



Cognitive Reasoning: A Personal View

Ulrich Furbach¹ · Steffen Hölldobler^{2,3} · Marco Ragni⁴ · Claudia Schon¹ · Frieder Stolzenburg⁵ 

Received: 29 May 2019 / Accepted: 24 June 2019 / Published online: 29 June 2019
© Gesellschaft für Informatik e.V. and Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

The adjective *cognitive* especially in conjunction with the word *computing* seems to be a trendy buzzword in the artificial intelligence community and beyond nowadays. However, the term is often used without explicit definition. Therefore we start with a brief review of the notion and define what we mean by *cognitive reasoning*. It shall refer to modeling the human ability to draw meaningful conclusions despite incomplete and inconsistent knowledge involving among others the representation of knowledge where all processes from the acquisition and update of knowledge to the derivation of conclusions must be implementable and executable on appropriate hardware. We briefly introduce relevant approaches and methods from cognitive modeling, commonsense reasoning, and subsymbolic approaches. Furthermore, challenges and important research questions are stated, e.g., developing a computational model that can compete with a (human) reasoner on problems that require common sense.

Keywords Reasoning · Cognitive Reasoning · Commonsense Reasoning

The authors are mentioned in alphabetical order.

✉ Frieder Stolzenburg
fstolzenburg@hs-harz.de

Ulrich Furbach
uli@uni-koblenz.de

Steffen Hölldobler
sh@iccl.tu-dresden.de

Marco Ragni
ragni@informatik.uni-freiburg.de

Claudia Schon
schon@uni-koblenz.de

¹ Universität Koblenz-Landau, Universitätsstr. 1,
56070 Koblenz, Germany

² Technische Universität Dresden, 01062 Dresden, Germany

³ North-Caucasus Federal University, Stavropol,
Russian Federation

⁴ Cognitive Computation Lab, University of Freiburg,
Georges-Köhler-Allee 52, 79110 Freiburg, Germany

⁵ Harz University of Applied Sciences, Friedrichstraße 57-59,
38855 Wernigerode, Germany

1 Introduction to Cognitive Reasoning

In this paper we give a personal view to the relatively new research field of *cognitive reasoning*. Although the notion has been used several times in the scientific community, surprisingly there is apparently no generally agreed definition of the term and field. At least no rigorously formal definition of the compound term exists. We found books such as [4] and [30] and research papers like [57], none of which contains a concise definition. There are lectures like *Humanoid Cognitive Reasoning* by Gordon Cheng at TU München in the summer term 2019,¹ workshops like the one on *Formal and Cognitive Reasoning*,² or the project *CoRg—Cognitive Reasoning*,³ but again, the corresponding webpages do not provide an explicit definition of the term.

On the other hand, we found many nouns to which the adjective *cognitive* was attached like, e.g., cognitive computing, cognitive robotics research, cognitive intelligence, cognitive semantics, cognitive programs, cognitive programming, cognitive architectures, cognitive assistant, cognitive compatibility, or cognitive models. Sadly, no concise definitions seem to exist here as well.

¹ <http://ics.ie.tum.de>, accessed: 13-May-2019.

² <http://www.fernuni-hagen.de/wbs/dkbbkik2019.html>, accessed: 13-May-2019.

³ <http://gepris.dfg.de>, accessed: 13-May-2019.

1.1 From Cognitive Computing ...

From the just mentioned terms, maybe the term *cognitive computing* has gained the most attention in recent years [69]. It was established by IBM, intended to supersede more or less the notions artificial intelligence and machine learning. It became a trendy buzzword after IBM's Watson project which was able to beat the best human players in the quiz game "Jeopardy" [28]. A definition is provided by IBM: "Cognitive Computing are systems that learn at scale, reason with purpose and interact with humans naturally. It is a mixture of computer science and cognitive science—that is, the understanding of the human brain and how it works. By means of self-teaching algorithms that use data mining, visual recognition, and natural language processing, the computer is able to solve problems and thereby optimize human processes."⁴ It appears to us very likely that the marketing people from IBM led the way with this formulation.

From a more technical perspective, cognitive computing can be characterized as a method at the intersection of artificial intelligence and cognitive science. It is inspired by cognitive processes in the human mind and brain and has been developed to foster an easier and more natural interaction level for the human user. What are the characteristics of human information processing?

- *Multiple knowledge formats*: There is a modality-specific type of information, e.g., visual information is maintained and processed differently from aural processes. This is one basic principle that is realized in different cognitive architectures such as ACT-R [2, 3].
- *Multiple reasoning mechanisms*: It is commonly accepted that a human reasoner can employ a heuristic process (often called System 1) and an analytic process (often called System 2) [26, 27, 47] (see also Sect. 2.1).
- *Cooperation of modules*: Different types of information are processed in the human mind. Regions in the human brain can be typically associated with task-specific modules. Those can form bottlenecks.
- *Time-critical*: Decisions and inferences are often drawn under time constraints.

The term cognitive computing has also been used to refer to new hardware and/or software that mimics the functioning of the human brain.⁵ In this context, neuromorphic hardware [45, 56] is considered which is the result of the exploration of unconventional physical substrates and nonlinear

phenomena.⁶ The goal in this context is to simulate the brain, which usually means building artificial neural networks, in particular recurrent networks with high connectivity. This leads to the field of reservoir computing [52], where a reservoir denotes a set of hidden neurons with more or less random dynamics.

