Neurosymbolic Artificial Intelligence 0 (0) 1
IOS Press


# A Path Towards High-Level Reasoning Through Cognitive Neuro-Symbolic Systems

Alessandro Oltramari
*Bosch Center for Artificial Intelligence, PA, USA*
*E-mail: alessandro.oltramari@us.bosch.com*

**Abstract.** Knowledge-infusion methods are key to enhance neural models and improve their performance in downstream tasks, but they are not sufficient to enable high-level reasoning, which can be defined as the capability to generalize over knowledge acquired via (direct or mediated) experience, and to exhibit robust behavior in novel situations. Such form of reasoning is a basic skill in humans, who seamlessly use it in a broad spectrum of tasks, from language communication to decision making in complex situations. When it manifests in understanding and manipulating the everyday world of objects and their interactions, we talk about *common sense* or *commonsense reasoning*. In this paper, we propose to enable high-level reasoning at the computational level by augmenting cognitive architectures with external neuro-symbolic components. In particular, we illustrate an AI system centered on ACT-R, whose hybrid mechanisms are well-suited for integration with neuro-symbolic algorithms and resources.

Keywords: cognitive architecture, high-level reasoning, neuro-symbolic systems

## 1. Introduction

A large part of neuro-symbolic systems is based on transforming symbolic knowledge into sub-symbolic representations that are suitable for infusion in data-driven learning algorithms: Knowledge Graph Embedding (KGE), among the others, 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: for instance, in [2] the authors leverage KGE methods to label unseen entities in autonomous driving datasets. 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 with dense, sub-symbolic expressions. Concatenation, non-linear mapping, attention-like mechanisms, gating mechanisms, are further methods to adapt knowledge structures to neural computations (e.g., [3–5]).
While knowledge-infusion is instrumental to augment the performance of neural models, it is not sufficient to enable *high-level reasoning*, which is typically required by complex tasks such as natural language understanding, activity recognition, decision making in complex scenarios: latent, sub-symbolic expressions can only augment training signals with features derived from explicit semantic content, but this infusion process does neither carry any information about the reasoning mechanisms needed to process the learned knowledge, nor instruct the neural models on how those should unfold.
But, what do we mean with *high-level reasoning* and why is it important to endow AI with such feature?

## 2. Problem statement

We can define *high-level reasoning* as the capability to generalize over knowledge acquired via (direct or mediated) experience, and to exhibit robust behavior in novel situations. This working definition is loosely inspired to Kahneman's SYSTEM 2 mode of thought [6]. When *high-level reasoning* manifests in understanding and manipulating the everyday world of objects and their interactions, we talk about *common sense* or *commonsense reasoning*. State-of-the-art AI systems don't possess such capability: for instance, Large Language Models (LLMs) have recently become popular by demonstrating remarkable fluency in conversing with humans, but they still make trivial mistakes when probed for commonsense competence (see section 3); on a different level, one of the motivations why the promise of autonomous cars hasn't panned out yet concerns performance degradation outside training data, which prevents self-driving vehicles to safely adapt to unseen scenarios. Humans, at the opposite, are very good at generalizing from a few examples, and at filling the gaps in experience with reasoning: for instance, when asked about what happens after a bottle of red wine is forcefully thrown against a concrete wall, even small children can answer with utmost certainty that the bottle will shatter and the wall will be wet and red-stained – they can also easily infer that the impact between *any* fragile material and *any* hard surface *typically* ends with the former being substantially altered, if not destroyed; moreover, student drivers only need limited training to learn how to safely maneuver a car, adapting their knowledge and skills to novel situations. Compared to current AI systems based on GPU accelerated computing, human reasoning capabilities are impressive, even more so when we factor in what Herbert Simon used to call 'bounded rationality' [7], i.e., the notion that human cognition operates with limited knowledge and is subject to time constraints – widely considered a heritage of evolution (e.g., see [8]).
In this context, we firmly believe that adopting a cognitive stance toward designing AI systems [9] is key to enable high-level reasoning capabilities at the computational level: accordingly, we propose to complement cognitive architectures [10, 11] with neuro-symbolic methods. In this paper we illustrate the blueprints of a *cognitive neuro-symbolic reasoning* system centered on the ACT-R[1] cognitive architecture [12], whose hybrid (symbolic and sub-symbolic) mechanisms are well-suited for integration with neuro-symbolic algorithms and resources (see section 4). Note that the proposed approach is applicable to any cognitive architecture whose properties are equivalent to or compatible with ACT-R, such as SOAR [13] and SIGMA [14]: in fact, these three architectures have been grouped into the so-called 'Standard Model of the Mind' [15]. The Standard Model of the Mind, an idea that has its roots in physics, doesn't prescribe any strict requirement on how to implement cognitively-inspired AI systems, but rather aims to play the role of a conceptual framework of reference for developing them.

