Evaluating Cognitive Theories using MPTs

Workshop on Human Reasoning and Computational Logic
2017

Nicolas Riesterer
Cognitive Computation Lab
Albert-Ludwigs-Universität Freiburg

Overview

- Motivating cognitive modeling
- Multinomial Processing Trees (MPTs)
- MPTs for Human Reasoning
- Conclusion & Outlook

Why Modeling?

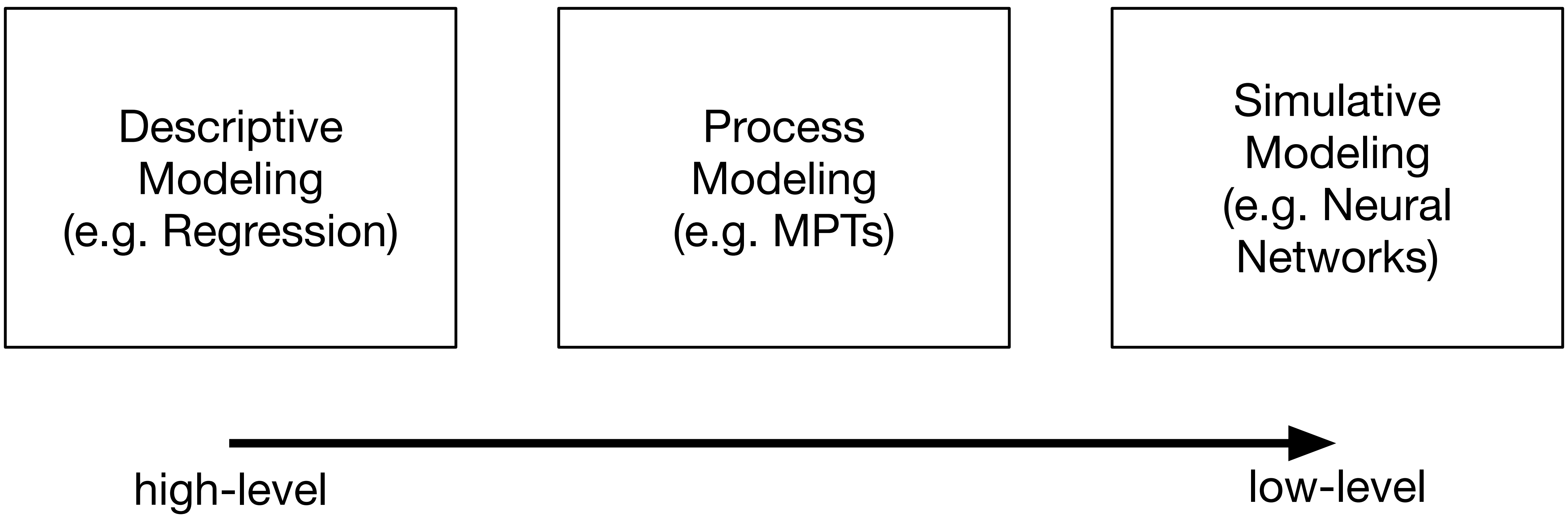
- Constrain scientific thinking
- Data can only be understood if a model is assumed
- Verbal formalisation of cognitive theories lack necessary precision
- Computational implementation requires complete and unambiguous definition
- Objective model comparison impossible without quantitative mathematical metrics

Case: Syllogistic Reasoning

Heuristic	Rule-Based	Model-Based
Atmosphere	PSYCOP	Euler Circles
Matching	Verbal Substitutions	Venn Diagrams
Illicit Conversion	Source-Founding	Verbal Models
Probability Heuristics	Monotonicity	Mental Models

Khemlani, S., & Johnson-Laird, P. N. (2012)

Cognitive Modeling



Cognitive Modeling

- Unambiguous and comprehensive definitions of theories
- Objective mathematical evaluation
- Compare different models on fair grounds
- Challenges:
    - Unknown signal-to-noise-ratio
    - “There is no perfect model”

