Tutorial on Cognitive Logics

Gabriele Kern-Isberner[1] and Marco Ragni[2,3] and Kai Sauerwald[4]

[1] Department of Computer Science, TU Dortmund University
[2]Danish Institute of Advanced Study, South Denmark University, Denmark
[3]Cognitive Computation Lab, University of Freiburg, Germany
[4]Knowledge Based Systems, FernUniversität in Hagen, Germany
gabriele.kern-isberner@cs.uni-dortmund.de
ragni@sdu.dk, kai.sauerwald@fernuni-hagen.de

September 13, 2020

Outline

Why are cognitive models of human thinking relevant?

- Smart devices, AI systems do (rarely) adapt to a specific users information process
    - They lack a theory of mind
- Tutorial systems rarely predict which errors you will do
- Human thinking is not yet understood, it is not transferable to systems

Talk Overview

Section 1

Some observations on the human reasoning process

Observation 1: The Wason Selection Task [Was68]

A	2	D	7
✓	✗	✗	✓

- Given:
  - Four cards with a letter on one and a number on the other side
  - A rule to check: *If there is a vowel on one side then there is an even number on the other side of the card*
- Decide:
  - Exactly which cards to turn in order to check that the rule holds?

A rule: *If a vowel is on one side then an even number is on the other side*

Observation 1': The deontic case [CG]

Again 4 cards; on one side person's age/backside drink.

If a person is drinking beer, then the person must be over 19 years of age.

Which cards must be turned to prove that the conditional holds?

	beer	coke	22yrs	16yrs
Experimental Results	95%	2.5%	2.5%	80%

- Isomorphic to the previous problem. But, most get it right!
- Observations:
  - Humans can reason classically logically, but not always
  - Even for isomorpic problems human reasoning is not equivalent

Meta-analysis of WST [RKJL18]

- Pubmed, Science Direct, or Google Scholar search with keywords: (conditional reasoning) or (selection task) or (Wason card)
- Inclusion of studies that report
  - Rules: if p, then q; every p
  - Individual selection patterns (No aggregation!)
  - At least the four canonical selections: p, pq, $p\bar{q}$, $pq\bar{q}$ per Ss
- Inclusion of 228 experiments with N = 18,000 Ss:
  - Abstract: 104 exp; Everyday: 44 exp; Deontic: 80 exp
- Aggregated results for the canonical selections in %

	p	pq	$pq\bar{q}$	$p\bar{q}$
Abstract	36	39	5	19
Everyday	23	37	11	29
Deontic	13	19	4	64

Data: `https://www.cc.uni-freiburg.de/data/`

Observation 2a: Belief Bias [EBP83]

All frenchmen drink wine
Some wine drinkers are gourmets

Some frenchmen are gourmets

Although the argument is widely accepted, it is not valid!

All frenchmen drink wine
Some wine drinkers are italians

Some frenchmen are italians

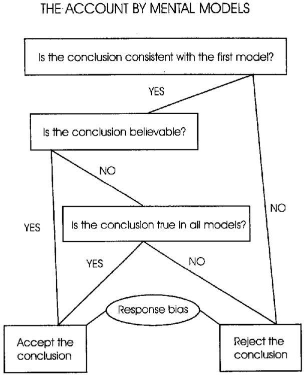
- Belief (in conclusion) Bias Effect!

Observation 2: Belief Bias – a meta-analysis

Conclusion	Syllogism	
	Believable	Unbelievable
Valid	No cigarettes are inexpensive.	No addictive things are inexpensive.
	Some addictive things are inexpensive.	Some cigarettes are inexpensive.
	Therefore, some addictive things are not cigarettes.	Therefore, some cigarettes are not addictive.
	P("valid") = 92%	P("valid") = 46%
Invalid	No addictive things are inexpensive.	No cigarettes are inexpensive.
	Some cigarettes are inexpensive.	Some addictive things are inexpensive.
	Therefore, some addictive things are not cigarettes.	Therefore, some cigarettes are not addictive.
	P("valid") = 92%	P("valid") = 8%

Example and numbers taken from [TKS+18].

Belief Bias – a meta-analysis [TKS+18]



Can be explained by

- Background knowledge
- Erroneously reasoning about consistency instead of deductive reasoning
- Humans focusing on the conclusion instead on the reasoning process

Picture from [KMN00]

- Data can be found here: `https://osf.io/8dfyv/`

Observation 2: Knowledge frame [TK83]

Linda is 31 years old, single, outspoken and very intelligent. As a student she concerned herself thoroughly with subjects of discrimination and social justice and participated in protest against nuclear energy.

Rank the following statements by their probabilities.

- Linda works as a bank teller.
- Linda works as a bank teller and is an active feminist.

- Result: More than 80% judge Linda works as a bank teller and is an active feminist to be more likely than Linda works as a bank teller.
- BUT: $p(a \wedge b) \leqslant p(a)$ or $p(b)$
- Hence, most answer falsely from the perspective of probability!
- Instead humans use the so called representativity heuristic.

Observation 3: Nonmonotonicity

- If Lisa has an essay to write, Lisa will study late in the library
- If the library is open, Lisa will study late in the library
- Lisa has an essay to write

  - 🔴 Lisa will study late in the library
  - 🟡 Nothing follows
  - 🟢 Can't say or I have another solution

The Suppression Task [Byr89]

- *If she has an essay to write, she will study late in the library.*
- *If the library is open, she will study late in the library.*
- *She has an essay to write.*

95% of all subjects conclude (modus ponens): Only 38% of all subjects conclude:

- She will study late in the library.

A logic is called non-monotonic if the set of (logical) conclusions from a knowledge base is not necessarily preserved when new information is added to the knowledge base.