## 3. Motivations

Over the last decade, the integration of deep learning in computer vision systems has yielded tremendous advancements. For instance, neural models can achieve high accuracy 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., [16, 17]). The examples documented in [17] are of particular interest, as they indicate how commonsense contextualization, by means of incorporating a priori structured knowledge into deep networks, can mitigate the effect of those perturbations, resulting in more robust performance [18]. In general, a visual model suitably infused with knowledge extracted from semantic resources like CONCEPTNET [19] 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 natural language, and to tasks like automated question answering, the key role played by knowledge-based contextualization for neural language models stands evident.[2] For instance, it has been demon-

[1]Abbreviation of 'Adaptive Control of Thought, Rational'.

[2]We use 'language model', 'neural language model' and 'large language model' as interchangeable terms, in virtue of their common neural architecture, based on multi-headed self-attention mechanisms [20]; however, their computational power significantly differs as function of the size of the specific implementations (e.g., BERT has 6 blocks with 12 heads, GPT-3 has 24 blocks and 48 heads), and of the training datasets (CHAT-GPT has been trained on a massive corpus – 570 GB – of text data).

Fig. 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.

strated that using KG triples to disambiguate textual elements in a sentence, and embed the corresponding concepts and relations in neural language models [21], significantly improves performance (e.g., [22]). In fact, despite of the impressive results that LLMs are producing in Natural Language Processing (NLP) [23–25], basic reasoning capabilities are still largely missing. 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 [26]. Let's expand on this argument, and consider some examples.

In ProtoQA [27], GPT-3 [28] 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, revealing a substantial lack of *analogical reasoning* [29]. The same issues are observed when CHATGPT, a recent popular version of GPT-3 optimized for conversations, is considered: the main difference is that CHATGPT is capable of generating plausible answers when the question is submitted literally, but often fails to do so when the verbal expression 'about as big as' is paraphrased with alternative forms like '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 content – seem to indicate that the model cannot perform the necessary (analogical) reasoning steps needed to correctly answer to the question. Along these lines, recent work [30] 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 [21], 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 [31], a comprehensive benchmark challenge designed by Facebook Research, neural language systems exhibit variable accuracy in grasping temporal ordering entailed by prepositions like 'before' and 'after'. Similarly, in bAbI task 17, which concerns *spatial reasoning*, LLM-based systems fail to infer basic positional information that require interpreting the semantics of 'to the left/right of', 'above/below', etc. If such 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 [32], 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, emotions, and of the events that trigger them.

The qualitative analysis presented above suggests that neural models cannot perform high-level reasoning. Are neuro-symbolic approaches sufficient to overcome these limitation? Latent expressions can augment training sig-