1.2 ...to Cognitive Reasoning

But now: What is *cognitive reasoning*?

According to [49] (see also Wikipedia⁷) "*reason* is the capacity of consciously making sense of things, establishing and verifying facts, applying logic, and changing or justifying practices, institutions, and beliefs based on new or existing information".

Reasoning is associated with thinking, cognition, and intellect. It may be subdivided into forms of classical logical reasoning: deductive reasoning, inductive reasoning, abductive reasoning; non-classical logical reasoning: analogical reasoning, commonsense reasoning, defeasible reasoning, probabilistic reasoning; and other modes of reasoning such as counterfactual reasoning, intuitive reasoning, and verbal reasoning. Reasoning is based on knowledge and beliefs, but knowledge may be incomplete and beliefs may be incorrect; inconsistencies may arise. Knowledge must be acquired or learned, beliefs must be revised and updated, and preferences, likes and dislikes must be taken into account.

The Oxford Dictionary defines *cognition* as the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses. Cognition refers to functions of the (human) brain and other biological processes. However, cognitive theories need not necessarily involve the brain or biological processes, but may describe their behavior in terms of information flow or function.

In the light of the above notions we define *cognitive reasoning* as modeling the human ability to draw meaningful conclusions despite incomplete and inconsistent knowledge. The modeling involves the (symbolic or subsymbolic) representation of knowledge, the acquisition and updating of the knowledge, as well as all computational processes of deriving conclusions from the given knowledge. All processes from the acquisition and updating of knowledge to the derivation of conclusions must be implementable and executable on appropriate hardware (cf. [54]).

Note that the notion of modeling, especially in cognitive science, has a long tradition. As pointed out by McClelland [55] "...[models] are explorations of ideas about the nature of cognitive processes". With this rather broad understanding of the notion of modeling it is not mandatory to build

⁴ <http://www.ibm.com/blogs/nordic-msp/artificial-intelligence-machine-learning-cognitive-computing/>, accessed: 13-May-2019.

⁵ http://en.wikipedia.org/wiki/Cognitive_computing, accessed: 13-May-2019.

⁶ <http://www.cognitive-comp.org/>, accessed: 13-May-2019.

⁷ <http://en.wikipedia.org/wiki/Reason>, accessed: 13-May-2019.

a theory about how humans are generating a certain cognitive ability. We also include approaches modeling just the ability itself.

2 Approaches and Methods

This section presents various methods for the design of cognitive reasoning systems.

2.1 Cognitive Modeling

Cognitive modeling is an interdisciplinary field using methods from computer science, artificial intelligence and cognitive psychology to develop computational models for human cognitive processes. For many years, a core goal has been to develop a unified theory of cognition [2, 3], i.e., a theory of cognition that comprises a general data structure and realizes the idea of a general problem solver [64]. The theory should be *computational* in that solutions to human reasoning tasks are computed, *integrated* in that different human reasoning tasks can be modeled by the theory without changing the theory, and *adequate* in that it correspond as closely as possible to what humans are doing when solving a reasoning task. Models of cognitive reasoning can comprise several levels: these can be heuristic [34, 35], probabilistic [65], logically rule-based inspired (including non-monotonic and non-classical logics) [14, 76], or model-based (e.g., [14, 46, 83]).

Today such a system would be of great interest due to its application at the interface between human and artificial agents. However, even after many years, the current research appears to be only in its initial stages. In a meta-study on human syllogistic reasoning, Khemlani and Johnson-Laird have compared 12 theories and have concluded that “the existence of 12 theories of any scientific domain is a small disaster ...if psychologists could agree on an adequate theory of syllogistic reasoning, then progress towards a more general theory of reasoning would seem to be feasible ...if researchers were unable to account for syllogistic reasoning, then they would have little hope of making sense of reasoning in general” [48].

This is not the place to discuss the major cognitive theories in detail and we refer the interested reader to [48] for an overview. But we would like to briefly introduce a novel cognitive theory which has recently outperformed the 12 theories mentioned above on human syllogistic reasoning [24, 66]: the *weak completion semantics*. The weak completion semantics is an integrated and computational cognitive theory, based on ideas initially proposed by Stenning and van Lambalgen [82]. It is mathematically sound [43], has been applied to various human reasoning tasks like the suppression task [20], the selection task [21], the belief-bias

effect [70], and ethical decision tasks [40], and can be implemented in a connectionist setting [22] (see also Sect. 2.3).

Given a human reasoning task, the first step within the weak completion semantics is to construct a logic program representing the task. The construction of these programs is based on several principles, some of which are well-established like using licenses for inferences [82], existential import, or Gricean implicature, whereas others are novel like unknown generalization. If interpreted under the three-valued logic of [53], the programs have a unique supported model which can be computed by iterating the semantic operator introduced in [82]. Reasoning is performed and answers are computed with respect to these models. Skeptical abduction is added if some observations in the given human reasoning task can not be explained otherwise. For an introduction into the weak completion semantics see [23].

