

# Tutorial on Cognitive Logics

Gabriele Kern-Isberner<sup>1</sup> and Marco Ragni<sup>2</sup>

<sup>1</sup>Department of Computer Science, Technical University Dortmund  
gabriele.kern-isberner@cs.tu-dortmund.de

<sup>2</sup>Cognitive Computation Lab, University of Freiburg  
ragni@cs.uni-freiburg.de

August 27, 2019

**tu** technische universität  
dortmund



The right logic?

## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

## Outline I

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

The right logic?

## Section 1

How human reasoning deviates from classical logic

## Basics of propositional logic

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

5 / 148

## 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  $\frac{A \Rightarrow B, B \Rightarrow C}{A \Rightarrow C}$

6 / 148

## Classical properties/axioms: Contraposition

From  $A \models B$  conclude  $\neg B \models \neg A$

$Penguin \models Bird$  Penguins are birds.  
 $\neg Bird \models \neg Penguin$  Non-birds are non-penguins. :)

$Human\_being \not\models \neg Millionaire$   
 Humans usually are not millionaires.  
 $Millionaire \not\models \neg Human\_being$   
 Millionaires usually are not human. :(

7 / 148

## 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. :)

$Penguin \not\models Bird$  Penguins are birds.  
 $Bird \not\models Fly$  Birds can fly.  
 $Penguin \not\models Fly$  Penguins can fly. :(

8 / 148

## Classical properties/axioms: Monotony

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

$$\frac{\text{Penguin} \models \text{Bird} \quad \text{Penguins are birds.}}{\text{Penguin} \wedge \text{Black} \models \text{Bird} \quad \text{Black penguins are birds.} \quad :)}{}$$

$$\frac{\text{Bird} \sim \text{Fly} \quad \text{Birds can fly.}}{\text{Bird} \wedge \text{Penguin} \sim \text{Fly} \quad \text{Penguin-birds can fly.} \quad :(}{}$$

From the common sense perspective, classical logic is inadequate, now let's have a look on the cognitive perspective ...

9 / 148

Cognitive perspective

## Section 2

Cognitive perspective

11 / 148

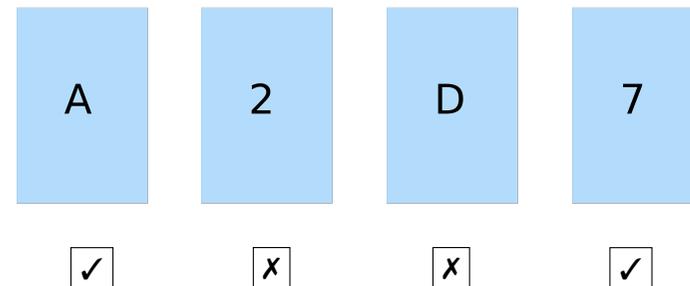
## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

10 / 148

Cognitive perspective Characteristics of human cognition

## Observation 1: The Wason Selection Task [21]



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

12 / 148

## Observation 1': The Deontic Case [3]

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 isomorphic problems human reasoning is **not** equivalent

13 / 148

## Observation 2a: Belief Bias [5]

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!**

15 / 148

## Meta-analysis of WST [17]

- 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 can be found here: [17] and <https://www.cc.uni-freiburg.de/data/14/148>

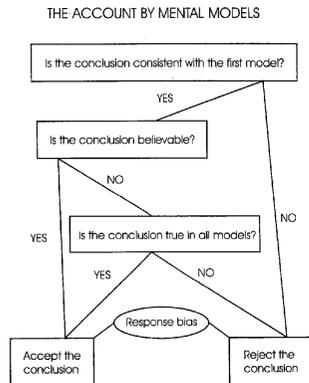
## Observation 2: Belief Bias – Meta-Analysis

Conclusion	Syllogism	
	Believable	Unbelievable
Valid	No cigarettes are inexpensive.	No addictive things are inexpensive.
	Some addictive things are inexpensive.	Some cigarettes are inexpensive.
Invalid	Therefore, some addictive things are not cigarettes.	Therefore, some cigarettes are not addictive.
	$P(\text{"valid"}) = .92$	$P(\text{"valid"}) = .46$
Invalid	No addictive things are inexpensive.	No cigarettes are inexpensive.
	Some cigarettes are inexpensive.	Some addictive things are inexpensive.
Invalid	Therefore, some addictive things are not cigarettes.	Therefore, some cigarettes are not addictive.
	$P(\text{"valid"}) = .92$	$P(\text{"valid"}) = .08$

Example and numbers taken from [19].

16 / 148

## Observation 2: Belief Bias – Meta-Analysis [19]



- 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 taken from [12].

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

17 / 148

## 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 an essay to write
  - Lisa will study late in the library
  - Nothing follows
  - Can't say or I have another solution

19 / 148

## Observation 2b: Knowledge frame [20]

Linda is 31 Jahre 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:  $p(a \wedge b) \leq p(a)$  or  $p(b)$
  - Hence, most answer falsely from the perspective of probability!
  - Instead humans use the so called **representativity heuristic**.

18 / 148

## The Suppression Task [2]

- *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 60%** 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 [18, 9]

20 / 148

## 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% <i>L</i> )	She will study late in the library (96% <i>L</i> )	She will study late in the library (38% <i>L</i> )
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% $\neg L$ )

Additional arguments lead to the suppression of previously drawn conclusions.

Alternative Arguments lead to the suppression of previously drawn conclusions.