- Everyday reasoning is often non-monotonic [SVL08, JL06]

Suppression Task

Facts	Conditional	Alternative Argument	Additional Argument
	If she has an essay to finish, then she will stay late in the library	If she has a textbook to read, then she will stay late in the library	If the library stays open, then she will stay late in the library
She has an essay to finish	She will study late in the library (96% L)	She will study late in the library (96% L)	She will study late in the library (38% L)
She does not have an essay to finish	She will not study late in the library (46% $\neg L$)	She will not study late in the library (4% $\neg L$)	She will not study late in the library (63% $\neg L$)

Suppression Task and classical logic

If she has an essay to finish	then she will stay late in the library	$l \leftarrow e$
If she has a textbook to read	then she will stay late in the library	$l \leftarrow t$
If the library stays open	then she will stay late in the library	$l \leftarrow o$

Clauses	Facts	Classical Logic	Exp. Findings	
$l \leftarrow e$	e	$\models l$	96% L	Modus Ponens
$l \leftarrow e \quad l \leftarrow t$	e	$\models l$	96% L	Modus Ponens
$l \leftarrow e \quad l \leftarrow o$	e	$\models l$	38% L	Modus Ponens
$l \leftarrow e$	$\neg e$	$\not\models \neg l$	46% $\neg L$	Denial of the Antecedent
$l \leftarrow e \quad l \leftarrow t$	$\neg e$	$\not\models \neg l$	4% $\neg L$	Denial of the Antecedent
$l \leftarrow e \quad l \leftarrow o$	$\neg e$	$\not\models \neg l$	63% $\neg L$	Denial of the Antecedent

Classical logic does not adequately represent the suppression task.

For more see [DHR12].

Intermediate summary

- Instead of analyzing aggregated values, single responses provide the "real" inference process.
  $\Rightarrow$ Always look at the RAW data of an individual human
- Human reasoners generate patterns that can not be reproduced by classical logic.
- Some answer patterns have implications for other answer patterns (see, [RKJL18]).
- Three-valued approaches are required [RDKH16].

Formal inference methods

Do formal nonmonotonic inference approaches show this behavior?

- Change of perspective:
  - From: Use formal inference systems as a norm for correct human behavior ($\rightarrow$ deviations of human reasoning)
  - To: Use human "commonsense" reasoning to evaluate formal inference methods ($\rightarrow$ cognitive-adequacy of formalisms)
- There are many nonmonotonic formalisms, e.g.,
  - System P
  - System Z
  - Reiter Default Logic
  - c-Representations
  - Logic Programming with Weak Completion Semantics

$\Rightarrow$ See next section

Talk Overview

Section 2

Basics on formal inference methods

Basics of propositional logic

$\mathcal{L} = \mathcal{L}(\Sigma)$	propositional language $\mathcal{L}$ over a set of atoms Σ
$\neg, \wedge, \vee$	junctors for negation, conjunction, disjunction
$A \Rightarrow B$	$\equiv \neg A \vee B$ material implication
Ω	set of interpretations/models/possible worlds over Σ
$\omega \models A$	ω is a model of $A (\in \mathcal{L})$
$\mathit{Mod}(A)$	set of models of A
$A \models B$	iff $\mathit{Mod}(A) \subseteq \mathit{Mod}(B)$ classical deduction
$\mathit{Cn}(A)$	$= \{B \in \mathcal{L} \mid A \models B\}$ classical consequence operator

Classical inference rules

Modus ponens $$\frac{A \Rightarrow B,\ A}{B}$$

Modus tollens $$\frac{A \Rightarrow B,\ \neg B}{\neg A}$$

Monotony $$\frac{A \Rightarrow B}{A \wedge C \Rightarrow B}$$

Transitivity $$\frac{\begin{array}{c} A \Rightarrow B \\ B \Rightarrow C \end{array}}{A \Rightarrow C}$$

Classical properties/axioms: Transitivity

From $A \models B$ and $B \models C$ conclude $A \models C$

$Penguin \models Bird$	*Penguins are birds.*	
$Bird \models Animal$	*Birds are animals.*	
$Penguin \models Animal$	*Penguins are animals.*	:)

$Penguin \mid\sim Bird$	*Penguins are birds.*	
$Bird \mid\sim Fly$	*Birds can fly.*	
$Penguin \mid\sim Fly$	*Penguins can fly.*	:(

Classical properties/axioms: Monotony

From $A \models C$ conclude $A \wedge B \models C$

$Penguin \models Bird$	*Penguins are birds.*	
$Penguin \wedge Black \models Bird$	*Black penguins are birds.*	:)

$Bird \mid\sim Fly$	*Birds can fly.*	
$Bird \wedge Penguin \mid\sim Fly$	*Penguin-birds can fly.*	:(

What is nonmonotonic logic?

In nonmonotonic logics, conclusions don't behave monotonically – if information is added to the knowledge base, it might happen that previous conclusions are given up, like in the famous Tweety example:

Tweety the penguin

Birds fly, penguins are birds, but penguins don't fly

$$bird \mid\!\sim fly, penguin \wedge bird \mid\!\sim \neg fly$$

Why nonmonotonic logic?

Nonmonotonic reasoning is indispensable for applications dealing with uncertain, incomplete information and should better be termed rational commonsense reasoning:

Nonmonotonic inference . . .

. . . "*is not to add certain knowledge where there is none, but rather to guide the selection of tentatively held beliefs in the hope that fruitful investigations and good guesses will result.*"

D. McDermott & J. Doyle, *Nonmonotonic logic*, 1980

The relevance of uncertain reasoning

Many applications today use classical logic or even weaker logics[1], but ...

Certainty is a treacherous illusion!

- Crucial and popular strategies of classical logics do not hold for uncertain reasoning: Modus ponens, contraposition, transitivity/syllogism, monotony, ...
- Inconsistencies and contradictions can not be resolved.

Costly or even disastrous consequences may result from ignoring uncertainty.

[1] E.g., for business rules often production rule engines are used.