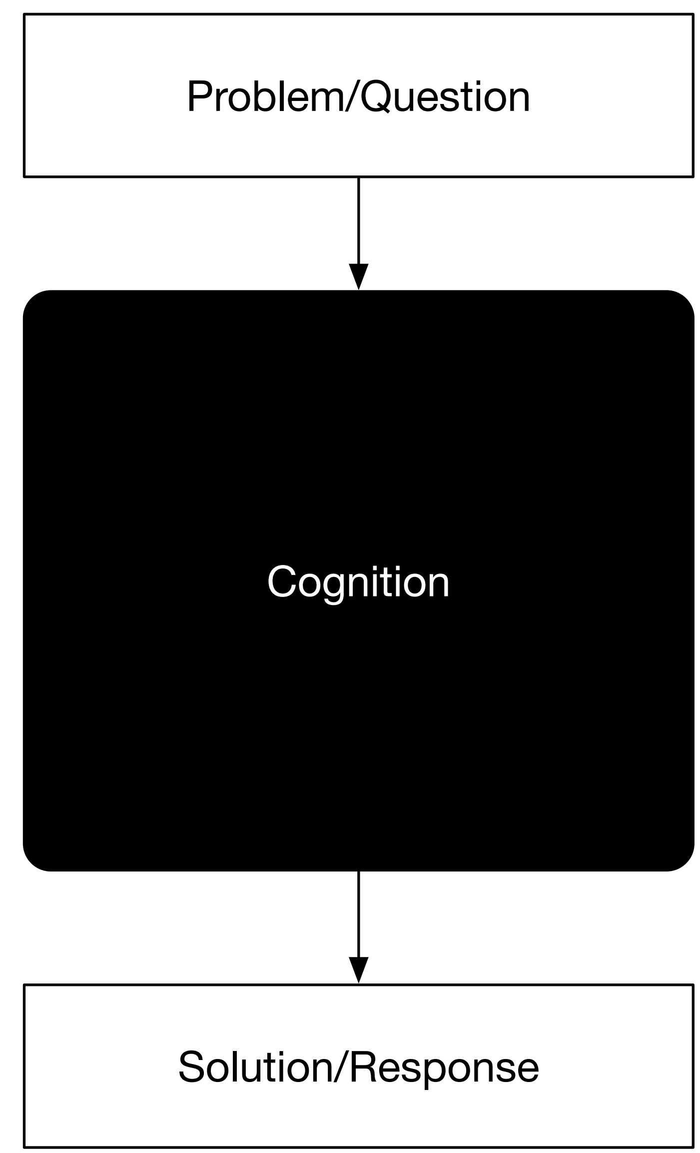
Cognitive Modeling

- No insight into workings of the mind

- Only input and output can be observed

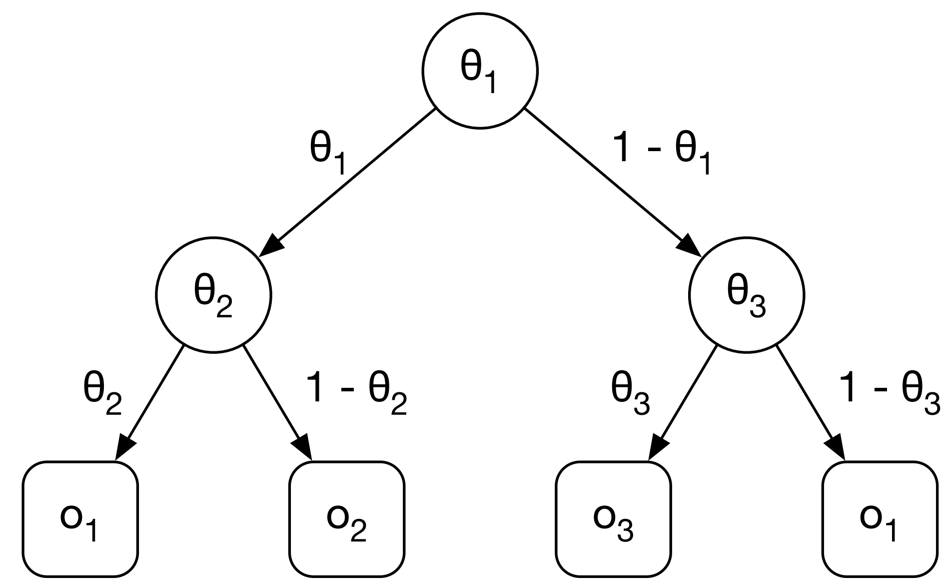
- Cognitive Models explain the processes in the black box

