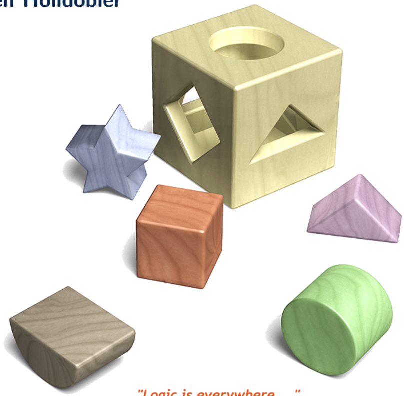
The Syllogistic Reasoning Task:
Reasoning Principles and Heuristic Strategies in Modeling Human Clusters

Richard Mörbitz, Emmanuelle Dietz, Steffen Hölldobler
International Center for Computational Logic
Technische Universität Dresden
Germany



"Logic is everywhere ..."

Modelling Human Reasoning

- Motivation: understanding human reasoning on a high level (psychology)
- Goal: provide a formalism that can model episodes of human reasoning
- In particular, predict the outcomes of psychological experiments, e.g.
  - Wason's selection task
  - Byrne's suppression task
  - Syllogistic reasoning
- Why syllogisms?
  - Well known and researched, exist since Aristotle
  - Not trivial, but also not too complex as a task
- Meta-analysis on syllogistic reasoning: Khemlani and Johnson-Laird (2012)
  - Data from six psychological studies (156 participants)
  - Comparison of twelve cognitive theories

The Syllogistic Reasoning Task

- What conclusion follows from two given premises?
  E.g. *All artists are bakers.*
  Some chemists are not bakers.
- Opinions vary:
  - Some chemists are not artists (Oca)
    - FOL, Humans, Mental Models, Verbal Models, PSYCOP
  - No valid conclusion (NVC)
    - Humans, Mental Models, Verbal Models
  - Other answers (not significant)
- FOL cannot model human syllogistic reasoning adequately
- *The* human reasoner might not exist (Oca vs. NVC)

Formalization of the Syllogistic Reasoning Task

- The premises of a syllogism are categorized by two criteria:

	affirmative	negative
$\forall$	A	E
$\exists$	I	O

Mood (type of the quantifier)

	b-c	c-b
a-b	1	3
b-a	4	2

Figure (arrangement of the terms)

- "All a are b, some c are not b" is AO3
- There are $4 \times 4 \times 4 = 64$ logically distinct premise pairs
- There are $4 \times 2 + 1 = 9$ possible answers
- Data: 64×9 matrix of answer vectors for each premise pair
- Evaluation: match of theory's prediction to participants' data

Three-valued Contextual Programs

- We use the three-valued logic defined by Łukasiewicz ($\top, \bot, \mathsf{U}$)
- We consider datalogic programs with the following types of clauses:

Rule: $A \leftarrow B_1 \wedge \cdots \wedge B_m \wedge \neg B_{m+1} \wedge \cdots \wedge \neg B_{m+n}.$

Fact: $A \leftarrow \top.$

Negative assumption: $A \leftarrow \bot.$

Unknown assumption: $A \leftarrow \mathsf{U}.$

(A and B_i are atoms)

- We denote the three-valued interpretation of a program $\mathcal{P}$ as $I = \langle I^\top, I^\bot \rangle$
- The context operator [Dietz Saldanha et al., 2017] may be applied to atoms:

$$\mathrm{ctxt}(A) = \begin{cases} \top & I(A) = \top, \\ \bot & \textit{otherwise}. \end{cases}$$

The Weak Completion Semantics

- For a logic program $\mathcal{P}$, the *weak completion*, $wc\mathcal{P}$, is obtained by
  1. For each atom A in $g\mathcal{P}$, replace all clauses of the form $A \leftarrow \text{body}_1, \ldots A \leftarrow \text{body}_n$ by $A \leftarrow \text{body}_1 \vee \cdots \vee \text{body}_n$
  2. Replace all occurrences of $\leftarrow$ by $\leftrightarrow$
- We are interested in minimal models of $wc\mathcal{P}$
- We compute them with the $\Phi_\mathcal{P}$ operator [Stenning and van Lambalgen, 2008]