A word on Tweety and penguins

The famous Tweety example deals with the important subclass-superclass-problem, like in this (less funny) example:

Example – Cancer

Cancer patients are usually adults.
Neuroblastoma is a form of cancer.
Lena is suffering from neuroblastoma.
Lena is 1 year old.[a]

[a] Neuroblastoma occurs (basically) only in children and is here the most frequent cancer disease with solid tumors.

Tweety and penguins – intuitive example that allows immediate approvement or rejection of conclusions by active reasoners (without making them feel unhappy).

Basic strategies of (nonmonotonic) commonsense reasoning

Like in classical logic, and although Modus Ponens is invalid in general,

RULES

are the main carriers of nonmonotonic inference. However, syntax and/or semantics of rules are different from implications in classical logic.

Basically, two types of rules are used:

- Rules with default assumptions: Reiter's default logic, answer set programming, weak completion semantics
- Defeasible rules: Conditional reasoning, Poole's default logic

Defeasible rules and conditionals

Defeasible rules establish an uncertain, defeasible connection between antecedent A and consequent B of a rule and can be (logically) implemented by conditionals

$(B|A)$ – "If A then (usually, probably, plausibly ...) B"

- Conditionals encode semantical relationships (plausible inferences) between the antecedent A and the consequent B.
- Conditionals implement nonmonotonic inferences via "$(B|A)$ is accepted iff $A \mid\!\sim B$ holds".
- Conditionals occur in different shapes in many approaches (e.g., as conditional probabilities in Bayesian approaches),
- Conditionals seem to be similar to classical (material) implications "If A then (definitely) B", but are substantially different!

  Indeed, many fallacies observed when applying classical logic to uncertain domains are caused by mixing up implications and conditionals!

Conditionals and implications – example

Christmas on the northern hemisphere

- If Christmas were in summer, there would be no snow at Christmas.
  plausible, approved
- If Christmas were in summer, there would be no Christmas gifts.
  strange, why?
- If Christmas were in summer, there would be no gravitation.
  downright nonsense!

All these statements are logically true, when understood as (material) implications (because Christmas is in winter on the northern hemisphere, hence the antecedent is false!).

However, understood as conditionals, crucial differences appear!

What makes conditionals so special?

A conditional $(B|A)$ focusses on cases where the premise A is fulfilled but does not say anything about cases when A does not hold – conditionals go beyond classical logic, as they are three-valued entities.

A conditional leaves more semantical room for modelling acceptance in case its confirmation $A \wedge B$ is more plausible than its refutation $A \wedge \neg B$.

Conditional acceptance and preferential entailment $\mid\!\sim_{\prec}$ [Makinson 89]

Let $\prec$ be a (well-behaved) relation on models (expressing , e.g., plausibility via a total preorder $\preceq$).

$$(B|A) \text{ is accepted} \quad \text{iff} \quad A \mid\!\sim_{\prec} B$$

iff in the most plausible models of A (wrt $\prec$), B holds also.

$\mid\!\sim_{\prec}$ is a semantic-based nonmonotonic inference relation that is encoded by conditionals on the syntax level.

Ranking functions and conditionals

Ordinal conditional functions (OCF, ranking functions[2]) [Spohn 1988]

$\kappa : \Omega \to \mathbb{N}(\cup\{\infty\})$ (Ω set of possible worlds, $\kappa^{-1}(0) \neq \varnothing$)

$\kappa(\omega_1) < \kappa(\omega_2)$	ω_1 is more plausible than ω_2
$\kappa(\omega) = 0$	ω is maximally plausible
$\kappa(A)$	$:= \min\{\kappa(\omega) \mid \omega \models A\}$
$\mathit{Bel}(\kappa)$	$:= \{A \mid \kappa(\neg A) > 0\}$

Validating conditionals

$$\kappa \models (B|A) \text{ iff } \kappa(AB) < \kappa(A\overline{B})$$

κ accepts a conditional $(B|A)$ iff its verification AB is more plausible than its falsification $A\overline{B}$.

[2]Rankings can be understood as qualitative abstractions of probabilities

Ranking functions – example

Example (ranked flyers)

$\kappa(\omega) = 4$	$p\overline{b}\,f$
$\kappa(\omega) = 2$	$pbf \quad p\overline{b}\,\overline{f}$
$\kappa(\omega) = 1$	$pb\overline{f} \quad \overline{p}\,b\overline{f}$
$\kappa(\omega) = 0$	$\overline{p}\,bf \quad \overline{p}\,\overline{b}\,f \quad \overline{p}\,\overline{b}\,\overline{f}$

$$Bel(\kappa) = Cn(\overline{p}\,(f \vee \overline{b}\,\overline{f}\,)$$

$$\kappa(bf) = 0 < 1 = \kappa(b\overline{f}\,) \Longrightarrow \kappa \models (f|b),$$

but $\kappa(p\overline{f}\,) = 1 < 2 = \kappa(pf) \Longrightarrow \kappa \models (\overline{f}\,|p)$

(also $\kappa \models (b|p)$)

Talk Overview

Section 3

A novel framework for rational human reasoning based on conditionals and plausibility

Commonsense inference rules

From a conditional statement "If A then B",
Modus ponens and Modus tollens are logically valid inference rules:

(MP) From A, infer B

(MT) From $\neg B$, infer $\neg A$

However, people also use other inference rules in commonsense reasoning:

(AC) Affirmation of the Consequent: From B, infer A

(DA) Denial of the Antecedent: From $\neg A$, infer $\neg B$

Both (AC) and (DA) are logically invalid, but are they irrational?

Logical invalidity in the Suppression Task

In the Suppression Task [Byrne 1989], participants had to draw inferences with respect to the arguments

Suppression Task (plus Additional Argument)

"If Lisa has an essay to write, she will study late in the library."
"If the library stays open, she will study late in the library."
"Lisa has an essay to write."