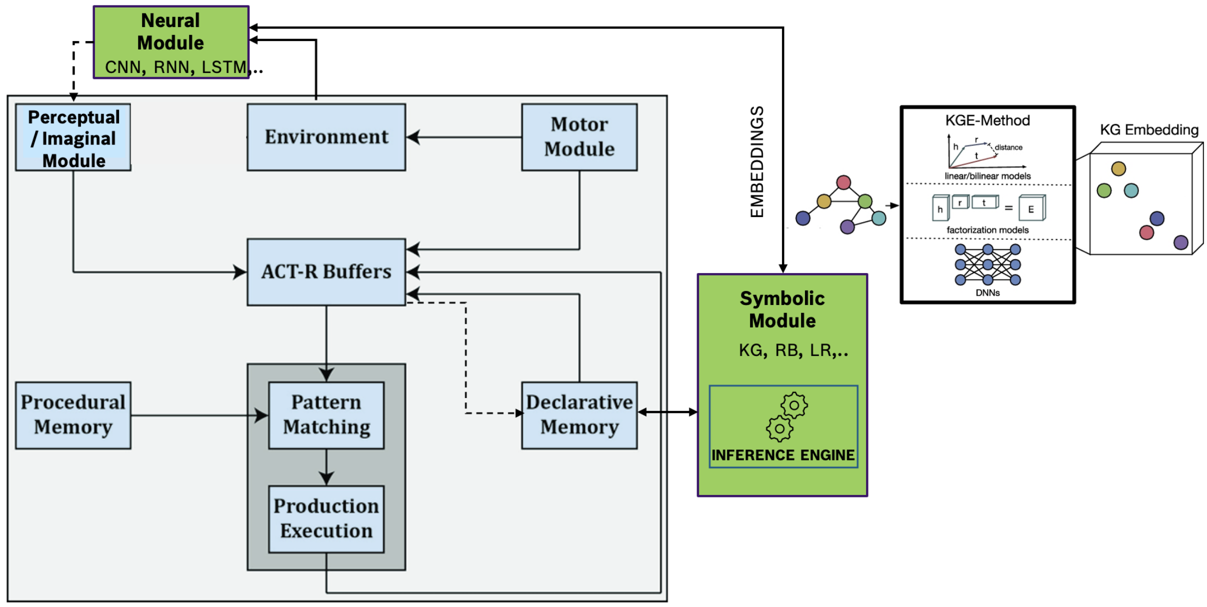


Fig. 2. ACT-R integrated with independent neural and symbolic modules.

nals with sub-symbolic features derived from explicit semantic content, but knowledge infusion *per se* doesn't determine how inference processes are conducted. Relevant work in this space shows how deep neural models can replicate logical reasoning (e.g., [33, 34]), but it doesn't follow that any form of logical reasoning that is provably reducible to learning, should also be systematically reduced to it (this would be a requirement only for tightly-coupled neuro-symbolic systems (e.g., [35, 36]). Combining semantic content with logic-based reasoning – e.g. Region-Connection-Calculus [37] for spatial reasoning, Allen's axioms for temporal reasoning [38] – and implementing these mechanisms by integrating cognitive architectures with neuro-symbolic approaches, is what we advocate for in this paper: in particular, in the next section we make the case for developing a cognitive neuro-symbolic reasoning framework, where the ACT−R architecture is extended with neuro-symbolic modules.

## 4. 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 [12], in particular, is designed as a hybrid 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 [39] and car driving [40], to highly technical tasks such as piloting an airplane [41], and monitoring a network to prevent cyber-attacks [42]. ACT-R has been used as a component in pipelines that include either learning algorithms (e.g., biologically-inspired neural networks [43]) or external semantic resources (e.g., [44, 45]): however, to the best of our knowledge, no effort exists to integrate the ACT-R cognitive architecture with neuro-symbolic components. We claim that such extension is instrumental to enhance AI-systems by enabling high-level reasoning.
Figure 2 provides a compact visualization of our proposed framework: the boxes in blue, enclosed in the grey rectangle, represent the default components of ACT−R, those in green the neuro-symbolic extensions. The integration would occur along three main directions:

– **knowledge ⇜⇝ 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 ⇝ perception**: the neural module, which can include convolutional, recurrent, long-short-term memory networks, generative models, etc., is trained or fine-tuned with data processed from the environment, providing relevant patterns of information to the perceptual or imaginal module. This integration bypasses the direct connection holding – in standard ACT-R – between the perceptual module and the environment.[3]
– **knowledge ⇝ neural**: adequately-selected embedding mechanisms govern knowledge-infusion in the neural module, enabling knowledge-based contextualization of patterns of information distilled from the environment, which are subsequently channeled into ACT-R buffers.

If the mutual connections between the two intertwined neuro-symbolic modules and ACT-R can be used to combine rich semantic contents with scalable learning functionalities, they don't *per se* bring about high-level reasoning: this capability also emerges from two features of the integrated framework, namely the cognitive architecture's own procedural module and a proper 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 situation 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 the spark plugs should be checked*. The variables in these rule conditions would need to be filled with actual empirical observations and measurements, as it is often the case when cognitive architectures are applied in real-world scenarios: in our example, such evidence could be actually gathered by a real technician using the recommendation system in a human-machine-teaming fashion, a type of approach that falls under the 'cognitive model as oracle' paradigm [46].
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 would realize two functions: 1) provide a combination of asserted and inferred knowledge that ACT-R declarative memory can process and pass to the production system; 2) support knowledge-infusion into neural modules. The first functionality would help to decouple basic forms of reasoning, e.g. temporal and spatial, from cognitive assessments performed by the production system on conditional actions. Such feature makes our proposed system efficient, as ACT-R productions are not well-suited to logical reasoning. The second functionality would allow pre-training or fine-tuning of a neural-model on asserted and inferred knowledge: this can provide ACT-R perceptual model with more informative patterns than just those obtained by processing raw data.