- Ultimately, cognitive models are just more or less informed guesses



Multinomial Processing Trees

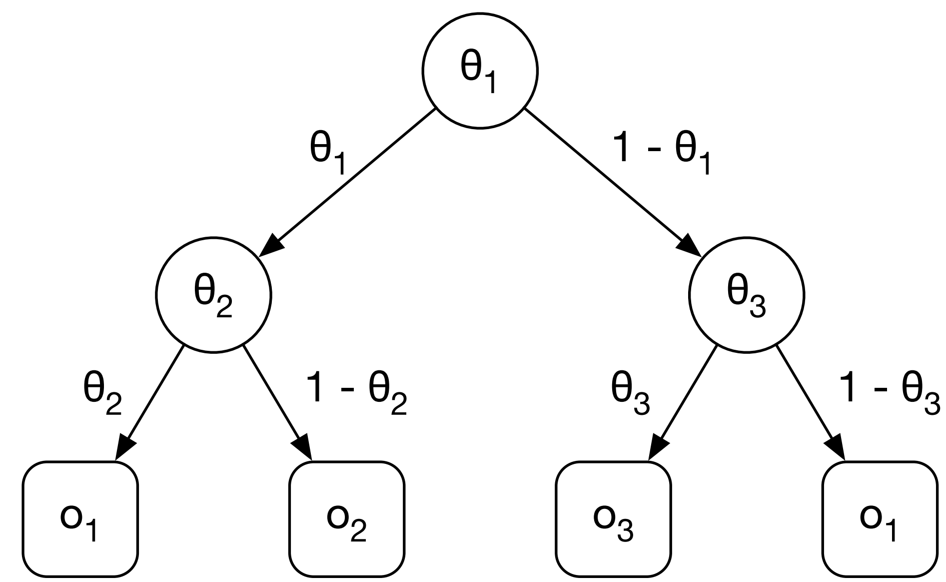
- Quantify multinomial distribution over set of outcomes
- Assume hierarchical, tree-like structure of latent variables representing cognitive processes
- No additional assumptions on parameters etc.



Generic MPT

Multinomial Processing Trees

- Directed acyclic graphs
- Nodes, leaves correspond to processes, observable outcomes
- Outgoing edges correspond to a set of possible process outcomes with assigned probabilities
- Simple (but general) case: Binary MPT, where processes can either succeed or fail

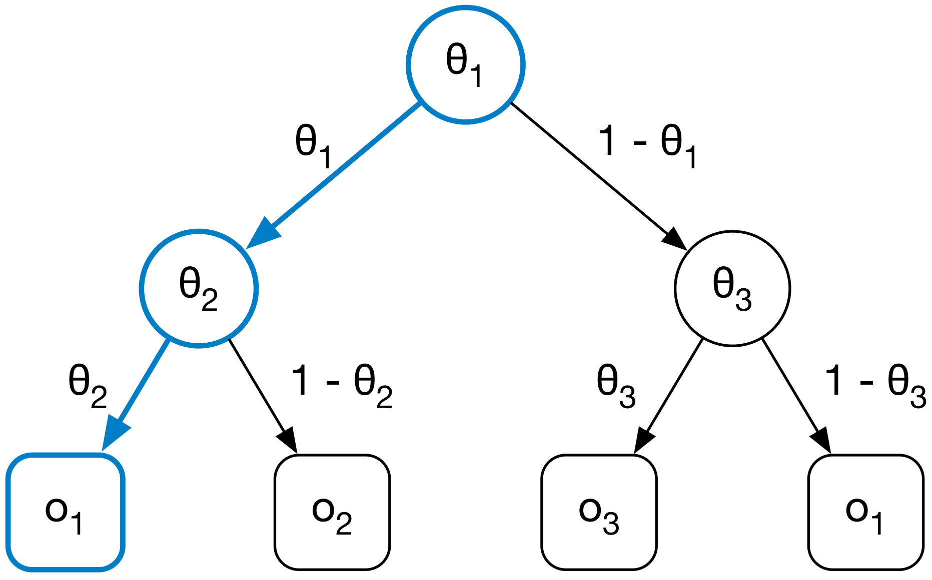


Generic MPT

Multinomial Processing Trees

- Paths (root to leaf) represent processing paths
- Edges with corresponding parameters define probability of reaching a specific outcome instance