21 / 148

## RQ 1: Are card selections (inference rules) cognitively dependent or independent?

- Some studies reported only percentage of selections for the 4 cards
- Caveat: Requires that selection of a card is independent from others
  - Some analysis report independence [4]
  - But other analysis correlations between pairs of selections [15, 13]
- Who is right ... and **how** can we test this?
- Idea:** Combine Shanon's measure of entropy with simulations of thousands of experiments
  - Entropy is a measure of unpredictability of the state (0 = certain)

$$H = - \sum p_i \log_2 p_i$$

- If  $H$ (card selections in experiment) reliably smaller as  $H$ (card selections in simulations) then card selection in experiment are dependent

23 / 148

## Suppression Task: 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% <i>L</i>	Modus Ponens
$l \leftarrow e \quad l \leftarrow t$	$e$	$\models l$	96% <i>L</i>	Modus Ponens
$l \leftarrow e \quad l \leftarrow o$	$e$	$\models l$	38% <i>L</i>	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.

22 / 148

## Entropy of the 228 experiments and 10K simulations

**Data:** Exp. with no. of Ss and frequencies of the four selections

**Result:** Proportion of experiments with lower/higher entropy

**foreach** *experiment* **do**

  Compute(*N*, *percentage*, *probs of selection for each of the 4 cards*)

  Compute(*Shanon's entropy H for the experiment*)

  Simulate(*10K experiments based on the probs of selecting each card*)

**end**

Three sorts of selection task	Mean entropy of experiments	Mean entropy of 10K simulations	Wilcoxon's W and p-value
Abstract	1.32	1.42	W = 469, p < .001
Everyday	1.51	1.66	W = 28, p < .001
Deontic	1.06	1.21	W = 68, p < .001

- Independence of card selections by Ss can be rejected!

24 / 148

## Independence assumption of theories

- Independence of card selections by Ss can be rejected!
- Theories assuming independence are built on false assumptions!
- (Eliminates 13 existing cognitive theories)

25 / 148

## Individual answer pattern matters

- Often focused on aggregated responses for each card
- Instead **each individual answer pattern** is more sensible
- Overall there can be  $2^4$  distinct answer patterns
- Conducted meta-analysis (43 Exp) for individual patterns

27 / 148

## RQ 2: Insufficiency of two-valued interpretation

Rule Number	Name	Premises	Conclusion	Logically correct?
1	MP	$p \rightarrow q, p$	$q$	Yes
2	DA	$p \rightarrow q, \bar{p}$	$\bar{q}$	No
3	AC	$p \rightarrow q, q$	$p$	No
4	MT	$p \rightarrow q, \bar{q}$	$\bar{p}$	Yes

Figure: Inference rules and abbreviations: MP: Modus Ponens, DA: Denial of Antecedent, AC: Affirmation of Consequence, MT: Modus Tollens.

26 / 148

## An interesting pattern

- How often is MP + MT + AC chosen in each study?  
→ And replicable by any two-valued valuation?

28 / 148

# Meta-Analysis of Wason Selection Task [16]

Six patterns in the meta-analysis of 46 articles. Ss = number of participants. All other values are percentages chosen by the participants

Publication	Ss	MP <i>p</i>	MP + MT <i>p, q̄</i>	MP + AC <i>p, q</i>	MP + MT + AC <i>p, q, q̄</i>	MP + AC + DA <i>p, q, p̄</i>	All <i>p, q, p̄, q̄</i>	Other
<b>Social</b>								
[3]	32	44	9	31	9	0	0	
[8]	50	6	82	2	6	0	0	
[22]	40	0	65	3	25	0	0	
[22]	40	0	45	18	8	0	0	1
[7]	60	27	17	23	10	0	0	2
[6]	25	16	16	36	12	0	0	2
Total	247	15	42	17	11	0	0	1
<b>Abstract</b>								
[10]	128	33	4	46	7	0	0	1
[14]	12	33	33	25	8	0	0	1
[23]	320	19	36	13	6	2	8	1
[8]	50	28	0	52	6	0	0	1
[1]	16	13	25	25	19	0	6	1
[11]	89	13	19	24	9	2	13	1
[18]	n/a	35	5	45	7	n/a	n/a	1
Total	615	18	13	40	7	1	2	1

# Example [16]

Table: n/a means that this is not relevant when evaluating the implication

p	q	$p \rightarrow q$	$p \rightarrow_{MP} q$	$p \rightarrow_{AC} q$
⊥	⊥	⊤	n/a	n/a
⊥	⊤	⊤	n/a	⊥
⊤	⊥	⊥	⊥	n/a
⊤	⊤	⊤	⊤	⊤

# Two-valued valuations [16]

- Each statement is mapped to 1 (true) or 0 (false)
- Idea: Search through the space of all possible valuations to explain the  $2^4$  reasoning patterns, especially if chosen
- We use valuations of the form  $p \rightarrow_{\chi} q$  where index  $\chi$  denotes the pattern that follows from the valuation of  $\rightarrow_{\chi}$
- Patterns indicate a different inference process, i.e., the 16 patterns are unique
  - E.g., MP + AC does not correspond to MP, since in the latter AC is not chosen

# Example [16]

Table: n/a means that this is not relevant when evaluating the implication

p	q	$p \rightarrow q$	$p \rightarrow_{MP} q$	$p \rightarrow_{AC} q$
⊥	⊥	⊤	n/a	n/a
⊥	⊤	⊤	n/a	⊥
⊤	⊥	⊥	⊥	n/a
⊤	⊤	⊤	⊤	⊤