Cognitive modeling can include specific assumptions about the architecture of human cognition. This is often realized by a symbolic or hybrid cognitive architecture that consists of modules and specific assumptions about the way information is processed, e.g., ACT-R [2, 3] or SOAR [63]. Some cognitive models are implemented in such architecture frameworks, using the specifics of the architecture’s data structure (e.g., chunks or cognitive bottlenecks) especially, if these models aim to explain or predict the outcome of human cognitive processes.

In the last years generally (at least) two modes of cognitive reasoning have been assumed: a fast, heuristic, and rather parallel process, often called *System 1* and a slow, analytic and sequential process often called *System 2* [26, 27, 47]. Recently, this analogy has been exploited for the connection between machine learning being a System 1 approach and (symbolic) artificial intelligence being a System 2 approach. Humans are quite flexible in using any of these approaches and so there is a large variety of models. A model is not only considered to be “good” if it predicts what a human reasoner does, but more is required: (i) the models need to make precise and testable hypotheses about cognitive processes in general and (ii) explain the process outcomes of the models (at the level of strong machine learning [61] including intermediate steps [80]).

2.2 Commonsense Reasoning

In recent years, numerous benchmarks for commonsense reasoning have been presented. Commonsense reasoning in general can be understood as the sort of everyday reasoning humans typically perform about the world [60]. The areas of the benchmarks are broad and include textual entailment [11], human causal reasoning [77], completing narratives [59], and linguistic tasks such as word sense disambiguation in particularly difficult cases [51], to name just a few. Although the focus of the individual benchmarks is different,

The trophy would not fit in the brown suitcase because it was too big (small).
 What was too big (small)?
 Answer 0: the trophy
 Answer 1: the suitcase

Fig. 1 Example from the Winograd Schema Challenge. The difficulty of these problems becomes clear when you consider that replacing the word *big* with the word *small* changes the answer to be given, although the sentence has not changed syntactically

most individual problems have a similar structure: A (short) text describes a situation for which a question is asked and is presented with several possible answers from which the correct answer is to be determined. See Fig. 1 for an example from the Winograd Schema Challenge [51].

Most of the approaches that consider these benchmarks are subsymbolic in nature (see Sect. 2.3). In the following, we will discuss the problems that symbolic approaches to commonsense reasoning benchmarks have to face.

What all benchmarks have in common is that they require a great deal of background knowledge, including people's everyday (commonsense) knowledge. This knowledge is broad and contains, for example, knowledge about everyday physical relationships like "usually, things you throw down fall down" but also knowledge about interpersonal relationships. Such background knowledge can be in the form of first-order logic knowledge bases, ontologies, knowledge graphs or text given in natural language. Ideally, symbolic approaches do not only rely on one possible source of background knowledge, but are able to combine heterogeneous sources of background knowledge.

Nutcracker, for instance, is a pipeline for natural language understanding based on the Boxer system [16] that translates text into first-order logic formulae and tries to find a proof for the given task with the help of an automated theorem prover. The Nutcracker system allows to include first-order logic knowledge bases as background knowledge into the pipeline. By default, it uses a knowledge base created from hypernymy and synonymy relations found in WordNet [58]. Even if the Nutcracker pipeline is equipped with a disambiguation tool and other axioms derived from FrameNet [8], Nutcracker's results on textual entailment benchmarks are not encouraging. The main problem is the lack of background knowledge [68]: Only a small part of the knowledge represented in WordNet and FrameNet is actually commonsense knowledge which causes the reasoner of the Nutcracker pipeline to struggle with incomplete knowledge.

Unlike the Nutcracker pipeline, which aims at solving problems by having a reasoner construct a proof, the reasoner used in the CoRg project, the first-order logic theorem prover Hyper [9, 10], is only used to conduct some useful inferences. This (potentially partial) model is being studied by machine learning techniques to find out what

possible answer the inferences points to. Involving machine learning is supposed to remedy the problem of incomplete background knowledge. Although WordNet together with OpenCyc [50] or Adimen-SUMO [1] are considered as background knowledge in the CoRg project, the results on the benchmarks considered cannot keep up with the results of subsymbolic methods. This is attributed to the still incomplete background knowledge, because OpenCyc and Adimen-SUMO represent taxonomic knowledge rather than commonsense knowledge. Recent experiments [12] describe an approach integrating word embeddings into the selection process of background knowledge resulting in a selection which is able to select parts useful for a certain commonsense problem from a big knowledge base.

2.3 Subsymbolic Approaches

Cognitive reasoning is often combined with machine learning, because logical reasoning alone is not sufficient to handle commonsense reasoning tasks, in particular if the problems are given in natural language and not by logic programs or similar formulations. Within commonsense reasoning tasks, recurrent networks with LSTM (long short-term memory) [39] are a promising strategy. They achieve a success rate of up to 84% on benchmarks like the machine comprehension task in SemEval 2018 [67]. It is a collection of narrative texts, questions of various types referring to these texts, and pairs of answer candidates for each question. It comprises 2119 such texts and a total of 13,939 questions. Most of the teams participating in the machine comprehension task used neural approaches in combination with ontological knowledge like ConceptNet [81]. One drawback of machine learning, especially neural networks, is that many examples are needed for training. However, this problem can at least partially be overcome by unsupervised pre-training with natural-language data (cf. [18, 72]).