$$P(B_{ij};\Theta) = \prod_{s=1}^{S} \theta_s^{a_{ijs}} (1-\theta_s)^{b_{ijs}}$$



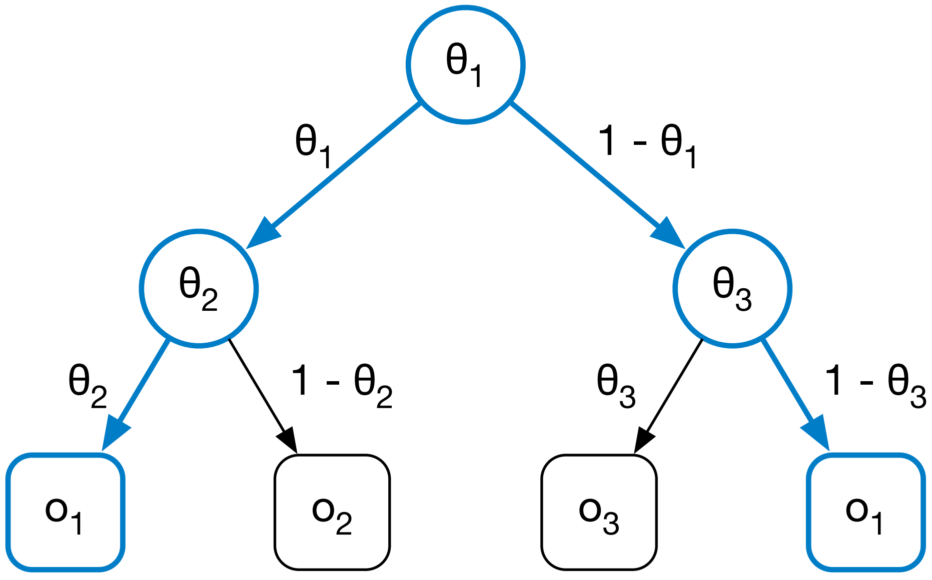
Generic MPT

Multinomial Processing Trees

- Outcome class probability defined as the sum over paths ending in its instances

- Outcome probabilities define multinomial distribution and sum up to 1

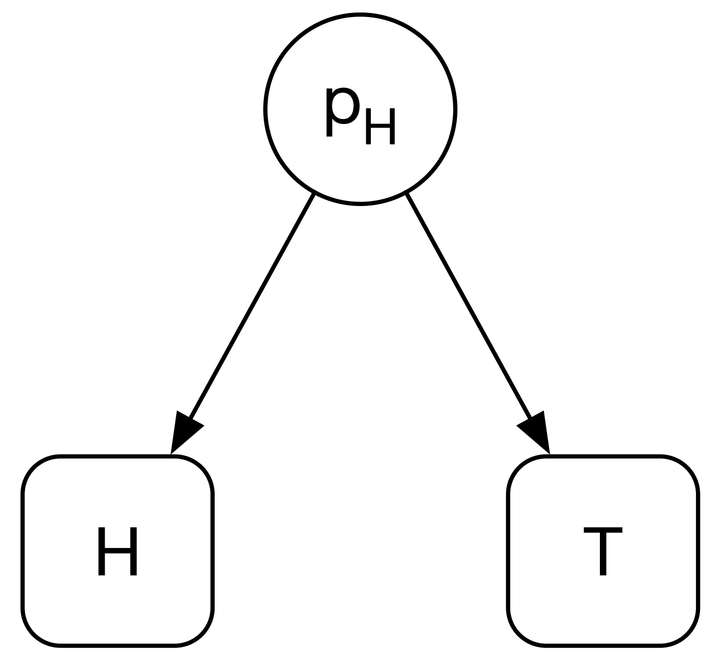
$$P(C_j;\Theta) = \sum_{i=1}^{I_j} P(B_{ij};\Theta)$$