  For a three-valued interpretation $I = \langle I^\top, I^\bot \rangle$, $\Phi_\mathcal{P}(I) = \langle J^\top, J^\bot \rangle$, where

$$
\begin{array}{lcl}
J^\top & = & \{A \mid A \leftarrow Body \in g\mathcal{P} \text{ and } Body \text{ is } true \text{ under } \langle I^\top, I^\bot \rangle\} \\
J^\bot & = & \{A \mid A \leftarrow Body \in g\mathcal{P} \text{ and} \\
& & \quad Body \text{ is } false \text{ under } \langle I^\top, I^\bot \rangle \text{ for all } A \leftarrow Body \in g\mathcal{P}\}
\end{array}
$$

- The least fixed point of $\Phi_\mathcal{P}$ is a minimal model of $wc\mathcal{P}$ ($\mathcal{M}_\mathcal{P}$)

Reasoning under the WCS

- Foundations by Hölldobler and Ramli [2009]
- Application to many episodes of human reasoning:
  - Wason's selection task [Dietz et al., 2012]
  - Byrne's suppresion task [Dietz et al., 2013]
  - The Belief-bias effect [Pereira et al., 2014a,b]
  - Reasoning about conditionals [Dietz and Hölldobler, 2015, Dietz et al., 2015b]
  - Spatial reasoning [Dietz et al., 2015a]
  - Syllogistic reasoning [Dietz, 2015]
- Recent application to syllogisms [da Costa et al., 2017]
  - Introduction of principles of reasoning
  - Still assuming one human reasoner

Basic Principles of Human Reasoning

- Encoding of premises motivated by findings from literature [da Costa et al., 2017]
- Basic principles are assumed to be used by all reasoners
- Quantified statements as conditionals (all moods)
  $z(X) \leftarrow y(X).$
- Licenses for inferences (all moods)
  $z(X) \leftarrow y(X) \wedge \neg ab_{yz}(X).$
- Existential import (all moods)
  $y(o_1) \leftarrow \top.$
  $ab_{yz}(o_1) \leftarrow \bot.$
- Unknown generalization (existential moods)
  $y(o_2) \leftarrow \top.$
- No refutation by counterexample (universal moods)
  $ab_{yz}(X) \leftarrow \bot.$

Advanced Principles of Human Reasoning

- Advanced principles are not necessarily applied by all reasoners
- Search for alternative conclusions (abduction)
  - Humans might not be satisfied if they come to *no valid conclusion*
  - Some may test additional conclusions
- Converse premise (moods I and E)
  - The premises *some a are b* and *some b are a* are logically equivalent
  - The corresponding converse premise may be added to the encoding as well
- Deliberate generalization (mood I)
  - Unknown generalization allows inference only for one object
  - Humans might reason a bit more relaxed
- Contraposition (mood A)
  - In FOL, “a implies b” is equivalent to “not b implies not a”

Logic Program of the Syllogism AO3 (basic principles)

- All a are b:

$b(X) \leftarrow a(X) \wedge \neg ab_{ab}(X).$	conditional, licenses
$a(o_1) \leftarrow \top.$	exImport
$ab_{ab}(o_1) \leftarrow \bot.$	exImport, licenses
$ab_{ab}(X) \leftarrow \bot.$	no refutation, licenses

- Some c are not b:

$b'(X) \leftarrow c(X) \wedge \neg ab_{cnb}(X).$	conditional, licenses
$c(o_2) \leftarrow \top.$	exImport
$ab_{cnb}(o_2) \leftarrow \bot.$	exImport, licenses
$c(o_3) \leftarrow \top.$	unknownGen
$b(X) \leftarrow \neg b'(X) \wedge \neg ab_{nbb}(X).$	transformation, licenses
$ab_{nbb}(o_2) \leftarrow \bot.$	doubleNeg, licenses
$ab_{nbb}(o_3) \leftarrow \bot.$	doubleNeg, licenses

Model and Conclusions of AO3

- Conclusions are entailed from the least fixed point of $\Phi_{\mathcal{P}}$ by logic formulas
- Least fixed point of $\Phi_{\mathcal{P}_{\text{AO3,basic}}}$ (AO3 with basic principles):

$$\mathcal{M}_{\mathcal{P}_{\text{AO3,basic}}} = \langle \{a(o_1), b(o_1), c(o_2), c(o_3), b'(o_2)\}, \{ab_{ab}(o_1), ab_{ab}(o_2), ab_{ab}(o_3), ab_{cnb}(o_2), ab_{nbb}(o_2), ab_{nbb}(o_3)\}\rangle$$