## Valuations

$$v(p \rightarrow_{\text{MP}} q) = \begin{cases} 0 & \text{if } v(p) = 1 \text{ and } v(q) = 0 \\ 1 & \text{if } v(p) = v(q) = 1 \\ 0/1 & \text{otherwise} \end{cases}$$

$$v(p \rightarrow_{\text{MT}} q) = \begin{cases} 0 & \text{if } v(p) = 1 \text{ and } v(q) = 0 \\ 1 & \text{if } v(p) = v(q) = 0 \\ 0/1 & \text{otherwise} \end{cases}$$

$$v(p \rightarrow_{\text{AC}} q) = \begin{cases} 0 & \text{if } v(q) = 1 \text{ and } v(p) = 0 \\ 1 & \text{if } v(p) = v(q) = 1 \\ 0/1 & \text{otherwise} \end{cases}$$

$$v(p \rightarrow_{\text{DA}} q) = \begin{cases} 0 & \text{if } v(q) = 0 \text{ and } v(p) = 1 \\ 1 & \text{if } v(p) = v(q) = 0 \\ 0/1 & \text{otherwise} \end{cases}$$

33 / 148

## Example [16]

**Table:** The pattern MP +MT +AC has the same evaluation as the "all four" pattern, which means it also satisfies DA

p	q	MP	MT	AC	MP+MT+AC	All four
⊥	⊥	⊤/⊥	⊤	⊤/⊥	⊤	⊤
⊥	⊤	⊤/⊥	⊤/⊥	⊥	⊥	⊥
⊤	⊥	⊥	⊥	⊤/⊥	⊥	⊥
⊤	⊤	⊤	⊤/⊥	⊤	⊤	⊤

35 / 148

## Example [16]

**Table:** The pattern MP+MT+AC has the same evaluation as the "all four" pattern, which means it also satisfies DA

p	q	MP	MT	AC	MP+MT+AC	All four
⊥	⊥	⊤/⊥	⊤	⊤/⊥	⊤	⊤
⊥	⊤	⊤/⊥	⊤/⊥	⊥	⊥	⊥
⊤	⊥	⊥	⊥	⊤/⊥	⊥	⊥
⊤	⊤	⊤	⊤/⊥	⊤	⊤	⊤

34 / 148

## Example

**Table:** The pattern MP+MT+AC has the same evaluation as the "all four" pattern, which means it also satisfies DA

p	q	MP	MT	AC	MP+MT+AC	All four
⊥	⊥	⊤/⊥	⊤	⊤/⊥	⊤	⊤
⊥	⊤	⊤/⊥	⊤/⊥	⊥	⊥	⊥
⊤	⊥	⊥	⊥	⊤/⊥	⊥	⊥
⊤	⊤	⊤	⊤/⊥	⊤	⊤	⊤

36 / 148

## Consequences [16]

## Lemma

The relation between the inference rules defined in the table are:

- $\rightarrow_{MT+AC}$  holds if and only if  $\rightarrow_{MP+DA}$  holds
- $\rightarrow_{MT+AC}$  implies  $\rightarrow_{DA}$
- $\rightarrow_{MP+DA}$  implies  $\rightarrow_{AC}$
- $\rightarrow_{MT+AC+DA}$  holds if and only if  $\rightarrow_{MP+DA+AC}$  holds
- $\rightarrow_{MP+MT+AC}$  holds if and only if  $\rightarrow_{MP+MT+DA}$  holds

## Corollary

There is no two-valued valuation for the patterns  $MT + AC$ ,  $MP + DA$ ,  $MT + AC + DA$ ,  $MP + DA + AC$ ,  $MP + MT + AC$ ,  $MP + MT + DA$

37 / 148

## Truth table for three-valued logics

p	q	$\rightarrow_{MP}$	$\rightarrow_{MT}$	$\rightarrow_{MP+MT}$	$\rightarrow_{AC}$	$\rightarrow_{DA}$
0	0	0/1	1	1	0/1	1
0	1	0/1	0/1	0/1	0	0
1	0	0	0	0	0/1	0/1
1	1	1	0/1	1	1	0/1
0	u	0/1	0/1	0/1	0/1	0
u	0	0/1	0	0	0/1	0/1
u	u	0/1	0/1	0/1	0/1	0/1
u	1	0/1	0/1	0/1	0	0/1
1	u	0	0/1	0	0/1	0/1

Figure: The light grey values are the ones chosen for  $\rightarrow_{MP+MT+AC}$ , while the dark grey values are the ones chosen for  $\rightarrow_{MP+MT+DA}$ .

39 / 148

## Three-valued logics [16]

- Core idea: Assign three values (0, u, 1) to p and q.
- As we model the Wason Selection Task the valuations of the Boolean functions are mapped to the set {0,1} with 1 "turn" and 0 "not turn".
- Allows for more freedom of interpretation, and allows us to find a uniquely represent all the patterns that were missing under binary logics.
- The lemma does not hold for ternary logics, since there are at least two unique truth tables that satisfy each of the 16 reasoning patterns.

38 / 148

## 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)
- Already known: Logic Programming with weak completion semantics shows suppression effect. [Stenning and Lambalgen, 2008]
- However there are many other approaches, e.g.,
  - System P
  - System Z
  - Reiter Default Logic
  - c-Representations
  - c-Representations + Revision

$\Rightarrow$  See Section 3

40 / 148

## Intermediate summary

- Instead of analyzing aggregated values, single responses provide the “real” inference process.
- Human reasoners generate patterns that can not be reproduced by classical logic approaches.
- Some answer patterns have implications for other answer patterns.
- Three-valued logics can explain the answer results.