From a behavioral point of view, pure machine learning approaches are able to solve commonsense reasoning tasks. However, there is no representation of the reasoning process and, hence, no explanation in these systems. The reason is that machine learning with deep learning neural networks [36] works as a black box only. Cognitive reasoning gives us explanations of answers in addition. *Explainable artificial intelligence* generally is artificial intelligence that is programmed to describe its purpose, rationale and decision-making process in a way that can be understood by most people [79]. In this context, *post-hoc* and *ante-hoc* analysis can be distinguished [44]: Models of the former type explain the given answer afterwards, e.g. by inspecting a learned neural network, while in the other case the model itself is explanatory, e.g. a first-order formulae set or a logic program. In order to achieve this goal, reasoning techniques from the fields of deduction, logics, and nonmonotonic

reasoning can be employed and combined with machine learning. Machine learning can be used as a subsystem to improve the reasoning process.

This has been called the *neural symbolic cycle* (see e.g. [7, 13, 30, 38]). Given declarative knowledge in the form of a logic program, it was first shown in [41] that such knowledge can be compiled into a feed-forward network of binary threshold units computing the immediate consequence operator [5] of the given logic program. Turned into a recurrent network by connecting the output layer of the feed-forward network to its input layer, the recurrent network computes the least fixed point of the immediate consequence operator, which is equal to be the least model of the given logic program. Reasoning is performed with respect to this least model. As first shown in [31] the logic program can also be compiled into a feed-forward network of sigmoidal units such that the network becomes trainable using backpropagation. In other words, the immediate consequence operator of a logic program can be learned. After training, however, an updated logic program needs to be extracted to obtain declarative knowledge again. There are various approaches for extracting rules from trained artificial neural networks [19] which can be applied to obtain explanations in the sense of explainable artificial intelligence. The initial approach has been extended in various ways (see e.g. [32, 33, 42, 61]) including commonsense or nonmonotonic reasoning approaches (cf. Sect. 2.1 and 2.2). Recently, the weak completion semantics has also been mapped onto artificial neural networks [22].

3 Challenges, Competitions, and Theoretical Research Questions

To foster research and scientific progress in the field of cognitive reasoning, we propose some interesting challenges. Our goal here is to start a discussion in the field based on a necessarily incomplete list of challenges. This strategy has been especially fruitful in leading to progress in mathematics or generally in artificial intelligence, e.g., in action planning⁸ or specifically in theorem proving [85]. We first propose some computational modeling challenges that have a clear task and some benchmark data.

3.1 Computational Modeling Challenges

We summarize in the following different domains that have been established.

1. Domain of commonsense reasoning:
 - (a) Task: develop a computational model that can compete with a (human) reasoner on problems that require common sense.
 - (b) Benchmarks: Winograd Challenge [51], TriangleCOPA [37] and many others.

As described in the previous section, Winograd consists of natural language based problems, whereas TriangleCOPA problems are already formulated in logic.
2. Domain of cognitive computational models for individual human reasoning:
 - (a) Task: develop a computational model that can predict an individual reasoner (on syllogistic, propositional, and relational problems).
 - (b) Benchmarks: For cognitive modeling a framework providing data for individual human reasoner has been developed: CCOBRA⁹
3. Domain of developing cognitive architectures:
 - (a) Task: develop a general, modular architecture that reflects and constrains the cognitive processes underlying human reasoning; often connecting symbolic and connectionist approaches.
 - (b) Benchmarks: not yet developed, but principles are formulated e.g., cognitive core criteria [25, 80].
4. Domain of natural-language based deep question answering
 - (a) Task: find out causal relationships from general natural language texts.
 - (b) Benchmarks: COPA (Choices of Plausible Alternatives) [78], SemEval [67], StoryClozeTest [59]
5. Domain of combining reasoning and machine learning in one integrated framework:
 - (a) Task: extract explanatory models from black-box approaches leading to general explainable artificial intelligence systems
 - (b) Benchmarks: TPTP [84]

Pure machine learning approaches are successful in some natural-language benchmarks (cf. Sect. 2.3) but usually lack explanations, whereas

⁸ <http://www.icaps-conference.org/index.php/Main/Competitions>, accessed: 13-May-2019.

⁹ http://orca.informatik.uni-freiburg.de/orca_sylwebsite/orca/, accessed: 13-May-2019.

automated reasoning often requires completely formalized input, e.g. by first-order logic. The challenge is to combine both worlds.

3.2 Research Questions

In the following we discuss some questions that are relevant for identifying the characteristics of human cognitive reasoning.

3.2.1 Mental Representations

Preference effect. If information is not specific enough allowing for the construction of several possible models then human reasoners in spatial reasoning often prefer to construct a specific representation that is minimal with respect to the number of mental operations [75]. This preference effect has been shown for other domains as well [74]. A challenge is to identify, when and how indeterminate descriptions lead to such preferred representations and to show how these depend on working memory descriptions.

Demarcation of System 1 and System 2 Reasoning. Humans can reason heuristically and analytically. Current research demonstrates that these reasoning strategies can be associated with different reasoning systems (see Sect. 2.1). But when and how is which system employed?

Number of Truth Values in Cognitive Reasoning. Two truth values are not sufficient to model human reasoning, but three truth values can, see [73]. A challenge is to identify how many truth values are sufficient to model individual human inferences.