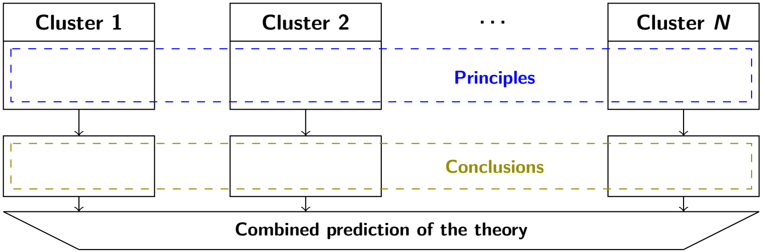
- No valid conclusion (NVC) is entailed
- If the encoding of the contraposition principle is added, we have:

$$\mathcal{M}_{\mathcal{P}_{\text{AO3,cp}}} = \langle \{a(o_1), ab_{ab}(o_2), b(o_1), c(o_2), c(o_3), a'(o_2), b'(o_2)\}, \{a(o_2), ab_{ab}(o_1), ab_{ab}(o_3), ab_{cnb}(o_2), ab_{nba}(o_1), ab_{nba}(o_2), ab_{nba}(o_3), ab_{nbb}(o_2), ab_{nbb}(o_3), b(o_2), a'(o_1)\}\rangle$$

- From this model "Some c are not a" (Oca) follows

Clusters of Reasoners

- **Prediction for the syllogism AO3 by the WCS:**
  - ***Some c are not a*** **(Oca) if the contraposition principle is applied**
  - ***No valid conclusion*** **(NVC) if it is not applied**
- **Participants' answers: both Oca (40 %) and NVC (20 %)**
- **We model clusters that differ in the application of advanced principles:**



Heuristic Strategies

- Atmosphere [Woodworth and Sells, 1935]
  - The premises create an atmosphere depending on their mood
  - A negative atmosphere makes a positive conclusion unlikely
  - An existential atmosphere makes a universal conclusion unlikely
- Matching [Wetherick and Gilhooly, 1995]
  - Defines an order of "conservativity" on moods: $E > O = I > A$
  - More conservative $\widehat{=}$ assertion about a smaller set of entities
  - A conclusion cannot be less conservative than any of the premises
- Biased conclusions in premises with figure 1
  - Recall the arrangement of terms in figure 1: (a-b, b-c)
  - Participants almost always answer Xac and never Xca for a mood X
  - Other answers are rare, too

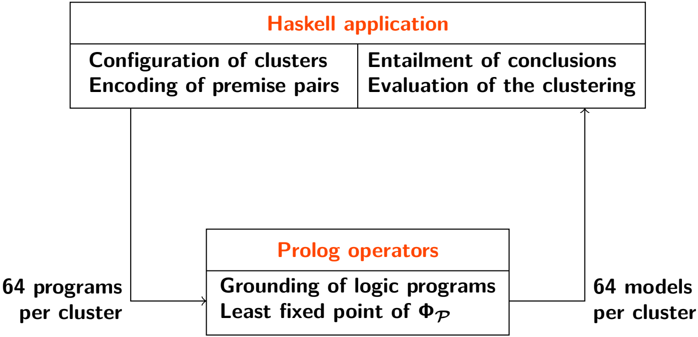
A Clustering Approach

- We model the following clusters of reasoners (in terms of used principles):
  1. Basic and converse interpretation (mood I)
  2. Basic, converse interpretation (mood I), and deliberate generalization
  3. Basic, converse interpretation (moods I and E) and contraposition
- We filter the predictions by the two heuristic strategies:
  1. Matching strategy
  2. Biased conclusions in premises with figure 1
- We predict 32 of 64 premise pairs without error

PSYCOP	Verbal Models	Mental Models	WCS (2017)	WCS (clusters)
77 %	84 %	78 % / 90 %	89 %	92 %

The WCS vs. the best cognitive theories of [Khemlani and Johnson-Laird, 2012].

Implementation



Conclusion

- The syllogistic reasoning task is an important episode of human reasoning
- FOL and well-known cognitive theories fail to model it adequately
- Our approach:
  - Identification of principles of reasoning
  - Clusters of humans in terms of used principles
  - Reasoning under the Weak Completion Semantics
  - Integration of heuristic strategies
- Future work: better clustering, quantitative predictions

Thank you for your attention. Questions?