41 / 148

## 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

$$bird \vdash fly, penguin \wedge bird \vdash \neg fly$$

43 / 148

## Classical consequences

Consequence operator of classical logic:

$$Cn : 2^{\mathcal{L}} \rightarrow 2^{\mathcal{L}}$$

$$Cn(\mathcal{F}) := \{G \in \mathcal{L} \mid \mathcal{F} \models G\}$$

$Cn$  is **monotone**, i.e., from  $\mathcal{F} \subseteq \mathcal{G}$  we conclude  $Cn(\mathcal{F}) \subseteq Cn(\mathcal{G})$

A set of formulas  $\mathcal{F}$  is **closed (deductively)** iff  $Cn(\mathcal{F}) = \mathcal{F}$ .

Deduction theorem relates logical consequence and validity:

$$F \models G \quad \text{iff} \quad \models F \Rightarrow G$$

42 / 148

## 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

44 / 148

## The relevance of uncertain reasoning

Many applications today use classical logic or even weaker logics<sup>1</sup>, 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.

<sup>1</sup>E.g., for business rules often production rule engines are used.

## Logical consequence vs. (uncertain) human inference

Logical consequence  $\models$   $\mathcal{F} \models \mathcal{G}$  iff  $Mod(\mathcal{F}) \subseteq Mod(\mathcal{G})$

Consequence operator  $Cn$  :  $2^{\mathcal{L}} \rightarrow 2^{\mathcal{L}}$   
 $Cn(\mathcal{F}) = \{G \in \mathcal{L} \mid \mathcal{F} \models G\}$

$\mathcal{F} \models \mathcal{G}$  gdw.  $\mathcal{G} \subseteq Cn(\mathcal{F})$

Commonsense reasoning  $\vdash$  ?????

Inference operator  $C$  :  $2^{\mathcal{L}} \rightarrow 2^{\mathcal{L}}$   
 $C(\mathcal{F}) = \{G \in \mathcal{L} \mid \mathcal{F} \vdash G\}$

$\mathcal{F} \vdash \mathcal{G}$  gdw.  $\mathcal{G} \subseteq C(\mathcal{F})$

## 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.<sup>a</sup>

<sup>a</sup>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 improvement or rejection of conclusions by active reasoners (without making them feel unhappy).

## Monotony

$\mathcal{F} \subseteq \mathcal{H}$  implies  $Cn(\mathcal{F}) \subseteq Cn(\mathcal{H})$

- $Cn$  considers **all** models.
- Also **Transitivity** and **Contraposition** are based (in principle) on **Monotony**.
- Monotony does not allow to revise inferences.

Expect a defeasible inference operation  $C$  to be nonmonotonic!

## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

49 / 148

## 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

51 / 148

## Section 3

## Formal models of commonsense reasoning

50 / 148

## Reiter's default rules

Let  $\varphi$ ,  $\psi_1, \dots, \psi_n$  and  $\chi$  be (classical) formulas.

## Reiter default rule

$$\delta = \frac{\varphi : \psi_1, \dots, \psi_n}{\chi}$$

with the reading

If  $\varphi$  is known, and  $\psi_1, \dots, \psi_n$  can be consistently assumed (i.e., none of  $\neg\psi_i$  is known), then conclude  $\chi$ .

$\varphi = \text{pre}(\delta)$	Precondition
$\chi = \text{cons}(\delta)$	Consequence
$\{\psi_1, \dots, \psi_n\} = \text{just}(\delta)$	Justifications

Reiter default theory  $(W, \Delta)$ :  $W$  classical formulas,  $\Delta$  defaults.

Semantics is given by **extensions** which are minimal sets of classical formulas closed under deduction and default application.

52 / 148

## Example: Suppression task [Byrne 1989]

In an empirical study [Byrne 1989]

- ( $\alpha$ ) If she has an essay to write,  $(e \rightarrow l)$   
 then she will study late in the library and  
 ( $\gamma$ ) If the library stays open,  $(o \rightarrow l)$   
 she will study late in the library and  
 ( $\delta$ ) She has an essay to write.  $(e)$

only 38% of the participants make a modus ponens inference and conclude that: *She will study late in the library.*  
 62% concluded that: *She may or may not study late in the library.*

Can this be modelled via Reiter?

53 / 148

## Extensions via process trees – definitions [Antoniou 1997]

- $\delta$  is **applicable** to a deductively closed set  $\mathcal{F}$  iff  $pre(\delta) \in \mathcal{F}$  and  $\neg B \notin \mathcal{F}$  for every  $B \in just(\delta)$ .
- (default) process  $\Pi = (\delta_1, \dots, \delta_m)$ : finite sequence of defaults  $\delta_i \in \Delta$  with the two sets

$$\begin{aligned} In(\Pi) &= Cn(W \cup \{cons(\delta) \mid \delta \in \Pi\}) \\ Out(\Pi) &= \{\neg A \mid A \in just(\delta), \delta \in \Pi\} \end{aligned}$$

such that each  $\delta$  is applicable to the *In*-set of the foregoing defaults.

- $\Pi$  is
  - successful iff  $In(\Pi) \cap Out(\Pi) = \emptyset$ ,
  - closed iff every  $\delta \in \Delta$  that is applicable to  $In(\Pi)$  is an element of  $\Pi$ .
- $\mathcal{E}$  is an **extension** of  $(W, D)$  iff  $\mathcal{E} = In(\Pi)$  for a closed and successful process  $\Pi$ .