Conditionals with Unknown Conclusions. How do humans evaluate conditional statements such as *if A then C*, when they cannot determine the truth value of the conclusion in the conditional, i.e., if it is unknown?

Counterfactuals with Unknown Antecedents. A *counterfactual* is a conditional statement with a false antecedent such as “if Oswald had *not* shot Kennedy, then someone else would have” [15]. Counterfactuals with unknown antecedents can be evaluated in two different ways: Firstly, we may extend the background knowledge such that the antecedent becomes false and, thereafter, revise the background knowledge such that the antecedent becomes true. Secondly, we may extend the background knowledge such that the antecedent becomes true. Under the weak completion semantics the two approaches may lead to different results. How do humans handle counterfactuals with unknown antecedents?

3.2.2 Inference Mechanisms

In all applications of the weak completion semantics, skeptical abduction has been adequate whereas credulous abduction has led to answers which were not given by humans.

The computation of skeptical conclusions is exponential in the number of abducibles. Hence, in extended reasoning episodes it appears to be unlikely that humans consider all possible explanations. Rather we believe that they consider only a bounded number of explanations and reason skeptically with respect to the considered explanations. But how do humans restrict the number of possible explanations? Do they prefer some explanations over others? If so, what does the preference relation look like?

3.2.3 Consciousness

Consciousness certainly plays an important role in the nature of human intelligence. It is a topic in literature, philosophy and, of course, in psychology. In philosophy the qualia problem, i.e., the problem of making a subjective personal conscious experience, is discussed extensively. The question is whether or not it is possible to experience “what is it to be a ...”—for the most famous example in [62]: “What is it like to be a bat?” Of course there are also arguments against qualia, a very prominent opponent is Daniel Dennett, who is arguing against qualia from a neuroscience viewpoint and by results from experiments that every experience is combined with a neural process [17].

In artificial intelligence and cognitive science there are various approaches towards an operational definition of consciousness. Don Perlis and his co-authors aim at operationalizations via building intentionality into artificial intelligence systems. By this, agents would be able to be aware that their actions are done by them and that their body is theirs [71]. Another approach offers consciousness as a method for handling the vast amount of knowledge a cognitive system has to deal with: The global workspace theory of John Baars [6] which can be easily described by a theater metaphor. There are agents (i.e. sensor inputs) on a stage which are trying to get into the bright spotlight of attention. There is a lot of personnel behind the scene, like authors, technicians or directors who help keeping the theater running. And there is a huge audience in the dark, which represents the knowledge about the world and about past experiences. This entire theater can be understood as consciousness. For more details and how this can be operationalized within a cognitive reasoning system, see [29].

Altogether, reasoning in cognitive systems obviously can benefit from both approaches towards an operational definition of consciousness as described above.

4 Summary

Cognitive reasoning is a highly interdisciplinary field: Methods from cognitive science, artificial intelligence, automated deduction, philosophy and psychology are applied. In

contrast, to a superficial perspective, it neither aims at just modeling human reasoning nor is it just focusing on a general rational optimal reasoning process. It is rather motivated by the insight that humans can demonstrate a specific way of thinking that has qualities that go beyond existing logical formalisms. But it provides features like explainability, non-monotonicity, and generalizability across domains that existing approaches from artificial intelligence and machine learning have not yet demonstrated.

