# The Search for Cognitive Models: Standards and Challenges

Marco Ragni and Nicolas Riesterer

Cognitive Computation Lab
Albert-Ludwigs-Universität Freiburg

**Abstract.** Cognitive modeling is the distinguishing factor of cognitive science and the method of choice for formalizing human cognition. In order to bridge the gap between logic and human reasoning, a number of foundational research questions need to be rigorously answered. The objective of this paper is to present relevant concepts and to introduce possible modeling standards as well as key discussion points for cognitive models of human reasoning.

**Keywords:** Cognitive Modeling; Human Reasoning; Logic

## 1 Introduction

All sciences are defined by their respective research objectives and methods. Cognitive science in particular is special in this regard, because it is an interdisciplinary field located at the boundaries of many other domains of research such as artificial intelligence, psychology, linguistics, computer science, and neuroscience. As a result, its goals and methods are diverse mixtures with influences from the neighboring fields.

The core research question of cognitive science focuses on investigating information processing in the *human mind* in order to gain an understanding of human cognition as a whole. To this end, it primarily employs the method of *cognitive modeling* as a means of capturing the latent natural processes of the mind by well-defined mathematical formalizations. The challenge of cognitive modeling is to develop models which are capable of representing highly complex and potentially unobservable processes in a computational manner while still guaranteeing their interpretability in order to advance the level of understanding of cognition.

This paper discusses high-level cognitive models of reasoning. In particular, it gives a brief introduction into the following three core research questions:

1. What characterizes a cognitive model?
2. What is a "good" cognitive model?
3. What are current challenges?

## 2 What is a cognitive model? Step 1: Model Generation

A theory of reasoning is defined as *cognitively adequate* [22] with respect to a reasoning task $T$ and a human reasoner $R$, if the theory is (i) *representationally adequate*, i.e., it uses the same mental representation as a human reasoner does, (ii) *operationally adequate*, i.e., the theory specifies the same operations the reasoner employs, and (iii) *inferentially adequate*, i.e., the theory draws the same conclusions based on the operations and mental representation as the human reasoner. While the inferential adequacy of a theory $T$ can be determined from the given responses of a reasoner for a given task, it is impossible to directly observe the operations and mental representations a reasoner applies. They can only be determined by means of *reverse engineering*, i.e. the identification of functionally equivalent representations and operations leading to the generation of a given reasoner's output.

A mental representation is localized within the specific cognitive architecture of the human mind, which the reasoning process operates on. Hence, we need to distinguish between cognitive architectures and a cognitive models. A *cognitive architecture* is a tuple $\langle \mathcal{D}, \mathcal{O} \rangle$ consisting of a data structure $\mathcal{D}$ (which can contain an arbitrary number of substructures) and a set of operations $\mathcal{O}$ specified in any formal language to manipulate the data structure. The goal of a cognitive architecture is to specify the often type dependent flow of information (e.g., visual or auditory) between different memory-related cognitive structures in the human mind. This imposes constraints on the data structures of the reasoner and the corresponding mental operations. An example for a cognitive architecture is ACT-R which uses so-called *modules*, i.e., data structures for specific types of information, and production rules as a set of general operations [2].

A *cognitive computational model* for a cognitive task $T$ in a given cognitive architecture specifies algorithms based on (a subset of) operations defined on the data structure of the underlying cognitive architecture. The application of those algorithm results in the computation of an input-output mapping for the cognitive task $T$ with the goal of representing human cognition.

## 3 What is a "good" cognitive model? Step 2: Model Evaluation

The definition of a cognitive computational model (cognitive model for short) is rather general and allows for a large space of possible model candidates. Driven by the motivation that a cognitive theory should be explanatory for human performance, a "good" cognitive model is never just a *simulation model*, i.e., a model that solely reproduces existing experimental data. Instead, it must always make explicit assumptions about the latent workings of the mind.