55 / 148

## Reiter and Suppression task [Byrne 1989; Ragni, Eichhorn, Kern-Isberner 2016]

A suitable Reiter default theory for modelling this problem would be

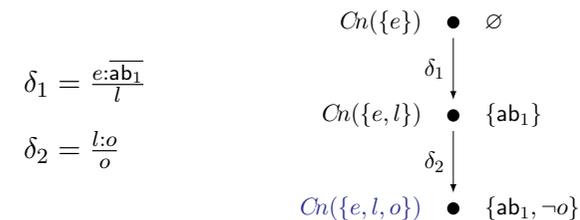
$T_{supp} = (W_{supp}, \Delta_{supp})$  with  $W_{supp} = \{e\}$  and

$$\Delta_{supp} = \left\{ \delta_1 = \frac{e : \overline{ab_1}}{l}; \delta_2 = \frac{l : o}{o}; \delta_3 = \frac{\overline{o} : ab_1}{ab_1} \right\}$$

where  $ab_1$  is an *abnormality predicate* which expresses that nothing abnormal is known.

54 / 148

## Process tree: Suppression task [Ragni, Eichhorn, Kern-Isberner 2016]



$Cn(\{e, l, o\})$  is the only extension of this theory, i.e.,  
**Modus ponens cannot be suppressed.**

56 / 148

## Inference relation for default logics

Let  $(W, \Delta)$  be a default theory.

A classical formula  $\phi$  follows nonmonotonically from  $W$  by exploiting  $\Delta$ ,

$$W \vdash_{\Delta}^{\text{Reiter}} \phi$$

if  $\phi$  is contained in all extensions of  $(W, \Delta)$ .

$$C_{\Delta}^{\text{Reiter}}(W) = \{\phi \mid W \vdash_{\Delta}^{\text{Reiter}} \phi\}$$

is the corresponding inference operator.

In the suppression task example, we have

$$e \vdash_{\Delta}^{\text{Reiter}} l,$$

so no suppression effect occurs.

57 / 148

## Logic Programming: Suppression task [Dietz et al., 2012]

Program	$l \leftarrow e \wedge \overline{ab_1}$ $l \leftarrow o \wedge \overline{ab_3}$ $ab_1 \leftarrow \overline{o}$ $ab_3 \leftarrow \overline{e}$ $e \leftarrow \top$
wcs	$l \leftrightarrow (e \wedge \overline{ab_1}) \vee (o \wedge \overline{ab_3})$ $ab_1 \leftrightarrow \overline{o}$ $ab_3 \leftrightarrow \overline{e}$ $e \leftrightarrow \top$
Least Model	$(\{e\}, \{ab_3\})$

Weak completion semantics can model the suppression effect.

59 / 148

## Logic programming

In logic programming, the (commonsense) implication “if she has an essay to finish, she will study late in the library” should be encoded by the clause

$$l \leftarrow e \wedge \overline{ab_1}$$

The so-called **weak-completion semantics** [Hölldobler and Kencana Ramli, 2009] works as follows:

- 1 Replace all clauses with the same head by a disjunction of the body elements, i.e.,  $A \leftarrow B_1, \dots, A \leftarrow B_n$  by  $A \leftarrow B_1 \vee \dots \vee B_n$ .
- 2 Replace all occurrences of  $\leftarrow$  by  $\leftrightarrow$ .

58 / 148

## 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  $A \sim B$  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!*

60 / 148

## 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!

61 / 148

## Ranking functions and conditionals

### Ordinal conditional functions (OCF, ranking functions<sup>2</sup>) [Spohn 1988]

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

$\kappa(\omega_1) < \kappa(\omega_2)$      $\omega_1$  is more plausible than  $\omega_2$

$\kappa(\omega) = 0$      $\omega$  is maximally plausible

$\kappa(A)$      $:= \min\{\kappa(\omega) \mid \omega \models A\}$

$Bel(\kappa)$      $:= \{A \mid \kappa(\neg A) > 0\}$

### Validating conditionals

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

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

<sup>2</sup>Rankings can be understood as qualitative abstractions of probabilities

63 / 148

## What makes conditionals so special?

A conditional  $(B|A)$  focuses 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 $\sim_{\prec}$ [Makinson 89]

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

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

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

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

62 / 148

## Ranking functions – example

### Example (ranked flyers)

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

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

$$\begin{aligned} \kappa(bf) = 0 < 1 = \kappa(\bar{b}\bar{f}) &\implies \kappa \models (f|b), \\ \text{but } \kappa(\bar{p}\bar{f}) = 1 < 2 = \kappa(pf) &\implies \kappa \not\models (\bar{f}|p) \\ &\text{(also } \kappa \models (b|p)) \end{aligned}$$

64 / 148

## Verification und falsification of conditionals

$\omega \in \Omega$  a possible world,  $(B|A)$  a conditional

- $\omega$  verifies  $(B|A)$  iff  $\omega \models AB$ ;
- $\omega$  falsifies  $(B|A)$  iff  $\omega \models A\bar{B}$ ;
- $\omega$  satisfies  $(B|A)$  iff  $\omega \models A \Rightarrow B$  (classical counterpart to  $(B|A)$ ).

Verification implies satisfaction.

## Conditionals and tolerance

A conditional knowledge base  $\Delta$  is consistent iff there is  $\kappa$  such that  $\kappa \models \Delta$ .

**Goal:** Finding a simple criterion to decide whether  $\Delta$  is consistent or not.

$(B|A)$  is tolerated by  $\Delta = \{(B_1|A_1), \dots, (B_n|A_n)\}$ , if there is  $\omega \in \Omega$  such that

$$\omega \models AB \wedge \bigwedge_{i=1}^n (A_i \Rightarrow B_i),$$

i.e., if there is  $\omega \in \Omega$  that verifies  $(B|A)$  and satisfies all conditionals in  $\Delta$ .