Here, the majority of the participants (students without tuition in logic)

- did not apply MP (38%) nor MT (33%),
- but did apply AC (63%) and DA (54%).

This inference behaviour (no MP nor MT, but AC and DA) was deemed to be completely irrational, i.e., rationality is usually assessed according to classical logic. However, obviously, the "irrational" inference behaviour was triggered by the additional information
→ Context of reasoning tasks must be taken into account!

Sensitivity of inference behavior

Different wordings and slightly different information can change human inferences drastically –

- What do people understand from the reasoning task?
  → implicit assumptions, background knowledge
- Additional information may suggest implicitly exceptions, alternatives, strengthening etc
  → nonmonotonic reasoning
- "If . . . then"-statements often are not strict
  → conditionals

Rationality needs context!

(My) Crucial hypothesis for cognitive logics

Rationality of statements can be assessed only if context is taken into account!

My most favourite example – rational or irrational???

At BRAON 2017, one of the (famous) *Madeira Workshops on Belief Revision, Argumentation, Ontologies, and Norms* locally and generally organized by *Eduardo Fermé*, Eduardo introduced himself presenting some slides and saying:

> *I have a picture of myself on my first slide because there are no cangaroos on Madeira.*

Everyone understood, and laughed . . .

Context: Dongmo Zhang from Australia introduced himself immediately before, and instead of a picture of himself, he had a picture of a cute cangaroo on his slide.

Eduardo Fermé
University of Madeira

Belief Revision
KRR

Dongmo Zhang

WESTERN SYDNEY UNIVERSITY

Affiliation: School of Computing, Engineering and Mathematics, Western Sydney University, Australia

Area of expertise: Belief revision, reasoning about action, multi-agent systems, knowledge representation and reasoning

A picture (optional):

Conditional theory of rational reasoning

People deviate so systematically from (MP) and (MT) and apply so frequently (AC) and (DA) that cognitive logics have to find a model for this. Obviously, classical logic is not cognitively adequate for cognitive logics.
Instead, we suggest:

[Eichhorn, Kern-Isberner & Ragni AAAI-2018]

- Using a (nonmonotonic) conditional logic as normative theory to evaluate human inferences
- Result: (basically) all irrationality can be eliminated!

The aim of that paper was to devise a novel (descriptive and/or normative) theory of a generic rational reasoner that emerges from a group of people.

Generic rational reasoner

When exploring rationality, we encounter the following

Dilemma of assessing rationality

Thesis: Overall, humans reason and behave rational in the sense that they are successful survivors. However,

- not all individuals reason rationally all the times – even worse, maybe each individual reasons and behaves irrationally at least from time to time . . .
- no individual reasoner can be a norm for their own rational reasoning.

Possible solution of this dilemma: Observe groups of people and try to extract a generic reasoning behaviour by

- aggregating reasoning behaviour over the group, and
- finding a formal theory to model this generic rational reasoner

Inference patterns

Basic idea: Consider all four inference rules (MP, MT, AC, DA) together in a 4-tuple to model generic inference behaviour:

Definition

An inference pattern ϱ is a 4-tuple that for each inference rule MP, MT, AC, and DA indicates whether the rule is used (positive rule, e.g., MP) or not used (negated rule, e.g., $\neg\mathrm{MP}$) in an inference scenario.

Inference patterns – examples

- Suppression Task: (MP (38%), MT (33%), AC (63%), DA (54%)) yields the inference pattern $\varrho_{Supp} = (\neg\mathrm{MP}, \neg\mathrm{MT}, \mathrm{AC}, \mathrm{DA})$.
- Counterfactuals [Thompson & Byrne 2002]: "If the car had been out of $\underline{g}$as, then it would have $\underline{s}$talled."
  Overall inferences: (MP (78%), MT (85%), AC (41%), DA (50%)), yielding the inference pattern $\varrho_{Counter} = (\mathrm{MP}, \mathrm{MT}, \neg\mathrm{AC}, \mathrm{DA})$.
  Since DA was observed with exactly half of the participants, one might also argue for the inference pattern
  $\varrho^{alt}_{Counter} = (\mathrm{MP}, \mathrm{MT}, \neg\mathrm{AC}, \neg\mathrm{DA})$.

$\rightarrow$ Basics of nonmonotonic logics and conditionals

Remember the basics of nonomotonic logics and plausibility:

Total preorders $\preccurlyeq$ on possible worlds Ω expressing plausibility are of crucial importance both for nonmonotonic reasoning and conditionals:

$\omega_1 \preccurlyeq \omega_2$	ω_1 is deemed at least as plausible as ω_2
$A \preccurlyeq B$	iff minimal models of A are at least as plausible as all models of B
$A \mid\!\sim_{\preccurlyeq} B$	iff $AB \prec A\overline{B}$ – in the context of A, B is more plausible than $\overline{B}$; iff the conditional $(B\|A)$ is accepted
Ψ	epistemic state equipped with a total preorder $\preccurlyeq_\Psi$ (you might think of Ψ as a ranking function)
Bel (Ψ)	$= Th(\min(\preccurlyeq_\Psi))$ most plausible beliefs in Ψ

Inference patterns → conditionals → plaus. constraints

With each inference rule, we associate a nonmonotonic inference relation resp. a conditional which implies a plausibility contraint:

Rule	Inference	Conditional	Plaus. constraint
MP	$A \mid\!\sim B$	$(B\|A)$	$A\,B \prec A\,\overline{B}$
MT	$\overline{B} \mid\!\sim \overline{A}$	$(\overline{A}\|\overline{B})$	$\overline{A}\,\overline{B} \prec A\,\overline{B}$
AC	$B \mid\!\sim A$	$(A\|B)$	$AB \prec \overline{A}\,B$
DA	$\overline{A} \mid\!\sim \overline{B}$	$(\overline{B}\|\overline{A})$	$\overline{A}\,\overline{B} \prec \overline{A}\,B$

Inference patterns → conditionals → plaus. constraints (cont'd)

Negated inference rules (e.g., ¬MP) are implemented simply by negating the constraint (e.g., $A\,\overline{B} \preccurlyeq A\,B$), being implemented by weak conditionals:

Definition

A weak conditional $(\!|B|A|\!)$ is accepted if $AB \preceq A\overline{B}$.