References

1. Álvarez J, Lucio P, Rigau G (2012) Adimen-SUMO: reengineering an ontology for first-order reasoning. *Int J Semant Web Inform Syst (IJSWIS)* 8(4):80–116
2. Anderson JR (2007) How can the human mind occur in the physical universe?. Oxford University Press, New York
3. Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y (2004) An integrated theory of the mind. *Psychol Rev* 111(4):1036–1060. <https://doi.org/10.1037/033-295X.111.4.1036>
4. Anshakov OM, Gergely T (2010) Cognitive reasoning—a formal approach. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-540-68875-4>
5. Apt K, van Emden M (1982) Contributions to the theory of logic programming. *J ACM* 29:841–862
6. Baars BJ (1997) In the theatre of consciousness. global workspace theory, a rigorous scientific theory of consciousness. *J Conscious Stud* 4(4):292–309
7. Bader S (2009) Neural-symbolic integration. Ph.D. thesis, Technische Universität Dresden, Faculty of Computer Science
8. Baker CF, Fillmore CJ, Lowe JB (1998) The Berkeley framenet project. In: C. Boitet, P. Whitelock (eds) 36th annual meeting of the association for computational linguistics and 17th international conference on computational linguistics, COLING-ACL '98, August 10–14, 1998, Université de Montréal, Montréal, Quebec, Canada. Proceedings of the Conference., pp. 86–90. Morgan Kaufmann Publishers / ACL. <http://aclweb.org/anthology/P/P98/P98-1013.pdf>
9. Baumgartner P, Furbach U, Niemelä I (1996) Hyper tableaux. In: European workshop on logics in artificial intelligence, Springer, pp 1–17
10. Bender M, Pelzer B, Schon C (2013) System description: E-KRHyper 1.4. In: International conference on automated deduction, Springer Nature Switzerland AG, pp 126–134
11. Bentivogli L, Dagan I, Magnini B (2017) The recognizing textual entailment challenges: datasets and methodologies. In: Ide N, Pustejovsky J (eds) Handbook of linguistic annotation. Springer, Berlin
12. Bergk T, Furbach U, Schon C (2019) Names are not just sound and smoke: word embeddings for axiom selection. In: Conference on automated deduction, CADE 27. Springer Nature Switzerland AG
13. Besold TR, d'Ávila Garcez AS, Bader S, Bowman H, Domingos PM, Hitzler P, Kühnberger KU, Lamb LC, Lowd D, Lima PMV, de Penning L, Pinkas G, Poon H, Zaverucha G (2017) Neural-symbolic learning and reasoning: a survey and interpretation. CoRR. [arXiv:abs/1711.03902](https://arxiv.org/abs/1711.03902)
14. Braine MDS, O'Brien DP (1998) Mental logic. Erlbaum, Mahwah
15. Byrne RM, Tasso A (1999) Deductive reasoning with factual, possible, and counterfactual conditionals. *Mem Cognit* 27(4):726–740
16. Curran JR, Clark S, Bos J (2007) Linguistically motivated large-scale NLP with C&C and Boxer. In: Proceedings of the ACL 2007 demo and poster sessions, Prague, Czech Republic. Association for Computational Linguistics, Stroudsburg PA, USA, pp. 33–36
17. Dennett D (1993) Consciousness explained. Penguin books. Penguin Adult. https://books.google.de/books?id=d2P_QS6AwgoC. Accessed 27 June 2019
18. Devlin J, Chang M, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. CoRR. [arXiv:1810.04805](https://arxiv.org/abs/1810.04805)
19. Diederich J, Tickle AB, Geva S (2010) Quo vadis? Reliable and practical rule extraction from neural networks. In: Koronacki J, Ras ZW, Wierzbichon ST, Kacprzyk J (eds) Advances in machine learning I: dedicated to the memory of professor Ryszard S. Michalski, studies in computational intelligence, vol 262. Springer, Berlin, pp 479–490. https://doi.org/10.1007/978-3-642-05177-7_24
20. Dietz EA, Hölldobler S, Ragni M (2012) A computational logic approach to the suppression task. In: N Miyake, D Peebles, RP Cooper (eds) Proceedings of the 34th annual conference of the cognitive science society, Cognitive Science Society. Curran Associates Inc. proceedings.com, pp 1500–1505
21. Dietz EA, Hölldobler S, Ragni M (2013) A computational logic approach to the abstract and the social case of the selection task. In: Proceedings eleventh international symposium on logical formalizations of commonsense reasoning. <http://commonsensereasoning.org/2013/proceedings.html>
22. Dietz Saldanha EA, Hölldobler S, Kencana Ramli CDP, Palacios Medinacelli L (2018) A core method for the weak completion semantics with skeptical abduction. *J Artif Intell Res* 63:51–86
23. Dietz Saldanha EA, Hölldobler S, Lourêdo Rocha I (1994) The weak completion semantics. In: C Schon, U Furbach (eds) Proceedings of the workshop on bridging the gap between human and automated reasoning—is logic and automated reasoning a foundation for human reasoning?, vol 1994, pp 18–30. CEUR-WS.org (2017). <http://ceur-ws.org/Vol-1994/>
24. Dietz Saldanha EA, Hölldobler S, Mörbitz R (2018) The syllogistic reasoning task: reasoning principles and heuristic strategies in modeling human clusters. In: Seipel D, Hanus M, Abreu S (eds) Declarative programming and knowledge management, vol 10997. Lecture notes in artificial intelligence. Springer, Berlin, Heidelberg, pp 149–165
25. Eliasmith C (2013) How to build a brain: a neural architecture for biological cognition. Oxford University Press, Oxford
26. Evans JSB (2003) In two minds: dual-process accounts of reasoning. *TRENDS Cogn Sci* 7(10):459–459
27. Evans JSB (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Ann Rev Psychol* 59:255–278
28. Ferrucci DA, Brown EW, Chu-Carroll J, Fan J, Gondek D, Kalyanpur A, Lally A, Murdock JW, Nyberg E, Prager JM, Schlaefer N, Welty CA (2010) Building watson: an overview of the DeepQA project. *AI Mag* 31(3):59–79
29. Furbach U, Schon C (2018) Reasoning and Consciousness. Teaching a Theorem Prover to let its Mind Wander. In: The third conference on artificial intelligence and theorem proving, AITP2018. <http://aitp-conference.org/2018/aitp18-proceedings.pdf>
30. d'Ávila Garcez A, Lamb L, Gabbay D (2009) Neural-symbolic cognitive reasoning. Springer, Berlin, Heidelberg
31. d'Ávila Garcez A, Zaverucha G, de Carvalho L (1997) Logic programming and inductive learning in artificial neural networks. In: Herrmann C, Reine F, Strohmaier A (eds) Knowledge representation in neural networks. Logos Verlag, Berlin, pp 33–46