## Conditionals – example [Goldszmidt & Pearl 1996]

Conditional knowledge base  $\Delta$ :

$r_1 : (f b)$	<i>birds fly</i>
$r_2 : (b p)$	<i>penguins are birds</i>
$r_3 : (\bar{f} p)$	<i>penguins don't fly</i>
$r_4 : (w b)$	<i>birds have wings</i>
$r_5 : (a f)$	<i>animals that fly are airborne</i>

$$\omega = pb\bar{f}w\bar{a}$$

- $\omega$  verifies  $r_2, r_3, r_4$ ,
- $\omega$  falsifies  $r_1$ ,
- $\omega$  satisfies  $r_2, r_3, r_4, r_5$ .

## Tolerance – example

$$\Delta = \{r_1 : (f|b), r_2 : (b|p), r_3 : (\bar{f}|p), r_4 : (w|b), r_5 : (a|f)\}$$

$r_1$  is tolerated by  $\Delta$ :

E.g.,

$$\omega = \bar{p}bfwa \models bf \wedge (p \Rightarrow b) \wedge (p \Rightarrow \bar{f}) \wedge (b \Rightarrow w) \wedge (f \Rightarrow a)$$

Likewise,  $r_4$  and  $r_5$  are tolerated by  $\Delta$ .

However,  $r_2$  is not tolerated by  $\Delta$  because

$$pb \wedge (b \Rightarrow f) \wedge (p \Rightarrow \bar{f}) \wedge (b \Rightarrow w) \wedge (f \Rightarrow a) \equiv \perp$$

Likewise,  $r_3$  is not tolerated by  $\Delta$ .

## Consistent conditional knowledge bases

## Theorem (Adams 1975)

A conditional knowledge  $\Delta$  is consistent iff each (non-empty) subset  $\Delta' \subseteq \Delta$  contains a conditional that is tolerated by  $\Delta'$ .

This implies

## Theorem

$\Delta$  is consistent iff there is a partition  $\Delta = (\Delta_0, \Delta_1, \dots, \Delta_k)$  such that each conditional in  $\Delta_i$  is tolerated by  $\bigcup_{j=i}^k \Delta_j$ .

69 / 148

## Consistency test and partitioning – example

$$\Delta = \{r_1 : (f|b), r_2 : (b|p), r_3 : (\bar{f}|p), r_4 : (w|b), r_5 : (a|f)\}$$

$\Delta_0 = \{r_1, r_4, r_5\}$  because  $r_1, r_4, r_5$  are tolerated by all conditionals in  $\Delta$ , but  $r_2$  and  $r_3$  are not.

$r_2$  and  $r_3$  tolerate each other, hence  $\Delta_1 = \{r_2, r_3\}$ .

Therefore, we obtain

$$\Delta = (\Delta_0, \Delta_1)$$

71 / 148

## Consistency Test Algorithm [Pearl 1990]

**Input** : Conditional knowledge base  $\Delta = \{(B_1|A_1), \dots, (B_n|A_n)\}$ ;

**Output** : Partition  $\Delta = (\Delta_0, \Delta_1, \dots, \Delta_k)$  (as described above)  
iff  $\Delta$  is consistent.

- 1 Set  $i := 0$ .
- 2 **While**  $\Delta \neq \emptyset$ 
  - 1 Find the subset  $\Delta_i$  consisting of all conditionals in  $\Delta$  that are tolerated by  $\Delta$ ;
  - 2 if there is no such conditional then **Halt**:  $\Delta$  is inconsistent;
  - 3 Else set  $\Delta := \Delta - \Delta_i$ ,  $i := i + 1$ ;
- 3 **Return**  $\Delta = (\Delta_0, \Delta_1, \dots, \Delta_k)$

Complexity:  $O(n^2)$  SAT-tests in 2.1.

70 / 148

## System Z [Pearl 1990]

We use the partitioning of  $\Delta = \{r_i : (B_i|A_i)\}_{1 \leq i \leq n}$  to define a “best”  $\kappa$  that is a model of  $\Delta$ :

$$\Delta = (\Delta_0, \Delta_1, \dots, \Delta_k),$$

induces a ranking  $Z$  of the conditionals  $r_i = (B_i|A_i) \in \Delta$ :

$$Z(r_i) = j \quad \text{iff } r_i \in \Delta_j;$$

this gives us the ranking function  $\kappa^Z$  defined by

$$\kappa^Z(\omega) = \begin{cases} 0, & \text{if } \omega \text{ does not falsify any conditional in } \Delta, \\ \max_{1 \leq i \leq n} \{Z(r_i) \mid \omega \models A_i \bar{B}_i\} + 1, & \text{otherwise} \end{cases}$$

$\kappa^Z$  imposes penalty points on the worlds for falsifying conditionals.

72 / 148

## System Z (cont'd)

## Theorem (System Z)

$\kappa^z$  is a model of  $\Delta$ , i.e.,  $\kappa^z \models \Delta$ , and is minimal among all models of  $\Delta$ , i.e., for all other  $\kappa$  such that  $\kappa \models \Delta$ , there is at least one  $\omega \in \Omega$  with  $\kappa(\omega) > \kappa^z(\omega)$ .

$\kappa^z$  implements maximal plausibility among all models of  $\Delta$ .

Z-inference wrt  $\Delta$ ,  $\vdash_{\Delta}^z$ , is defined as follows:

$$A \vdash_{\Delta}^z B \quad \text{iff} \quad \kappa^z(AB) < \kappa^z(A\bar{B})$$

$\vdash_{\Delta}^z$  is one of the best existing nonmonotonic inference systems.