¬Rule	Weak Conditional	Plaus. constraint
¬MP	$(\!|\overline{B}|A|\!)$	$A\,\overline{B} \preceq A\,B$
¬MT	$(\!|A|\overline{B}|\!)$	$A\,\overline{B} \preceq \overline{A}\,\overline{B}$
¬AC	$(\!|\overline{A}|B|\!)$	$\overline{A}B \preceq AB$
¬DA	$(\!|B|\overline{A}|\!)$	$\overline{A}B \preceq \overline{A}\,\overline{B}$

Rationality in terms of nonmonotonic/conditional logic

reasoning pattern ϱ $\longrightarrow$ set of plausibility constraints $\mathcal{C}(\varrho)$
$\longrightarrow$ set of (weak) conditionals Δ_ϱ

$\mathcal{C}(\varrho)$ is satisfiable

- iff there is a plausibility relation (i.e., a (total) preorder) $\preceq$ on possible worlds that satisfies all constraints in $\mathcal{C}(\varrho)$
- iff the associated set of (weak) conditionals Δ_ϱ is consistent

$\longrightarrow$ novel definition of rationality in terms of conditional consistency:

Definition

- An inference pattern $\varrho \in \mathcal{R}$ is called rational iff there is a plausibility relation $\preceq$ that satisfies $\mathcal{C}(\varrho)$.
- Otherwise, the inference pattern is irrational.

...and irrationality disappears

Only 2 out of 16 patterns are irrational:

- $(\mathrm{MP}, \neg\mathrm{MT}, \neg\mathrm{AC}, \mathrm{DA})$: $\overline{A}\,\overline{B} \prec \overline{A}B \preccurlyeq AB \prec A\overline{B} \preccurlyeq \overline{A}\,\overline{B}$ – unsatsifiable
- $(\neg\mathrm{MP}, \mathrm{MT}, \mathrm{AC}, \neg\mathrm{DA})$: $\overline{A}\,\overline{B} \prec A\overline{B} \preccurlyeq AB \prec \overline{A}B \preccurlyeq \overline{A}\,\overline{B}$ – unsatisfiable

How often do they appear in practical reasoning tasks?

In over 60 empirical studies investigated so far, hardly any irrational patterns could be found (less than 2%).

(more on this later)

Implicit assumptions and background knowledge

With the help of conditionals and nonmonotonic logics/plausibility logics as a normative theory, we are able to model human reasoning much better. Using this framework, we can also deal with the following two issues:

- What implicit assumptions are used?
  How do people understand the task?
  $\rightarrow$ beliefs;
- What (conditional) beliefs are people actually using for the task?
  $\rightarrow$ elaborating on sets of conditionals giving rise to the total preorders compatible with the respective inference pattern
  $\rightarrow$ reverse engineering human reasoning

Example Suppression Task: beliefs

$$\varrho_{Supp} = (\neg\mathrm{MP}, \neg\mathrm{MT}, \mathrm{AC}, \mathrm{DA}) \rightarrow \begin{array}{ccc} A\overline{B} & \preceq & AB \\ A\overline{B} & \preceq & \overline{A}\,\overline{B} \\ AB & \prec & \overline{A}B \\ \overline{A}\,\overline{B} & \prec & \overline{A}B \end{array}$$

$$\rightarrow \qquad A\overline{B} \preceq \left\{ \begin{array}{c} AB \\ \overline{A}\,\overline{B} \end{array} \right\} \prec \overline{A}B$$

Choosing minimal, i.e., most conservative total preorder $\preceq^{min}_{Supp}$:

$$A\overline{B} \approx^{min}_{Supp} AB \approx^{min}_{Supp} \overline{A}\,\overline{B} \prec^{min}_{Supp} \overline{A}B$$

Example Suppression Task: beliefs (cont'd)

From this, we compute the beliefs

$$Bel(\preceq_{Supp}^{min}) = Cn(A\overline{B} \vee AB \vee \overline{A}\,\overline{B}) = Cn(B \Rightarrow A).$$

Here, we have $A = e$ (essay writing), $B = l$ (studying in the library), hence

$$Bel(\preceq_{Supp}^{min}) = Cn(l \Rightarrow e), \text{not } Cn(e \Rightarrow l)!$$

This explains the rationality of the inference pattern:

Participants might have understood the given conditional information in its inverse form, and hence applied AC and DA which, in fact, amount to MP and MT for the inverse conditional.

Example counterfactuals: beliefs

Constraints for the inference pattern $\varrho_{Counter} = (\mathrm{MP}, \mathrm{MT}, \neg\mathrm{AC}, \mathrm{DA})$:

$$\{AB \prec A\overline{B}, \overline{A}\,\overline{B} \prec A\overline{B}, \overline{A}B \preccurlyeq AB, \overline{A}\,\overline{B} \prec \overline{A}B\}$$
$$\equiv \quad \overline{A}\,\overline{B} \prec \overline{A}B \preccurlyeq AB \prec A\overline{B}$$

In this example, $Bel(\varrho_{Counter}) = Cn(\overline{A}\,\overline{B})$.

→ Finding: In the counterfactual case, people believe not only that the antecedent is false[3], but also that the consequent is false!