Based on several criteria from the literature [18] the following list can serve as a starting point for defining principles of "good" cognitive modeling:

1. *The model has transparent assumptions.* All operations and parameters are transparent and the model's responses can be explained by model operations.
2. *The model's principles are independent from the test data.* A model cannot be developed on the same data it is tested on. To avoid overfitting to a specific dataset, fine-tuning on the test data is not allowed.
3. *The model generates quantitative predictions.* The model computes the same type of answers a human reasoner gives based on the input she receives. Model predictions can be compared with the test data by mathematical discrepancy functions often applied in mathematical psychology and AI, such as the Root-Mean-Square Error (RMSE), statistical information criteria, or others (see below).
4. *The model predicts behavior of an individual human reasoner.* Often, models predict an average reasoner. However, aggregating data increases the noise and eliminates individual differences.
5. *The model covers several relevant reasoning phenomena and predicts new phenomena.* The goal of a cognitive model is not to just fit data perfectly, but to explain latent cognitive reasoning processes in accordance with the results obtained from psychological studies. Ultimately, models are supposed to offer an alternative view on cognition allowing for the derivation of new phenomena that can be validated or falsified by conducting studies on human reasoners.

These points also introduce an ordering based on the importance of the modeling principles. Points 1 and 2 are general requirements we consider to be mandatory for any serious modeling attempts. Points 4 and 5 are important for general cognitive models which are supposed to shed light on the inner workings of the mind. For the reverse engineering process and a comparison of different models that share all points 1-5, criteria 3 is the most important one.

There are different methods for assessing the quality of models. On their very basis, they all share the idea of defining a discrepancy metric that can be used to quantify the value of a specific model in comparison with others. Most fundamentally, the RMSE defines the discrepancy based on the distance between the model predictions and outcomes observed in real world experiments. More sophisticated statistical approaches based on the likelihood of data, such as the $\chi^2$ or $G^2$ metrics, can be interpreted as test statistics with significant results indicating large differences to the data [3]. However, since models do not only differ with respect to the goodness of fit, but also with respect to their complexity, further information must often be integrated into the model comparison process. Akaike's Information Criterion ($AIC$) [1] and the Bayesian Information Criterion ($BIC$) [21] are metrics based on $G^2$, that incorporate the number of free parameters as an indication of complexity. $FIA$ is an information theoretic approach that quantifies complexity based on the minimum description length principle [8]. Furthermore, there are purely Bayesian approaches to the problem of model comparison. By relying on Bayes' Theorem, the Bayes Factor ($BF$) measures the relative fit of two models by integrating uncertainties about the data and parameters under the model. It quantifies whether the data provides more evidence for or against one model being compared with an alternative [15].

## 4 Challenges

The field of cognitive science can benefit greatly from interdisciplinary work and results. This ranges from the application of recent advances in modeling methods from computer science and statistics, and extends all the way to exploiting the knowledge gained in the field of theoretical psychology.

However, in order to foster this collaborative approach that could potentially result in faster and more goal-oriented progress, the field needs to address several open questions and relevant challenges:

1. **What are relevant benchmark problems?** In computer science and AI, well-defined benchmark problems have been great aids to the field. By organizing annual competitions and generally maintaining low barriers for entry, progress could be boosted in various domains, such as planning or satisfiability solving for logic formulae. Additionally, the rigorous definition of benchmarks allowed for a fair comparison between different approaches based on well-defined criteria triggering a competitive spirit for improving the state-of-the-art of the respective domains.
   We see the necessity to introduce the field of cognitive science and especially the domain of human reasoning to the concept of competition, as well. Without defining benchmark problems and providing the data to approach them in a clear and direct manner, the field risks to drown in the continuously increasing stream of cognitive theories arguing to explain parts of human reasoning. In order to guarantee progression, we consider the definition of explicit criteria for model comparison and their application based on commonly accepted and publicly available datasets mandatory.
   While psychological experiments can provide benchmark problems, they need to be differentiated with respect to priority. So far, no criteria for the identification of *relevant* problems have been introduced in the literature. However, they are necessary for the development of a generally accepted benchmark. The following list compiles experiments and phenomena as well as general remarks that should be taken into account when formalizing a benchmark problem:
   (a) Phenomena/experiments that have often been modeled and/or cited are:
      – Conditional and propositional reasoning:
         • Simple conditional inferences [17]
         • Counterfactual reasoning [5]
         • Rule testing: The Wason Selection Task [**?**]
         • Illusions in propositional reasoning [9]
         • Suppression effect [4]
      – Relational reasoning:
         • Preference effect [19]

