.
[33] A. Oltramari, C. Lebiere, Using ontologies in a cognitive-grounded system: automatic action recognition in video surveillance, in: Proceedings of the 7th International Conference on Semantic Technology for Intelligence, Defense, and Security, Citeseer, 2012.
[34] B. Emond, Wn-lexical: An act-r module built from the wordnet lexical database, in: Proceedings of the Seventh International Conference on Cognitive Modeling, 2006, pp. 359–360.
[35] C. Lebiere, E. Cranford, M. Martin, D. Morrison, A. Stocco, Cognitive architectures and their applications, in: Proceedings of IEEE CIC, 2022.
[36] Yoshua Bengio, Deeplearning.ai, 2022. https://www.deeplearning.ai/the-batch/yoshua-bengio-wants-neural-nets-that-reason/.