Generic MPT

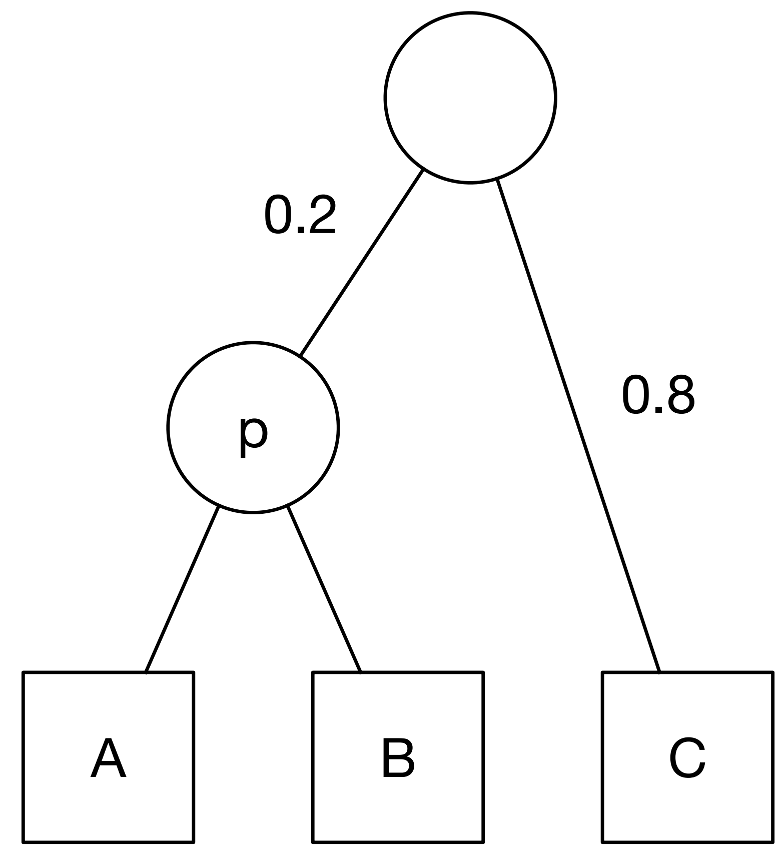
MPT Example: Coin Toss

- Two outcome classes
- One free parameter



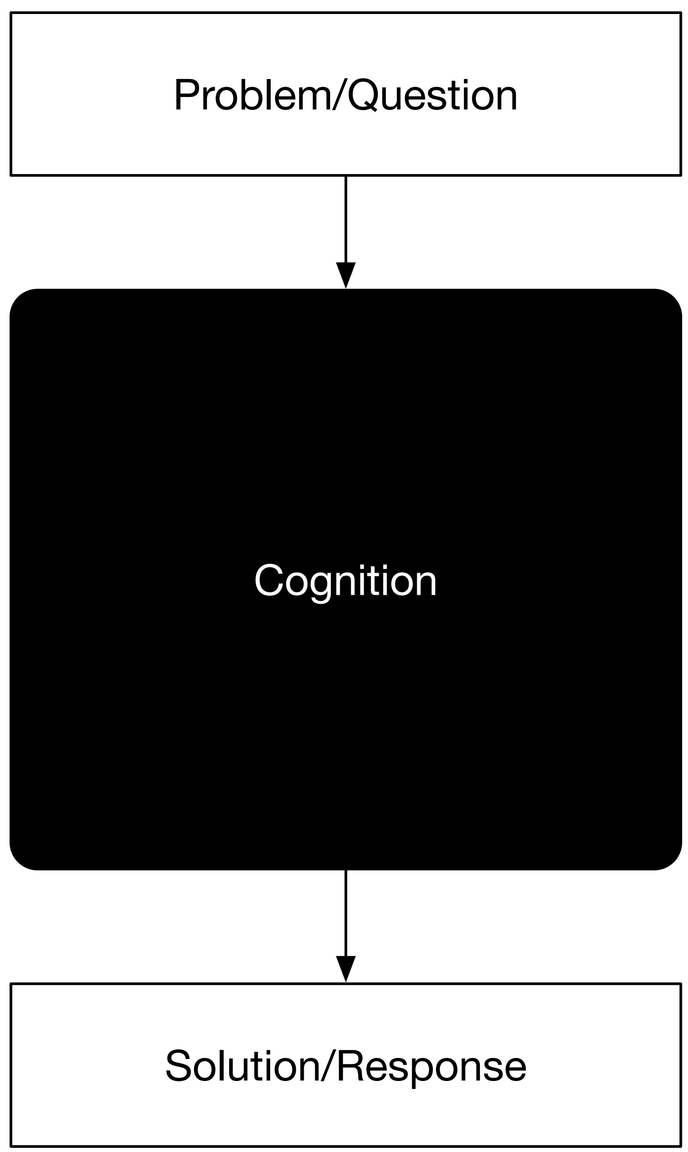
Properties

- Orders on parameters
- Orders on answers

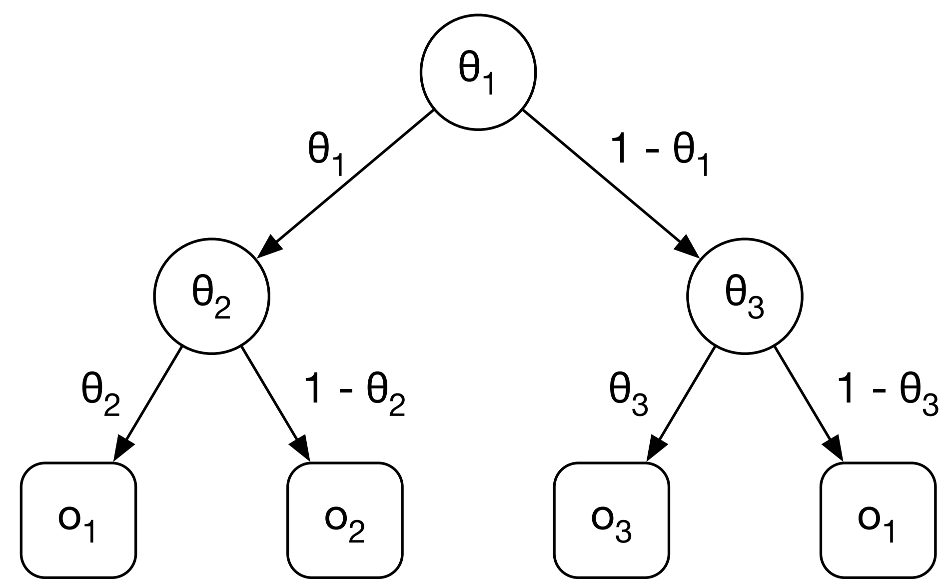


Answer ordering: C < A, C < B

MPTs for Reasoning



Tasks