    - Pseudo-transitive relations [7]
    - Complex relations [6]
    - Indeterminacy effect [4]
    - Visual impedence effect [13]
  - Syllogistic reasoning:
    - Reasoning patterns on the 64 syllogisms [10, 11]
    - Belief-bias effect [12]
    - Generalized quantifiers [16]

   (b) Data from the literature that can be included in the a benchmark needs to include a description of the information a reasoner received as well as her response. Aggregated data of individual reasoners can help to formulate an intuition or can give an indication for an effect. However, for developing a profound model, answers of the individual reasoners are necessary.

2. **How to translate existing descriptive theories into computational cognitive models?**
   Most cognitive theories are not defined algorithmically. Instead they are often based on verbal descriptions alone. However, for purposes of fair mathematical comparison, a formalization of these theories is required. The challenge here is to develop a model implementation of the theory that is as close to the original theory as possible while making all additional assumptions made by the modeler explicit. There currently is no accepted methodology for general theory implementation.
3. **How could a general cognitive modeling language be specified?**
   The field of action planning greatly benefits from having a general Planning Domain Definition Language (PDDL). On one hand, PDDL allows for the easy definition and introduction of new problems. On the other hand, it forces planners to be defined generally without exploiting domain-dependent shortcuts and heuristics.
   Especially when considering the goal to construct a model for unified cognition, finding a common cognitive modeling language might be beneficial. However, the task of defining a language which is accepted by most modelers is not an easy endeavor as the list of potential reasoning domains is quite extensive, and each has its own specific set of requirements. Additionally, there are very different modeling approaches beyond the purely symbolic methods that are commonly found in planning introducing even more complexity for the language desired. Examples include models based on artificial neurons, hybrid approaches, Bayesian models, and abstract description based models such as Multinomial Processing Trees (MPTs).
4. **What are properties of the human data structures that influence the reasoning process?**
   While working memory is resource bounded, long-term memory is not. But there are additional cognitive features that can have an influence on reasoning such as the background knowledge, cognitive bottlenecks, parallel processing, etc. These limitations are often not represented in cognitive theories

but crystallized in cognitive architectures [14]. However, general approaches for developing and comparing these architectures have yet to be identified.

## 5 Desirables: Standards, Networks, and Competitions

### 5.1 Cognitive Modeling Standards

Cognitive models are usually developed in a post-hoc fashion with the goal to fit to an existing set of experimental data. Alternatively, cognitive models can be created following a mixture of data- and theory-based approaches with undefined overlap. Irrespective of the motivation and development process, a fair comparison of models must be based on well-defined criteria (such as those introduced in Section 3). Generally, the research community of the field needs to settle on which criteria are mandatory, which are desirable, and which are not worthwile to pursue further. In order to develop and maintain this set of modeling standards, close communication between researchers is necessary.

### 5.2 Cognitive Modeling Network

Researchers dealing with similar tasks are scattered among many diverse disciplines and research communities with few to none overlap. Amongst others, researchers developing cognitive models for reasoning can be found in

- MathPsych community (MathPsych conference) and a mailing list
- Cognitive modeling community (ICCM conference)
- Knowledge representation and reasoning community (AI conferences like IJCAI, AAAI, KR) and
- Reasoning community (with the Thinking conference and the annual Londoner Reasoning Workshop) and a mailing list

However, there often is no overlap between the individual communities. A joint effort to combine the approach is neccessary.