73 / 148

## System Z – example 2

$$\Delta = \{r_1 : (f|b), r_2 : (b|p), r_3 : (\bar{f}|p), r_4 : (w|b), r_5 : (a|f)\}$$

Partitioning:

$$\Delta_0 = \{r_1, r_4, r_5\}$$

$$\Delta_1 = \{r_2, r_3\},$$

therefore

$$Z(r_1) = Z(r_4) = Z(r_5) = 0,$$

$$Z(r_2) = Z(r_3) = 1.$$

75 / 148

## System Z – example 1

$\Delta = \{(f|b), (b|p), (\bar{f}|p)\}$  with the following partitioning:

$$\Delta_0 = \{(f|b)\}, \quad \Delta_1 = \{(b|p), (\bar{f}|p)\},$$

hence  $Z(f|b) = 0$  and  $Z(b|p) = Z(\bar{f}|p) = 1$ .

$\omega$	$\kappa^z(\omega)$	$\omega$	$\kappa^z(\omega)$
$pbf$	2	$\bar{p}bf$	0
$pb\bar{f}$	1	$\bar{p}b\bar{f}$	1
$p\bar{b}f$	2	$\bar{p}\bar{b}f$	0
$p\bar{b}\bar{f}$	2	$\bar{p}\bar{b}\bar{f}$	0

$$pb \vdash_{\Delta}^z \bar{f}$$

74 / 148

## System Z – example 2 (cont'd)

$\omega$	$\kappa^z(\omega)$	$\omega$	$\kappa^z(\omega)$	$\omega$	$\kappa^z(\omega)$	$\omega$	$\kappa^z(\omega)$
$pbfwa$	2	$pbf\bar{w}\bar{a}$	2	$pb\bar{f}\bar{w}\bar{a}$	2	$pb\bar{f}\bar{w}a$	2
$pb\bar{f}wa$	1	$pb\bar{f}w\bar{a}$	1	$pb\bar{f}w\bar{a}$	1	$pb\bar{f}w\bar{a}$	1
$p\bar{b}fwa$	2	$p\bar{b}f\bar{w}\bar{a}$	2	$p\bar{b}\bar{f}\bar{w}\bar{a}$	2	$p\bar{b}\bar{f}\bar{w}a$	2
$p\bar{b}\bar{f}wa$	2	$p\bar{b}\bar{f}w\bar{a}$	2	$p\bar{b}\bar{f}w\bar{a}$	2	$p\bar{b}\bar{f}w\bar{a}$	2
$\bar{p}bfwa$	0	$\bar{p}bf\bar{w}\bar{a}$	1	$\bar{p}b\bar{f}\bar{w}\bar{a}$	1	$\bar{p}b\bar{f}\bar{w}a$	1
$\bar{p}b\bar{f}wa$	1	$\bar{p}b\bar{f}w\bar{a}$	1	$\bar{p}b\bar{f}w\bar{a}$	1	$\bar{p}b\bar{f}w\bar{a}$	1
$\bar{p}\bar{b}fwa$	0	$\bar{p}\bar{b}f\bar{w}\bar{a}$	1	$\bar{p}\bar{b}\bar{f}\bar{w}\bar{a}$	0	$\bar{p}\bar{b}\bar{f}\bar{w}a$	1
$\bar{p}\bar{b}\bar{f}wa$	0	$\bar{p}\bar{b}\bar{f}w\bar{a}$	0	$\bar{p}\bar{b}\bar{f}w\bar{a}$	0	$\bar{p}\bar{b}\bar{f}w\bar{a}$	0

$$b \vdash_{\Delta}^z a \text{ because } \kappa^z(ba) = 0 < 1 = \kappa^z(b\bar{a}).$$

76 / 148

## Drowning problem for system Z

- $\Delta$ :
- $r_1 : (f|b)$  *Birds fly*
  - $r_2 : (b|p)$  *Penguins are birds*
  - $r_3 : (\bar{f}|p)$  *Penguins don't fly*
  - $r_4 : (w|b)$  *Birds have wings*

Do penguins (as non-typical birds) wings?

Z-partitioning:

$$\Delta_0 = \{r_1, r_4\}, \Delta_1 = \{r_2, r_3\};$$

yielding the system-Z representation  $\kappa_z$ :

## C-representations [Kern-Isberner 2001]

An alternative to system Z:  $\Delta = \{(B_1|A_1), \dots, (B_n|A_n)\}$

c-representation of  $\Delta$  is defined by

$$\kappa_\Delta(\omega) = \sum_{\omega \models A_i \bar{B}_i} \kappa_i^-$$

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

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

holds, i.e.,

$$\kappa_j^- > \min_{\omega \models A_j B_j} \sum_{\substack{i \neq j \\ \omega \models A_i \bar{B}_i}} \kappa_i^- - \min_{\omega \models A_j \bar{B}_j} \sum_{\substack{i \neq j \\ \omega \models A_i \bar{B}_i}} \kappa_i^-$$

## Drowning problem for system Z (cont'd)

$\omega$	$r_i$ fals.	$\kappa_z(\omega)$	$\omega$	$r_i$ fals.	$\kappa_z(\omega)$
$pbfw$	$r_3$	2	$\bar{p}bfw$	—	0
$pbf\bar{w}$	$r_3, r_4$	2	$\bar{p}b\bar{f}\bar{w}$	$r_4$	1
$pb\bar{f}w$	$r_1$	1	$\bar{p}b\bar{f}w$	$r_1$	1
$pb\bar{f}\bar{w}$	$r_1, r_4$	1	$\bar{p}b\bar{f}\bar{w}$	$r_1, r_4$	1
$\bar{p}\bar{b}fw$	$r_2, r_3$	2	$\bar{p}\bar{b}fw$	—	0
$\bar{p}\bar{b}\bar{f}\bar{w}$	$r_2, r_3$	2	$\bar{p}\bar{b}\bar{f}\bar{w}$	—	0
$\bar{p}\bar{b}\bar{f}w$	$r_2$	2	$\bar{p}\bar{b}\bar{f}w$	—	0
$\bar{p}\bar{b}\bar{f}\bar{w}$	$r_2$	2	$\bar{p}\bar{b}\bar{f}\bar{w}$	—	0

$\kappa_z(pw) = 1 = \kappa_z(p\bar{w})$  – we cannot decide if penguins have wings or not

→ **Drowning problem**: In  $pb\bar{f}\bar{w}$ , two conditionals with the same Z-rank are falsified, hence one of them “drowns”.