E.g. “All A are B, All B are C”



Answer
E.g. “All A are C”

MPTs for Syllogistic Reasoning

- 64 syllogistic tasks usually modeled by independent MPTs
- Each tree defines distribution over 9 answers
- MPT structure based on theoretic assumptions

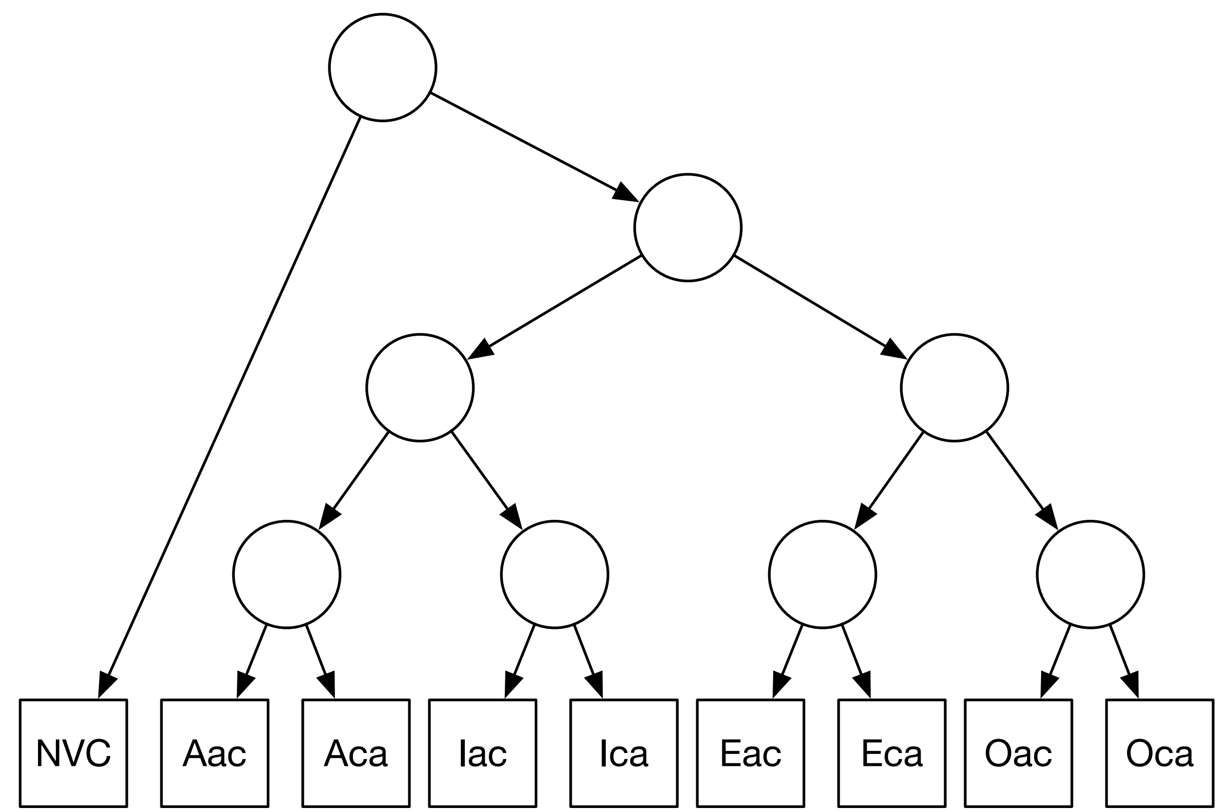
How to account for non-compliant answers in the data?