[3] This is usually assumed in the counterfactual case

C-representations [Kern-Isberner 2001]

For reverse engineering human reasoning, we build on an alternative to system Z: $\Delta = \{(B_1|A_1), \ldots, (B_n|A_n)\}$

c-representation of Δ is defined by

$$\kappa_\Delta(\omega) = \sum_{\omega \models A_i \overline{B_i}} \kappa_i^-$$

with parameters $\kappa_1^-, \ldots, \kappa_n^- \in \mathbb{N}_0$ chosen such that

$$\kappa_\Delta \models (B_j|A_j), 1 \leqslant j \leqslant n,$$

holds, i.e.,

$$\kappa_j^- > \min_{\omega \models A_j B_j} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B_i}}} \kappa_i^- - \min_{\omega \models A_j \overline{B_j}} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B_i}}} \kappa_i^-$$

For weak conditionals, one simply has to use $\geqslant$ instead of $>$.

Background beliefs and reasoning

$\kappa_\Delta(\omega) = \sum_{\omega \models A_i \overline{B_i}} \kappa_i^-$ with parameters $\kappa_1^-, \ldots, \kappa_n^- \in \mathbb{N}_0$ chosen such that

$$\kappa_j^- \begin{array}{c} > \\ \geqslant \end{array} \min_{\omega \models A_j B_j} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B_i}}} \kappa_i^- - \min_{\omega \models A_j \overline{B_j}} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B_i}}} \kappa_i^-$$

Using c-representations of (weak) conditional belief bases Δ and their parameters κ_i^-, we can further elaborate on the background (conditional) beliefs that people (may) have used for reasoning:

- Each κ_i^- symbolizes the impact of (weak) conditional $(B_i|A_i)$ on reasoning with c-representations;
- this impact has to obey a constraint that reveals the impact of $(B_i|A_i)$ in the interaction with the other conditionals from Δ.

$\rightarrow$ Each κ_i^- whose constraint is covered by other constraints can be eliminated.

Explanation generator

With the algorithm Explanation generator [Eichhorn, Kern-Isberner, Ragni, AAAI 2018] we are able to extract most basic conditionals from inference patterns:

Algo Explanation Generator

Input: Inference pattern $\varrho \in \mathcal{R}$

Output: Knowledge base of (weak) conditionals compatible with ϱ

1. Set up Δ_ϱ with a conditional for each rule in pattern ϱ
2. Set up the system of inequalities for Δ_ϱ and simplify:
   - For each inequality that is implied by the other inequalities, remove the line from the system of inequalities and the respective conditional from Δ_ϱ to obtain a (wrt. set inclusion) minimal explaining knowledge base Δ_ϱ^{expl}.
3. Return the knowledge base Δ_ϱ^{expl}.

Reverse engineering: Suppression Task

Here we have the inference pattern $\varrho_{Supp} = (\neg\mathrm{MP}, \neg\mathrm{MT}, \mathrm{AC}, \mathrm{DA})$
$\rightarrow \Delta_{Supp} = \{\delta_1 : (\!|\bar{l}|e|\!), \delta_2 : (\!|e|\bar{l}|\!), \delta_3 : (e|l), \delta_4 : (\bar{l}|\bar{e})\}$.

Schema of c-representation:

ω	$\kappa_{\Delta_{Supp}}(\omega)$	ω	$\kappa_{\Delta_{Supp}}(\omega)$
el	κ_1^-	$\bar{e}l$	$\kappa_3^- + \kappa_4^-$
$e\bar{l}$	0	$\bar{e}\bar{l}$	κ_2^-

System of constraints:

$$\kappa_1^- \geqslant \min_{e\bar{l}}\{0\} - \min_{el}\{0\} = 0 \qquad \kappa_3^- > \min_{el}\{\kappa_1^-\} - \min_{\bar{e}l}\{\kappa_4^-\}$$
$$\kappa_2^- \geqslant \min_{e\bar{l}}\{0\} - \min_{\bar{e}\bar{l}}\{0\} = 0 \qquad \kappa_4^- > \min_{\bar{e}\bar{l}}\{\kappa_2^-\} - \min_{\bar{e}l}\{\kappa_3^-\}$$

Reverse engineering: Suppression Task (cont'd)

In the end, the only relevant constraint is

$$\kappa_3^- + \kappa_4^- > \max\{\kappa_1^-, \kappa_2^-\}, \text{i.e., minimally } \kappa_3^- > 0 \text{ or } \kappa_4^- > 0$$

→ two KBs can explain the inference pattern ϱ_{Supp}:

- $\Delta_{Supp}^{expl} = \{(e|l)\}$
  "If Lisa is in the library, then she (usually) has an essay to write"
- $\Delta'^{expl}_{Supp} = \{(\bar{l}|\overline{e})\}$
  "If Lisa does not have an essay to write, then she (usually) is not in the library"

Again: Participants might have understood the given conditional information in its inverse (contraposed) form, and then $\varrho_{Supp} = (\neg\text{MP}, \neg\text{MT}, \text{AC}, \text{DA})$ appears to be rational.

Reverse engineering: counterfactuals

$\varrho_{counter} = (\mathrm{MP}, \mathrm{MT}, \neg\mathrm{AC}, \mathrm{DA})$
$\to \Delta_{counter} = \{\delta_1 : (s|g), \delta_2 : (\overline{g}|\overline{s}), \delta_3 : (\!|\overline{g}|s|\!), \delta_4 : (\overline{s}|\overline{g})\}$

Constraints:

$$\kappa_1^- + \kappa_2^- > \kappa_3^- \geqslant 0,\ \kappa_1^- + \kappa_2^- > 0,\ \kappa_3^- \geqslant \kappa_4^-,\ \kappa_4^- > 0$$

$\to \delta_2$ and κ_2^- can be eliminated
$\to \Delta_{counter}^{expl} = \{\delta_1 : (s|g), \delta_3 : (\!|\overline{g}|s|\!), \delta_4 : (\overline{s}|\overline{g})\}$:

- δ_1 "If the car is out of gas, then (usually) it stalls."
- δ_3 "If the car stalls, then it might not be out of gas." ($\to$ other possible, more plausible causes)
- δ_4 "If the car is not out of gas, then (usually) it will not stall."($\to$ possible, but not very plausible cause because drivers usually take care of gas (implicit assumption))

Reverse engineering: counterfactuals (alternative)

Let's look at the alternative inference pattern
$\varrho_{\text{counter-alt}} = (\mathrm{MP}, \mathrm{MT}, \neg\mathrm{AC}, \neg\mathrm{DA})$

$\rightarrow \Delta_{\text{counter-alt}} = \{\delta_1 : (s|g), \delta_2 : (\overline{g}|\overline{s}), \delta_3 : (\!|\overline{g}|s|\!), \delta_4' : (\!|s|\overline{g}|\!)\}$

$\rightarrow \Delta_{counter-alt}^{expl} = \{(s|g)\}$ and $\Delta'^{expl}_{counter-alt} = \{(\overline{g}|\overline{s})\}$, and

$$Bel(\Delta^{expl}_{\text{counter-alt}}) = Cn(g \Rightarrow s)$$

$\rightarrow$ classical-logical reasoner

Inference patterns in empirical studies

Focus on 22 studies with 35 experiments [Spiegel, BSc Thesis TU Dortmund 2018] –

Only six inference patterns were ever drawn at a frequency of more than 5%. The proportion of irrational patterns is only 1.1%.

Most frequent inference patterns:

(MP, MT, AC, DA)	perc.	meaning
TTTT	33.9	"credulous reasoner"
TTFF	23.6	"the logical reasoner"
TTTF	12.1	"partly logical reasoner"
TFTF	9.2	"reasoner rejecting negations"
TFTT	5.7	"bold reasoner" (all but MT)
TFFF	5.7	"basic reasoner (only MP)

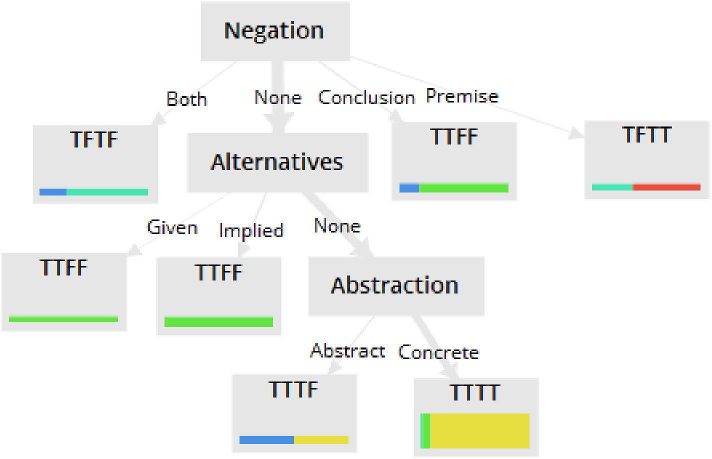
Features of tasks in empirical studies

Wordings, suggestions etc can have a major impact on human reasoning (formalized by inference patterns).

[Spiegel, GKI, Ragni, PRICAI 2019] investigated empirical studies and classified reasoning behavior ($\equiv$ inference pattern) by features that reasoning tasks may have:

Features	
age group	task type
negation	alternatives
abstraction	familiarity
meaning	(counter)factual
strictness	wording

A small decision tree



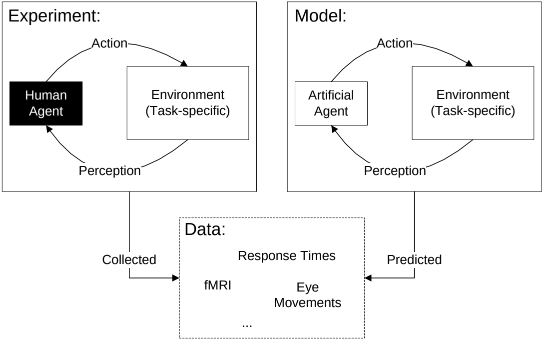
Decision tree based on three core features: negation, alternatives, abstraction

Talk Overview

Section 4

Cognitive aspects of Cognitive Logics

What does a cognitive model do?



- Reconstructive and generative models (Lüer & Spada, 1990):
  - Reconstructive: Conceptualising structures and processes that underly mental activity
  - Generative: The execution of a model not only describes psychological phenomena but also generates them
    $\Rightarrow$ Compare model predictions with empirical data

Phases of cognitive modeling

Four phases can be considered (e.g., Lewandowski & Farrell, 2011):

1. Task analysis:
   - What knowledge is needed to solve a task?
   - What are processes involved in generating the knowledge to solve a task
   - What are relevant structures an architecture used to specify a model?
2. Empirical data
   - Reconstruction of trace/statistical measure for one participant
   - Reconstruction of some statistical measure which considers all participants

Phases of cognitive modeling

3. Model implementation
   - Architecture selection (e.g. Neural Network, MPT, Logic)
   - Process specification
   - Parameter estimation (e.g. simulated annealing, maximum likelihood estimation)

4. Model validation
   - Parameter uncertainty
   - Model comparison
   - Model interpretation

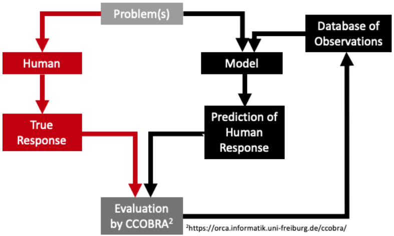
$\Rightarrow$ Mental representation ($\rightarrow$ conditionals) and the inference mechanism are core issues

How can we evaluate cognitive theories?