## 5. Related Work

The importance of integrating cognitive architectures into data-driven AI systems has been recently acknowledged by one of the key figures in deep learning, Yann LeCun: in a position paper published in 2022 [47], the scholar outlined a biologically-inspired cognitive architecture, where a so-called *configurator* orchestrates information provided by different modules, such as the *perception module* and the *world model module*, which replicate the functions emerging from prefrontal-cortical processes. Furthermore, a *motivation model* – designed to mimic the role of the amygdala in producing basic emotional states like pain and pleasure – is used to compute intrinsic costs associated with current and future actions, a mechanism that is instrumental to inform predictive capabilities. It's relevant to point out, here, that there is extensive research on mappings ACT-R modules to brain areas/processes,

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

e.g., [48], and that an established scientific community has been working on biologically-inspired cognitive architectures since the early 2000's (the upcoming BICA international conference has reached its 14th installation[4]).
In line with LeCun's effort to include computational models of cognition into current mainstream AI, a recent blog [49] provides an interesting overview of how LLMs could be used as main controllers for autonomous agents. It goes beyond the scope of our contribution to review the papers mentioned in the blog, but it's beneficial to position ourselves with respect to some of the most interesting approaches covered by it (see references [51-55]). There are two main reasons to place a cognitive architecture and not a LLM at the center of an AI framework: 1) the inner functioning of the former is accessible, whereas the latter is considered a 'black-box' [50]; 2) the former is designed to replicate the invariant mechanisms of human cognition, the latter – at best – can be used to ascribe cognitive properties to linguistic behavior. For example, in [51] the authors leveraged chain-of-thought prompting with PaLM 540B [52] for task-decomposition: despite of their reported success, using prompting to generate fine-grained reasoning steps does not always yield consistent results, even within different versions of the same model, as shown by [53] for GPT-4 [54]. The same paper also indicates that, when reasoning steps are correctly reproduced, they don't always match with the correct solution/answer to a problem/question. In [55], the authors designed *believable agents* for a sandbox environment inspired by the video-game 'The Sims'[5], using GPT-4. We found an interesting remark in the paper: agents built using cognitive architectures are typically designed for closed-world contexts, since procedural knowledge is typically hand-crafted; as a consequence, cognitive architectures would not have the same degree of flexibility that modern generative models demonstrate. However, this is a partial account of the state of the art: production compilation, ACT-R's rule learning mechanism, can be enabled to learn new, task-specific production rules that directly implement the relevant action for a particular state [56]; moreover, to assess what is relevant to act upon, given some input from the environment, mechanisms like instance-based learning (a type of reinforcement learning [57]) have been plugged into ACT-R agents [58]. Besides, one may also question the very nature of the alleged flexibility of generative model's based agents: in fact, such flexibility does not truly reflect the capability of an agent to adapt to changes in the world, but rather to variances that occur in the training data, which are *de facto* biased interpretations of reality. Incidentally, this lack of 'grounding' is at the origin of LLMs' widely-documented hallucination problem (for an introduction to this phenomenon, see [59]).
It's worth making a final consideration here: the framework introduced in section 4 is complemental to the body of work that investigates how neuro-symbolic systems can be leveraged to realize human-like cognitive reasoning (see for instance [60]): in our proposal, ACT-R is interfaced with neuro-symbolic components, whereas – in the approaches reviewed by Garcez et al. – neuro-symbolic frameworks are used to solve cognitive tasks. The difference lies on considering cognitive processes either as primitive or as emergent properties.

## 6. Conclusion

In the current debate on the limits of deep neural networks, the split is oftentimes between those who think that 'more/better data' is the panacea, and those who support designing systems that integrate learning approaches with other processing elements, such as knowledge representation and reasoning, statistical algorithms, human-in-the-loop methods. In this position paper, which belongs to 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 replicating complex human tasks at the machine level. We described the main components of a *cognitive neuro-symbolic reasoning system*, and outlined their respective functionalities.
Far from being exhaustive, this paper aims to open the path to what, we hope, will be a florid research thrust in the forthcoming issues of the Neuro-Symbolic AI journal. In the end, to paraphrase Yoshua Bengio [61] – like LeCun, a key figure of the new wave in AI – we don't assume or prove that using cognitive architectures is the only possibility to equip machines with high-level, human-like reasoning: however, through a diversity of scientific explorations, we'll increase our chances to find the ingredients we are missing.