Guessing

- Not all observations explained by theory
- Dual-Process idea: reasoning based on multiple strategies
- Integration of human guessing
- Risks:
  - Guessing alters theoretic interpretability
  - Impact on fitting quality needs to be assessed carefully

Guessing

- Guessing trees define distribution over all possible answers
- Guessing choices
    - *Uniform* guessing: uniform distribution over answers
    - *Global* guessing: Same guessing strategy for all tasks
    - *Individual* guessing: Individual guessing for tasks



General guessing tree structure

Why MPTs?

- MPT model class as foundation for fair comparison of theories
- Hierarchical structure lends itself to high-level description in most theories
- Allow for modeling of cognitive processes without the need to go into neural specifics
- Only assumption is tree structure and multinomial distribution
- Rich set of properties and results

Fitting MPTs

- Fitting via Maximum Likelihood Estimation (MLE) using the Expectation-Maximization (EM) algorithm
- Straight-forward transformation into the Bayesian approach to statistics (using Markov-Chain-Monte-Carlo methods)
- Data-driven error/likelihood minimisation similar to Machine Learning techniques

Goodness of Fit

$$G^2 = 2 \sum_{j=1}^{J} O_j \ln \frac{O_j}{N \cdot p_j}$$

- Essentially the deviation from an optimal model
- Guessing influences goodness of fit
- Model complexity influences goodness of fit

“With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.”

— John von Neumann

Comparing MPTs

- Goodness of Fit
- Information Criteria
  - Akaike Information Criterion (AIC)
  - Bayesian Information Criterion (BIC)
  - Fisher Information Approximation (FIA)
  - Deviance Information Criterion (DIC)
  - Bayes Factor
- Parameter values (e.g. reasoning parameter)

Information Criteria

$$AIC = G^2 + 2K$$
$$BIC = G^2 + K \ln N$$

- Penalties for model complexity (e.g. number of parameters)
- Different approaches exist (e.g. no. parameters, Minimum-Description-Length, etc.)