32. d'Avila Garcez AS, Broda K, Gabbay DM (2001) Symbolic knowledge extraction from trained neural networks: a sound approach. *Artif Intell* 125(1–2):155–207. [https://doi.org/10.1016/S0004-3702\(00\)00077-1](https://doi.org/10.1016/S0004-3702(00)00077-1)
33. d'Avila Garcez AS, Zaverucha G (1999) The connectionist inductive learning and logic programming system. *Appl Intell* 11(1):59–77. <https://doi.org/10.1023/A:1008328630915>
34. Gigerenzer G, Selten R (2002) Bounded rationality: the adaptive toolbox. MIT Press, Cambridge, MA, USA
35. Gigerenzer G, Todd P (1999) Simple heuristics that make us smart. Oxford University Press, New York
36. Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. Adaptive Computation and Machine Learning. MIT Press, Cambridge, London. <http://www.deeplearningbook.org>
37. Gordon AS (2016) Commonsense interpretation of triangle behavior. In: D Schuurmans, MP Wellman (eds) Proceedings of the thirtieth AAAI conference on artificial intelligence, February 12–17, 2016, Phoenix, Arizona, USA., pp 3719–3725. AAAI Press <http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/11790>
38. Hammer B, Hitzler P (eds) (2007) Perspectives of neural-symbolic integration. Springer, Berlin, Heidelberg
39. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
40. Hölldobler S (2018) Ethical decision making under the weak completion semantics. In: C Schon (ed) Proceedings of the workshop on bridging the gap between human and automated reasoning, vol 2261, pp 1–5. CEUR-WS.org. <http://ceur-ws.org/Vol-2261/>
41. Hölldobler S, Kalinke Y (1994) Towards a new massively parallel computational model for logic programming. In: Proceedings of the ECAI94 workshop on combining symbolic and connectionist processing, ECCAI - European Association for Artificial Intelligence, pp. 68–77
42. Hölldobler S, Kalinke Y, Störr H (1999) Approximating the semantics of logic programs by recurrent neural networks. *Appl Intell* 11(1):45–58. <https://doi.org/10.1023/A:1008376514077>
43. Hölldobler S, Kencana Ramli CDP (2009) Logic programs under three-valued Łukasiewicz's semantics. In: Hill PM, Warren DS (eds) LNCS, vol 5649. Springer-Verlag, Berlin Heidelberg, pp 464–478
44. Holzinger A (2018) Explainable AI (ex-AI). *Informatik Spektrum* 41(2):138–143. <https://doi.org/10.1007/s00287-018-1102-5> (**Aktuelles Schlagwort, in German**)
45. Indiveri G, Linares-Barranco B, Hamilton T, van Schaik A, Etienne-Cummings R, Delbruck T, Liu SC, Dudek P, Häfziger P, Renaud S, Schemmel J, Cauwenberghs G, Arthur J, Hynna K, Folowosele F, Saïghi S, Serrano-Gotarredona T, Wijekoon J, Wang Y, Boahen K (2011) Neuromorphic silicon neuron circuits. *Front Neurosci* 5:73. <https://www.frontiersin.org/article/10.3389/fnins.2011.00073>
46. Johnson-Laird PN (2006) How we reason. Oxford University Press, New York
47. Kahneman D (2011) Thinking, fast and slow. Macmillan Publishers, London
48. Khemlani S, Johnson-Laird PN (2012) Theories of the syllogism: a meta-analysis. *Psychol Bull* 138(3):427–457
49. Kompridis N (2000) So we need something else for reason to mean. *Int J Philos Stud* 8:271–295
50. Lenat DB (1995) CYC: a large-scale investment in knowledge infrastructure. *Commun ACM* 38(11):33–38
51. Levesque HJ (2011) The winograd schema challenge. In: Logical formalizations of commonsense reasoning, papers from the 2011 AAAI Spring Symposium, Technical Report SS-11-06, Stanford, California, USA, March 21–23, 2011. AAAI. <http://www.aaai.org/ocs/index.php/SSS/SSS11/paper/view/2502>
52. Liao Y, Li H (2017) Reservoir computing trend on software and hardware implementation. *Global J Res Eng (F)* 17(5). <https://engineeringresearch.org/index.php/GJRE/article/download/1654/1585>
53. Łukasiewicz J (1920) O logice trójwartościowej. *Ruch Filozoficzny* 5:169–171. English translation: On Three-Valued Logic. In: Jan Łukasiewicz Selected Works. (L. Borkowski, ed.), North Holland, 87–88, 1990
54. Marr D (1982) Vision: a computational investigation into the human representation and processing of visual information. W. H. Freeman and Company, New York
55. McClelland J (2009) The place of modeling in cognitive science. Topics in cognitive science. Wiley Online Library, Hoboken
56. Mead C (1990) Neuromorphic electronic systems. *Proceedings of the IEEE* 78(10):1629–1636. <https://ieeexplore.ieee.org/document/58356>
57. Michael L (2019) Cognitive reasoning and learning mechanisms. In: Proceedings of the 2019 towards conscious AI system symposium, vol. 2287. CEUR Workshop Proceedings
58. Miller GA (1995) WordNet: a lexical database for english. *Commun ACM* 38(11):39–41
59. Mostafazadeh N, Roth M, Louis A, Chambers N, Allen J (2017) LSDSem 2017 shared task: The story cloze test. In: Proceedings of the 2nd workshop on linking models of lexical, sentential and discourse-level semantics. Association for Computational Linguistics, pp 46–51
60. Mueller ET (2014) Commonsense reasoning, 2nd edn. Morgan Kaufmann, San Francisco
61. Muggleton SH, Schmid U, Zeller C, Tamaddoni-Nezhad A, Besold T (2018) Ultra-strong machine learning: comprehensibility of programs learned with ilp. *Mach Learn* 107:1119–1140. <https://doi.org/10.1007/s10994-018-5707-3>
62. Nagel T (1974) What is it like to be a bat? *Philos Rev* 83(4):435–450. <http://www.jstor.org/stable/2183914>
63. Newell A (1990) Unified theories of cognition. Harvard University Press, Cambridge
64. Newell A, Simon HA (1972) Human problem solving. Prentice-Hall, Englewood Cliffs
65. Oaksford M, Chater N (2007) Bayesian rationality: the probabilistic approach to human reasoning. Oxford University Press, Oxford Cognitive Science Series
66. Oliviera da Costa A, Dietz Saldanha EA, Hölldobler S, Ragni M (2017) A computational logic approach to human syllogistic reasoning. In: Gunzelmann G, Howes A, Tenbrink T, Davelaar EJ (eds) Proceedings of the 39th annual conference of the cognitive science society. Cognitive Science Society, Austin, pp 883–888888
67. Ostermann S, Roth M, Modi A, Thater S, Pinkal M (2018) SemEval-2018 task 11: Machine comprehension using commonsense knowledge. In: Proceedings of the 12th international workshop on semantic evaluation. Association for Computational Linguistics, pp 747–757
68. Ovchinnikova E (2012) Integration of world knowledge for natural language understanding. *Atlantis thinking machines*, vol 3. Atlantis Press, New York. <https://doi.org/10.2991/978-94-91216-53-4>
69. Pagel P, Portmann E, Vey K (2018) Cognitive computing. *Informatik Spektrum* 41(1–2) (**Edited special issues**)
70. Pereira LM, Dietz EA, Hölldobler S (2014) An abductive reasoning approach to the belief-bias effect. In: Baral C, Giacomo GD, Eiter T (eds) Principles of knowledge representation and reasoning: proceedings of the 14th international conference. AAAI Press, Cambridge, pp 653–656656
71. Perlis D, Brody J (2019) Operationalizing consciousness. In: Proceedings of the 2019 towards conscious AI system symposium, vol 2287