### 5.3 Competitions

As introduced in Section 4, competitions allow to compare different approaches and to test ideas. Additionally, the test data serves as a benchmark for future cognitive models and to aid the development of comprehensive models of unified cognition. One way is to embrace a more competitive perspective on model development. By introducing challenges on comprehensive benchmarks, models that perform best according to a predefined list of criteria (connecting strictly quantitative requirements with theoretical profoundness) are selected.

## 6 Conclusion

This paper introduced challenges and research questions the fields of cognitive science and cognitive modeling in particular need to address. In order to ensure progress in the understanding of the mind, models have to transcend the state of simulations focusing on fitting experimental data. The goal for modeling is to construct model candidates that account for prominent phenomena discovered in cognitive psychology. By comparing these models on fair grounds and extracting new phenomena from the computational formalizations that can in turn be validated or falsified on experimental data, the field can advance towards a unified model of cognition.

One aim of this paper is to make general cognitive modeling principles available to the diverse communities, to open the discussion of standards, to foster the interdisciplinary research, and to tackle one of the core problems of high-level cognition: human reasoning

## References


1. H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, Dec 1974.
2. J. R. Anderson. *How can the human mind occur in the physical universe?* Oxford University Press, New York, 2007.
3. William H Batchelder and David M Riefer. Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, 6(1):57–86, 1999.
4. R. M. J. Byrne. Suppressing valid inferences with conditionals. *Cognition*, 31:61–83, 1989.
5. R. M. J. Byrne. Mental models and counterfactual thoughts about what might have been. *Trends in cognitive sciences*, 6(10):426–431, 2002.
6. G. P. Goodwin and P. N. Johnson-Laird. Reasoning about relations. *Psychological Review*, 112:468–493, 2005.
7. G. P. Goodwin and P. N. Johnson-Laird. Transitive and pseudo-transitive inferences. *Cognition*, 108:320–352, 2008.
8. Peter D Grünwald. *The minimum description length principle*. MIT press, 2007.
9. S. Khemlani and P. N. Johnson-Laird. Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, 37(5):615–623, 2009.
10. S. Khemlani and P. N. Johnson-Laird. Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, January 2012.
11. S. Khemlani and P. N. Johnson-Laird. How people differ in syllogistic reasoning. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society*, 2016.
12. K. C. Klauer, J. Musch, and B. Naumer. On belief bias in syllogistic reasoning. *Psychological Review*, 107(4):852–884, 2000.
13. M. Knauff and P. N. Johnson-Laird. Visual imagery can impede reasoning. *Memory & Cognition*, 30:363–71, 2002.
14. I. Kotseruba, O. J. A. Gonzalez, and J. K. Tsotsos. A review of 40 years of cognitive architecture research: Focus on perception, attention, learning and applications. *arXiv preprint arXiv:1610.08602*, 2016.

15. Michael D Lee and Eric-Jan Wagenmakers. *Bayesian cognitive modeling: A practical course*. Cambridge university press, 2014.
16. M. Oaksford and N. Chater. A Rational Analysis of the Selection Task as Optimal Data Selection. *Psychological Review*, 101(4):608–631, 1994.
17. K. Oberauer. Reasoning with conditionals: A test of formal models of four theories. *Cognitive Psychology*, 53:238–283, 2006.
18. M. Ragni. *Räumliche Repräsentation, Komplexität und Deduktion: Eine kognitive Komplexitätstheorie[Spatial representation, complexity and deduction: A cognitive theory of complexity]*. PhD thesis, Albert-Ludwigs-Universität Freiburg, 2008.
19. M. Ragni and M. Knauff. A theory and a computational model of spatial reasoning with preferred mental models. *Psychological Review*, 120(3):561–588, 2013.
20. David M Riefer and William H Batchelder. Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95(3):318–339, 1988.
21. Gideon Schwarz. Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464, 03 1978.
22. Gerhard Strube. The role of cognitive science in knowledge engineering. *Contemporary knowledge engineering and cognition*, pages 159–174, 1992.