Technical Aspects of Encoding Principles

- Negated atoms cannot be the head of a logic program clause
- The basic principles alone cannot distinguish between positive and negative moods
- Negation by transformation

  $z'(X) \leftarrow y(X) \wedge \neg ab_{nyz}(X).$
  $z(X) \leftarrow \neg z'(X) \wedge \neg ab_{nzz}(X).$

- Under the WCS, it is possible to derive knowledge through double negation

$$\mathcal{P}_{\text{doubleNeg}} = \{b \leftarrow \neg a, c \leftarrow \neg b, a \leftarrow \top\}$$
$$\mathcal{M}_{\mathcal{P}_{\text{doubleNeg}}} = \langle\{a, c\}, \{b\}\rangle$$

- No derivation through double negation

  $ab_{nzz}(o_1) \leftarrow \bot.$
  $ab_{nzz}(o_2) \leftarrow \bot.$ (if *unknown generalization* is used as well)

Entailment of Conclusions (Deduction and Abduction)

- Deduction (forward reasoning): conclusions follow from the least fixed point

  Ayz $\exists X(\mathcal{P} \models_{wcs} y(X)) \wedge \forall X(\mathcal{P} \models_{wcs} y(X) \rightarrow \mathcal{P} \models_{wcs} z(X))$

  Eyz $\exists X(\mathcal{P} \models_{wcs} y(X)) \wedge \forall X(\mathcal{P} \models_{wcs} y(X) \rightarrow \mathcal{P} \models_{wcs} \neg z(X))$

  Iyz $\exists X(\mathcal{P} \models_{wcs} y(X) \wedge z(X)) \wedge \exists X(\mathcal{P} \models_{wcs} y(X) \wedge \mathcal{P} \not\models_{wcs} z(X)) \wedge \exists X(\mathcal{P} \models_{wcs} z(X) \wedge \mathcal{P} \not\models_{wcs} y(X))$[1]

  Oyz $\exists X(\mathcal{P} \models_{wcs} y(X) \wedge \neg z(X)) \wedge \exists X(\mathcal{P} \models_{wcs} y(X) \wedge \mathcal{P} \not\models_{wcs} \neg z(X))$

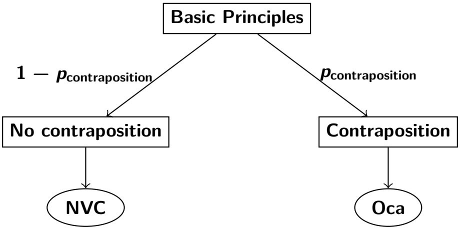
  NVC None of the above conclusions is entailed

- Abduction (backward reasoning): conclusions are computed by *abductive frameworks* that sequentially try to explain observations
  - Observations are facts that are not the head of a rule (existential import)
  - If they can be explained, new conclusions may follow as well
  - How are explanations defined? See the paper

[1] This third part of the conjunction is only used if the principle *converse interpretation* is applied.

Multinomial Processing Trees (MPT)

- Origin: Riefer and Batchelder (1988)
- Cognitive processes as inner nodes
- Response categories (here: conclusions) as leaves
- Probabilistic transition between processes (learnt parameters)



MPT for the syllogism AO3

Quantitative Evaluation

- We use the evaluation method of [Khemlani and Johnson-Laird, 2012]
- An answer is significant if it is given by more than 16 % of the participants
- Answers and predictions are encoded as vectors $\in \{0,1\}^9$
- For each premise pair, the quality of the prediction is as follows:

$$\text{score}(v_{answ}, v_{pred}) = \frac{1}{9} \times \sum_{i=1}^{9} [v_{answ}(i) = v_{pred}(i)]$$

- The overall quality of a theory is the average over all premise pairs

PSYCOP	Verbal Models	Mental Models	WCS	WCS (clusters)
77 %	84 %	78 % / 90 %	89 %	92 %

The WCS vs. the best cognitive theories of [Khemlani and Johnson-Laird, 2012].