72. Radford A, Narasimhan K, Salimans T, Sutskever I (2018) Improving language understanding by generative pre-training. <https://openai.com/blog/language-unsupervised/>. Accessed: 13 May 2019
73. Ragni M, Dietz EA, Kola I, Hölldobler S (2016) Two-valued logic is not sufficient to model human reasoning, but three-valued logic is: a formal analysis. In: U Furbach, C Schon (eds) Bridging 2016—bridging the gap between human and automated reasoning, CEUR Workshop Proceedings, vol 1651, pp 61–73. CEUR-WS.org. <http://ceur-ws.org/Vol-1651/>
74. Ragni M, Khemlani S, Johnson-Laird PN (2013) The evaluation of the consistency of quantified assertions. *Mem Cognit* 42(1):53–66
75. Ragni M, Knauff M (2013) A theory and a computational model of spatial reasoning with preferred mental models. *Psychol Rev* 120(3):561–588
76. Rips LJ (1994) *The psychology of proof: deductive reasoning in human thinking*. The MIT Press, Cambridge
77. Roemmele M, Bejan CA, Gordon AS (2011) Choice of plausible alternatives: an evaluation of commonsense causal reasoning. In: Logical formalizations of commonsense reasoning, Papers from the 2011 AAAI Spring Symposium, Technical Report SS-11-06, Stanford, California, USA, March 21–23, 2011. AAAI. <http://www.aaai.org/ocs/index.php/SSS/SSS11/paper/view/2418>
78. Roemmele M, Bejan CA, Gordon AS (2011) Choice of plausible alternatives: an evaluation of commonsense causal reasoning. In: AAAI spring symposium: logical formalizations of commonsense reasoning, pp 90–95
79. Rouse M, Wigmore I (2019) Definition explainable AI (XAI) (2018). <https://whatis.techtarget.com/definition/explainable-AI-XAI>. Accessed: 13 May 2019
80. Simon H, Wallach D (1999) Cognitive modeling in perspective. *Kognitionswissenschaft* 8:1–4
81. Speer R, Chin J, Havasi C (2017) ConceptNet 5.5: an open multilingual graph of general knowledge. In: AAAI conference on artificial intelligence, pp 4444–4451 <http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14972>
82. Stenning K, van Lambalgen M (2005) Semantic interpretation as computation in nonmonotonic logic: the real meaning of the suppression task. *Cogn Sci* 29:919–960
83. Stenning K, van Lambalgen M (2008) *Human reasoning and cognitive science*. MIT Press, Cambridge
84. Sutcliffe G (2017) The TPTP problem library and associated infrastructure. From CNF to TH0, TPTP v6.4.0. *J Autom Reason* 59(4):483–502
85. Sutcliffe G (2018) The 9th IJCAR automated theorem proving system competition—CASC-J9. *AI Commun* 31(6):495–507. <https://doi.org/10.3233/AIC-180773>