Simon and Wallach (1999) require a generative theories to have:

- Product correspondence: this requires that the cognitive model shows a similar overall performance as human data
- Correspondence of intermediate steps: this requires that assumed processes and steps in the model parallels separable stages in human processing
- Error correspondence: this requires that the same error patterns in the model emerge than in experimental data
- Correspondence of context dependency: this is a comparable sensitivity to known external influences

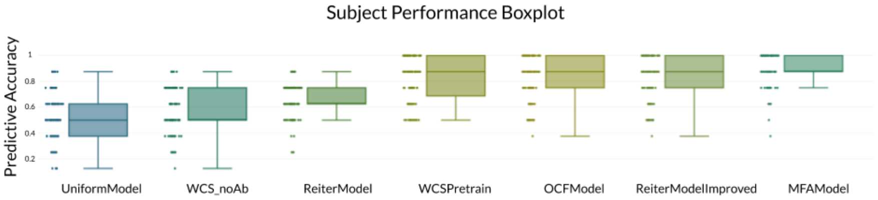
Cognitive Computation for Behavioral Reasoning Analysis (CCOBRA)



- Benchmarking tool integrating *individual* in prediction loop
- Models are evaluated based on their predictive accuracies
- CCOBRA offers pretrain, adapt, and predict methods
- Applied to syllogistic, relational, propositional reasoning [RBR20, RFB$^+$19]

`https://orca.informatik.uni-freiburg.de/ccobra`

Nonmotonic logics . . .



- Abduction in WCS is relevant
- Reiter with modus tollens and affirmation of consequence lead to ReiterModelimproved
- OCF performs identical to ReiterModelImproved

Summary

- Humans deviate from valid inferences by classical logic, but **nonmonotonic logics** are competitive.
- The extended version of Reiter's model is a functionally equivalent model to the OCF.
- Pre-trained WCS only slightly worse than Reiter Model Improved and OCF $\rightarrow$ missed **MP** predictions due to abnormalities, but, in contrast to them, successfully models **DA** by abduction.
- Decrease of predictive performance of WCS by almost 26% when not using abduction.
- Individualization relevant in all other problems relevant as well, e.g., in Wason Selection Task [RKJL18, BIMR19], etc.

Talk Overview

Section 5

Future challenges

You can make the difference!

- There exist many more reasoning problems in cognitive psychology
  - The need for a set of benchmark arises
- There are many logics and reasoning formalisms in AI
  - The need for implementations in a testable framework arises
  - and *the core point* is logics need to be made adaptive (or dynamic) that based on observations they can adapt in explain *black box processes*
- Ultimate goal: Cognitive logics are white-boxing the black-box process of individual human reasoning

Cognitive Logics Website

http://cognitive-logics.org/

Talk Overview

Section 6

References

References: Cognitive I

[BIMR19] Christian Breu, Axel Ind, Julia Mertesdorf, and Marco Ragni.
The weak completion semantics can model inferences of individual human reasoners.
In *Logics in Artificial Intelligence - 16th European Conference, JELIA 2019, Rende, Italy, May 7-11, 2019, Proceedings*, pages 498–508, 2019.

[Byr89] R.M.J. Byrne.
Suppressing valid inferences with conditionals.
Cognition, 31:61–83, 1989.

[CG] James R. Cox and Richard A. Griggs.
The effects of experience on performance in wason's selection task.
Memory & Cognition, 10(5):496–502.

References: Cognitive II

[DHR12] Emmanuelle-Anna Dietz, Steffen Hölldobler, and Marco Ragni.
A Computational Approach to the Suppression Task.
In N. Miyake, D. Peebles, and R. P. Cooper, editors, *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, pages 1500–1505, Austin, TX, 2012. Cognitive Science Society.

[EBP83] J St BT Evans, Julie L Barston, and Paul Pollard.
On the conflict between logic and belief in syllogistic reasoning.
Memory & cognition, 11(3):295–306, 1983.

[JL06] P. N. Johnson-Laird.
How we reason.
Oxford University Press, New York, 2006.

References: Cognitive III

[KMN00] K. C. Klauer, J. Musch, and B. Naumer.
On belief bias in syllogistic reasoning.
Psychological Review, 107(4):852–884, 2000.

[RBR20] Nicolas Riesterer, Daniel Brand, and Marco Ragni.
Predictive modeling of individual human cognition: Upper bounds and a new perspective on performance.
Topics in Cognitive Science, pages 1–15, 2020.

[RDKH16] Marco Ragni, Emmanuelle-Anna Dietz, Ilir Kola, and Steffen Hölldobler.
Two-valued logic is not sufficient to model human reasoning, but three-valued logic is: A formal analysis.
In *Bridging@ IJCAI*, pages 61–73, 2016.

References: Cognitive IV

[RFB+19] Marco Ragni, Paulina Friemann, Enver Bakija, Novian Habibie, Yannick Leinhos, Dennis Pohnke, Yvan Satyawan, Maya Schöchlin, and Rabea Turon.
Predicting individual spatial reasoners: A comparison of five cognitive computational theories.
pages 157–162, 2019.

[RKJL18] Marco Ragni, Ilir Kola, and P.N. Johnson-Laird.
On selecting evidence to test hypotheses: A theory of selection tasks.
Psychological Bulletin, 144(8):779–796, 2018.

[SVL08] Keith Stenning and Michiel Van Lambalgen.
Human reasoning and cognitive science.
MIT Press, 2008.

References: Cognitive V

[TK83] A. Tversky and D. Kahneman.
Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment.
Psychological Review, 90(4):293–315, 1983.

[TKS+18] Dries Trippas, David Kellen, Henrik Singmann, Gordon Pennycook, Derek J Koehler, Jonathan A Fugelsang, and Chad Dubé.
Characterizing belief bias in syllogistic reasoning: A hierarchical bayesian meta-analysis of roc data.
Psychonomic bulletin & review, 25(6):2141–2174, 2018.

[Was68] P.C. Wason.
Reasoning about a rule.
Quarterly Journal of Experimental Psychology, 20(3):273–281, 1968.