[4]See: https://bica2023.org/cfp/
[5]https://www.ea.com/games/the-sims

## References


[1] Q. Wang, Z. Mao, B. Wang and L. Guo, Knowledge Graph Embedding: A Survey of Approaches and Applications, *IEEE Transactions on Knowledge and Data Engineering* **29**(12) (2017), 2724–2743. doi:10.1109/TKDE.2017.2754499.

[2] R. Wickramarachchi, C. Henson and A. Sheth, CLUE-AD: A Context-Based Method for Labeling Unobserved Entities in Autonomous Driving Data, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37, 2023, pp. 16491–16493.

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

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

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

[6] D. Kahneman, *Thinking, fast and slow*, macmillan, 2011.

[7] H.A. Simon, A behavioral model of rational choice, *The quarterly journal of economics* (1955), 99–118.

[8] L.R. Santos and A.G. Rosati, The evolutionary roots of human decision making, *Annual review of psychology* **66** (2015), 321–347.

[9] A. Lieto, *Cognitive design for artificial minds*, Routledge, 2021.

[10] I. Kotseruba and J.K. Tsotsos, 40 years of cognitive architectures: core cognitive abilities and practical applications, *Artificial Intelligence Review* **53**(1) (2020), 17–94.

[11] P. Langley, J.E. Laird and S. Rogers, Cognitive architectures: Research issues and challenges, *Cognitive Systems Research* **10**(2) (2009), 141–160.

[12] J.R. Anderson, ACT: A simple theory of complex cognition., *American psychologist* **51**(4) (1996), 355.

[13] J.E. Laird, *The Soar cognitive architecture*, MIT press, 2019.

[14] P.S. Rosenbloom, A. Demski and V. Ustun, The Sigma cognitive architecture and system: Towards functionally elegant grand unification, *Journal of Artificial General Intelligence* **7**(1) (2016), 1.

[15] J.E. Laird, C. Lebiere and P.S. Rosenbloom, A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics, *Ai Magazine* **38**(4) (2017), 13–26.

[16] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno and 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.

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

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

[19] R. Speer, J. Chin and C. Havasi, Conceptnet 5.5: An open multilingual graph of general knowledge, in: *Thirty-first AAAI conference on artificial intelligence*, 2017.

[20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser and I. Polosukhin, Attention is all you need, *Advances in neural information processing systems* **30** (2017).

[21] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, *arXiv preprint arXiv:1810.04805* (2018).

[22] K. Ma, F. Ilievski, J. Francis, Y. Bisk, E. Nyberg and A. Oltramari, Knowledge-driven data construction for zero-shot evaluation in commonsense question answering, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, 2021, pp. 13507–13515.

[23] K. Ma, J. Francis, Q. Lu, E. Nyberg and 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.

[24] L. Bauer and 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. https://www.aclweb.org/anthology/2021.eacl-main.192.

[25] V. Shwartz, P. West, R. Le Bras, C. Bhagavatula and 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. doi:10.18653/v1/2020.emnlp-main.373. https://www.aclweb.org/anthology/2020.emnlp-main.373.

[26] M. McShane, Natural language understanding (NLU, not NLP) in cognitive systems, *AI Magazine* **38**(4) (2017), 43–56.

[27] M. Boratko, X.L. Li, R. Das, T. O'Gorman, D. Le and A. McCallum, ProtoQA: A Question Answering Dataset for Prototypical Common-Sense Reasoning, *arXiv preprint arXiv:2005.00771* (2020).

[28] R. Dale, GPT-3: What's it good for?, *Natural Language Engineering* **27**(1) (2021), 113–118.

[29] A. Ushio, L. Espinosa-Anke, S. Schockaert and 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).

[30] 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.

[31] J. Weston, A. Bordes, S. Chopra, A.M. Rush, B. van Merriënboer, A. Joulin and T. Mikolov, Towards ai-complete question answering: A set of prerequisite toy tasks, *arXiv preprint arXiv:1502.05698* (2015).

[32] M. Sap, H. Rashkin, D. Chen, R. Le Bras and Y. Choi, Social IQa: Commonsense Reasoning about Social Interactions, in: *Proc. of EMNLP-IJCNLP*, 2019, pp. 4463–4473.