Implementation

- Given a clustering approach, compute
  - The theory's predictions for all 64 premise pairs
  - The quantitative evaluation with respect to the participants' answers
- Component 1: implementation of $\Phi_{\mathcal{P}}$
  - Encoding of logic programs directly in Prolog
  - Prolog implementation grounds these programs and computes $\Psi_{\mathcal{P}}$
- Component 2: computing the predictions and score for a clustering
  - Representation of principles and clusters as Haskell datataypes
  - Haskell implementation generates the program for each cluster and premise pair
  - Prolog implementation is called to obtain the least fixed point
  - Conclusions are derived from that model and combined
  - Option to compute the fit to the participants' answers
  - Option to generate a MPT for each premise pair

Advanced Principles of Human Reasoning

- Advanced principles are not necessarily applied by all reasoners
- Search for alternative conclusions (abduction)
  - Humans might not be satisfied if they come to *no valid conclusion*
  - Some may test additional conclusions
- Converse premise (moods I and E)
  - The premises *some a are b* and *some b are a* are logically equivalent
  - The corresponding converse premise may be added to the encoding as well
- Deliberate generalization (mood I)
  $ab_{yz}(X) \leftarrow \text{ctxt}(z'(X)).$
  $ab_{yz}(o_2) \leftarrow \text{U}.$
- Contraposition (mood A)
  $y'(X) \leftarrow \neg z(X) \wedge \neg ab_{nzy}(X).$
  $y(X) \leftarrow \neg y'(X) \wedge \neg ab_{nyy}(X).$
  $ab_{nzy}(X) \leftarrow \bot.$

Bibliography I

Ana Oliveira da Costa, Emmanuelle-Anna Dietz Saldanha, and Steffen Hölldobler. Monadic reasoning using weak completion semantics. 2017.

Emanuelle-Anna Dietz. A computational logic approach to syllogisms in human reasoning. In *Bridging@ CADE*, pages 17–31, 2015.

Emmanuelle-Anna Dietz and Steffen Hölldobler. A new computational logic approach to reason with conditionals. In *International Conference on Logic Programming and Nonmonotonic Reasoning*, pages 265–278. Springer, 2015.

Emmanuelle-Anna Dietz, Steffen Hölldobler, and Marco Ragni. A computational logic approach to the suppression task. In *CogSci*, 2012.

Emmanuelle-Anna Dietz, Steffen Hölldobler, and Marco Ragni. A computational logic approach to the abstract and the social case of the selection task. In *Proceedings of the 11th International Symposium on Logical Formalizations of Commonsense Reasoning, COMMONSENSE*, 2013.

Emmanuelle-Anna Dietz, Steffen Hölldobler, and Raphael Höps. A computational logic approach to human spatial reasoning. In *Computational Intelligence, 2015 IEEE Symposium Series on*, pages 1627–1634. IEEE, 2015a.

Emmanuelle-Anna Dietz, Steffen Hölldobler, and Luís Moniz Pereira. On conditionals. In *GCAI*, pages 79–92, 2015b.

Bibliography II

Emmanuelle-Anna Dietz Saldanha, Steffen Hölldobler, and Luís Moniz Pereira. Contextual reasoning: Usually birds can abductively fly. 2017.

Steffen Hölldobler and Carroline Dewi Puspa Kencana Ramli. Logic programs under three-valued łukasiewicz semantics. In *International Conference on Logic Programming*, pages 464–478. Springer, 2009.

Sangeet Khemlani and P. N. Johnson-Laird. Psychological bulletin theories of the syllogism: A meta-analysis. 2012.

Luís Moniz Pereira, Emmanuelle-Anna Dietz, and Steffen Hölldobler. An abductive reasoning approach to the belief bias effect. In *KR*, 2014a.

Luis Moniz Pereira, Emmanuelle-Anna Dietz, and Steffen Hölldobler. Contextual abductive reasoning with side-effects. *Theory and Practice of Logic Programming*, 14 (4-5):633–648, 2014b.

David M Riefer and William H Batchelder. Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95(3):318–339, 1988.

Keith Stenning and Michiel van Lambalgen. Human reasoning and cognitive science. a bradford book. In *Cambridge MA*. MIT Press, 2008.

N. E. Wetherick and K. J. Gilhooly. ‘atmosphere’, matching, and logic in syllogistic reasoning. *Current Psychology*, 14(3):169–178, 1995.

R. S. Woodworth and S. B. Sells. An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, 18(4):451, 1935.