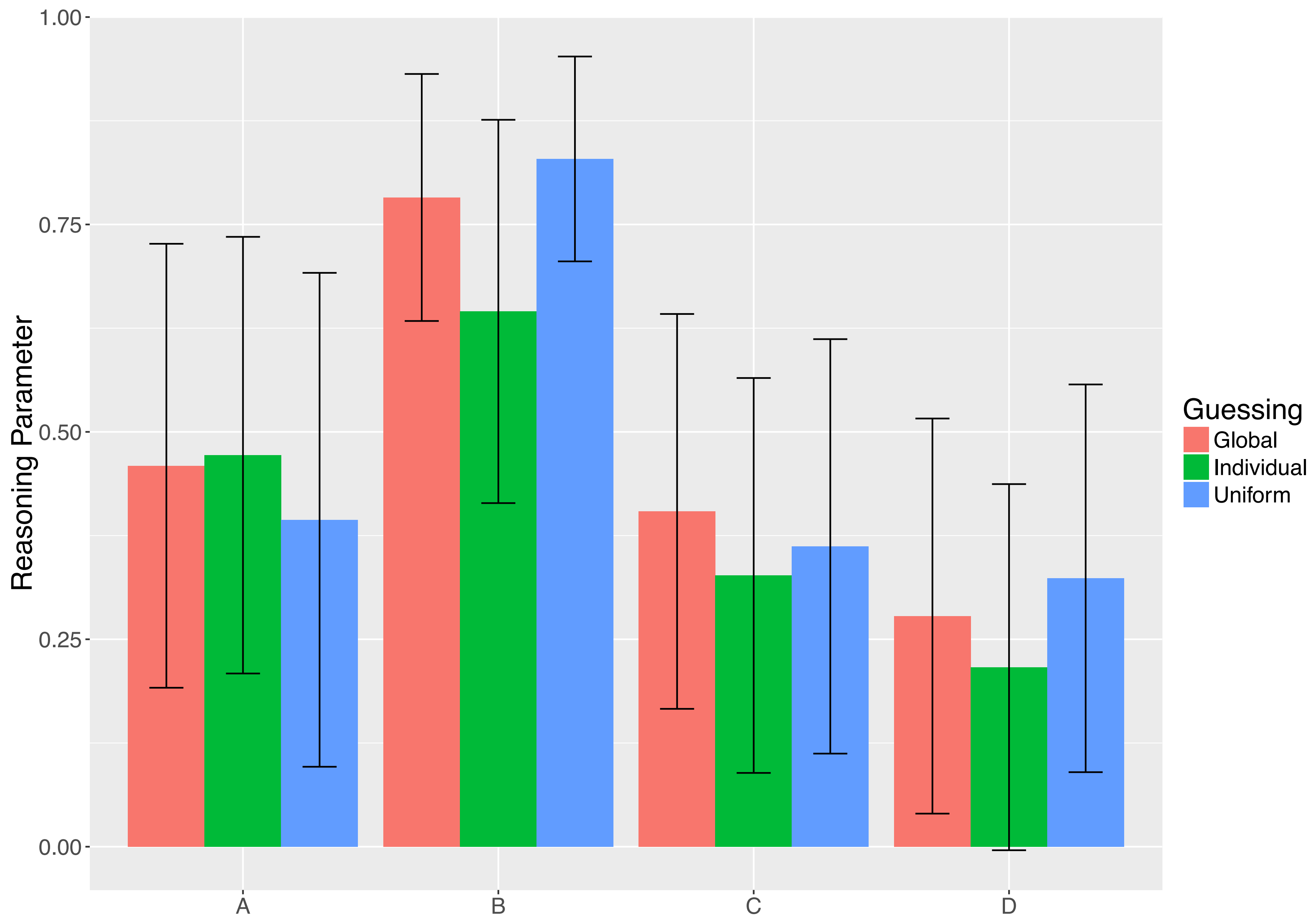
Model Selection

- Individual guessing introduces huge amount of parameters
- Guessing strategies add different amounts of free parameters (Uniform < Global < Individual)
- Individual guessing heavily penalized

Predictive Complexity

- AIC/BIC do not take a model's predictive scope into account
- A good theory should account for all of the data while occupying a minimal region of prediction space

Reasoning Parameters



Problems

- Very difficult to account for everything (Influence of guessing, over-/underfitting, consistency with the theory, etc.)
- How to handle disagreement between information criteria?
- Solution: Bayesian inferences to leave realm of frequentist approximations

Conclusion

Positive

- Data con only be understood properly using models
- Enables researchers to communicate more precisely and objectively
- MPTs lend themselves to high-level cognitive modelling

Negative

- Lots of assumptions to decide on (within model, guessing, framework)
- Information Criteria are far from being decisive
- There is no “perfect” model

Outlook

- Bring cognitive modeling closer to computer science, e.g. machine learning
- Formalisation of cognitive models using modern frameworks and fitting techniques, e.g. neural networks
- Automated optimisation and generation of cognitive theories
- Include partitioning of data into fitting procedure (training, testing, validation)

Thank you for your attention!

And don't hesitate to ask questions!