[33] M. Ebrahimi, A. Eberhart and P. Hitzler, On the Capabilities of Pointer Networks for Deep Deductive Reasoning, *arXiv preprint arXiv:2106.09225* (2021).
[34] A.d. Garcez, S. Bader, H. Bowman, L.C. Lamb, L. de Penning, B. Illuminoo, H. Poon and C. Gerson Zaverucha, Neural-symbolic learning and reasoning: A survey and interpretation, *Neuro-Symbolic Artificial Intelligence: The State of the Art* **342** (2022), 1.
[35] H. Kautz, The third ai summer: Aaai robert s. engelmore memorial lecture, *AI Magazine* **43**(1) (2022), 105–125.
[36] A.d. Garcez and L.C. Lamb, Neurosymbolic AI: The 3 rd wave, *Artificial Intelligence Review* (2023), 1–20.
[37] A.G. Cohn, B. Bennett, J. Gooday and N.M. Gotts, Qualitative spatial representation and reasoning with the region connection calculus, *geoinformatica* **1** (1997), 275–316.
[38] J.F. Allen and G. Ferguson, Actions and events in interval temporal logic, *Journal of logic and computation* **4**(5) (1994), 531–579.
[39] S. Somers, A. Oltramari and C. Lebiere, Cognitive Twin: A Cognitive Approach to Personalized Assistants., in: *AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering (1)*, 2020.
[40] M. Cina and 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.
[41] H. Chen, S. Liu, L. Pang, X. Wanyan and Y. Fang, Developing an improved ACT-R model for pilot situation awareness measurement, *IEEE Access* **9** (2021), 122113–122124.
[42] N. Ben-Asher, A. Oltramari, R.F. Erbacher and C. Gonzalez, Ontology-based adaptive systems of cyber defense, in: *Conference on Semantic Technology for Intelligence, Defense, and Security*, CEUR-WS, 2015.
[43] D.J. Jilk, C. Lebiere, R.C. O'Reilly and J.R. Anderson, SAL: An explicitly pluralistic cognitive architecture, *Journal of Experimental and Theoretical Artificial Intelligence* **20**(3) (2008), 197–218.
[44] A. Oltramari and 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.
[45] 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.
[46] C. Lebiere, E. Cranford, M. Martin, D. Morrison and A. Stocco, Cognitive architectures and their applications, in: *Proceedings of IEEE CIC*, 2022.
[47] Y. LeCun, A path towards autonomous machine intelligence version 0.9. 2, 2022-06-27, *Open Review* **62** (2022).
[48] J.P. Borst, M. Nijboer, N.A. Taatgen, H. van Rijn and J.R. Anderson, Using data-driven model-brain mappings to constrain formal models of cognition, *PLoS One* **10**(3) (2015), e0119673.
[49] LLM-powered Autonomous Agents.
[50] D. Castelvecchi, Can we open the black box of AI?, *Nature News* **538**(7623) (2016), 20.
[51] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q.V. Le, D. Zhou et al., Chain-of-thought prompting elicits reasoning in large language models, *Advances in Neural Information Processing Systems* **35** (2022), 24824–24837.
[52] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H.W. Chung, C. Sutton, S. Gehrmann et al., Palm: Scaling language modeling with pathways, *arXiv preprint arXiv:2204.02311* (2022).
[53] L. Chen, M. Zaharia and J. Zou, How is ChatGPT's behavior changing over time?, *arXiv preprint arXiv:2307.09009* (2023).
[54] R. OpenAI, GPT-4 technical report, *arXiv* (2023), 2303–08774.
[55] J.S. Park, J.C. O'Brien, C.J. Cai, M.R. Morris, P. Liang and M.S. Bernstein, Generative agents: Interactive simulacra of human behavior, *arXiv preprint arXiv:2304.03442* (2023).
[56] N.A. Taatgen, D. Huss and J.R. Anderson, How cognitive models can inform the design of instructions, in: *Proceedings of the seventh international conference on cognitive modeling*, Citeseer, 2006, pp. 304–309.
[57] R.S. Sutton and A.G. Barto, *Reinforcement learning: An introduction*, MIT press, 2018.
[58] C. Gonzalez, J.F. Lerch and C. Lebiere, Instance-based learning in dynamic decision making, *Cognitive Science* **27**(4) (2003), 591–635.
[59] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y.J. Bang, A. Madotto and P. Fung, Survey of hallucination in natural language generation, *ACM Computing Surveys* **55**(12) (2023), 1–38.
[60] A.S. Garcez, L.C. Lamb and D.M. Gabbay, *Neural-symbolic cognitive reasoning*, Springer Science & Business Media, 2008.
[61] Yoshua Bengio, DeepLearning.AI, 2022, https://www.deeplearning.ai/the-batch/yoshua-bengio-wants-neural-nets-that-reason/.