## C-representations and the Drowning problem

Considering again:

- $\Delta$ :
- $r_1 : (f|b)$  *Birds fly*
  - $r_2 : (b|p)$  *Penguins are birds*
  - $r_3 : (\bar{f}|p)$  *Penguins don't fly*
  - $r_4 : (w|b)$  *Birds have wings*

Compute (pareto-)minimal parameters  $\kappa_i^-$  for each conditional:

$$\kappa_1^- = \kappa_4^- = 1, \kappa_2^- = \kappa_3^- = 2;$$

minimal c-representation  $\kappa_\Delta$ :

## C-representations and the Drowning problem (cont'd)

$\omega$	$r_i$ fals.	$\kappa_{\Delta}(\omega)$	$\omega$	$r_i$ fals.	$\kappa_{\Delta}(\omega)$
$pbfw$	$r_3$	2	$\bar{p}bfw$	—	0
$pbf\bar{w}$	$r_3, r_4$	3	$\bar{p}b\bar{f}\bar{w}$	$r_4$	1
$p\bar{b}fw$	$r_1$	1	$\bar{p}\bar{b}fw$	$r_1$	1
$p\bar{b}\bar{f}\bar{w}$	$r_1, r_4$	2	$\bar{p}\bar{b}\bar{f}\bar{w}$	$r_1, r_4$	2
$p\bar{b}f\bar{w}$	$r_2, r_3$	4	$\bar{p}\bar{b}f\bar{w}$	—	0
$p\bar{b}\bar{f}w$	$r_2, r_3$	4	$\bar{p}\bar{b}\bar{f}w$	—	0
$p\bar{b}\bar{f}w$	$r_2$	2	$\bar{p}\bar{b}\bar{f}w$	—	0
$p\bar{b}\bar{f}\bar{w}$	$r_2$	2	$\bar{p}\bar{b}\bar{f}\bar{w}$	—	0

Here we have  $\kappa_{\Delta}(pw) = 1 < 2 = \kappa_{\Delta}(p\bar{w})$  – hence we can infer that penguins have wings.

## System Z vs. c-representations

$\omega$	$elo$	$el\bar{o}$	$e\bar{l}o$	$e\bar{l}\bar{o}$	$\bar{e}lo$	$\bar{e}l\bar{o}$	$\bar{e}\bar{l}o$	$\bar{e}\bar{l}\bar{o}$
$\kappa_1^Z(\omega)$	0	0	1	1	1	1	1	1
$\kappa_1^c(\omega)$	0	1	1	1	1	1	2	1
$\kappa_2^Z(\omega)$	0	1	1	1	1	1	1	1
$\kappa_2^c(\omega)$	0	1	2	1	1	2	1	1

## OCF inference and Suppression effect

Coming back to the Suppression effect:

$$KB_1 = \{\delta_1 : (l|e), \delta_2 : (l|o), \delta_3 : (e|\top)\}$$

$$KB_2 = \{\delta_1 : (l|e), \delta_4 : (o|l), \delta_3 : (e|\top)\}$$

$\Delta_0 = KB_i$  in each case, so  $Bel(\kappa_1^z) = Ch(el)$  and  $Bel(\kappa_2^z) = Ch(elo)$ , hence  $l$  is believed in both cases and therefore, no suppression effect occurs with system Z.

C-representations behave very similar as system Z in this example and don't show the suppression effect either.

## Mimicking weak completion semantics

In the modelling by logic programs, a strong connection is established between  $l$  and both  $o$  and  $e$ .

So, if we encode this in the knowledge base

$$\Delta_3 = \{(l|eo), (e|\top)\},$$

then we find both for system Z and c-representations that

$$Bel(\kappa_3^z) = Bel(\kappa_3^c) = Ch(elo \vee e\bar{l}\bar{o}),$$

which means that  $l$  is no longer believed!

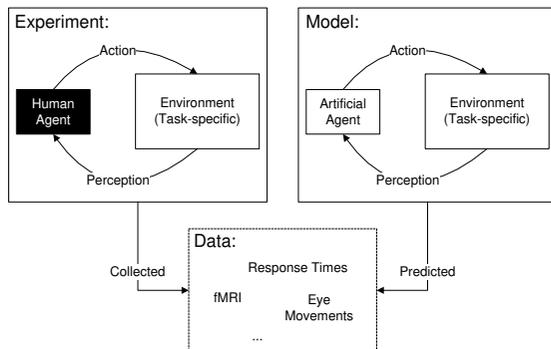
Occurrence of the suppression effect depends more on the modelling than on the chosen method!

## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

85 / 148

## What does a cognitive model?



- 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
    - ⇒ Compare model predictions with empirical data

87 / 148

## Section 4

### Cognitive aspects of Cognitive Logics

86 / 148

## 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

88 / 148

## Requirements of cognitively-adequate logics

- Explainability
- Prediction of intermediate steps
- Reverse engineering
- Implementability
- Product correspondence, i.e., same inferences
- Embedding “fallacies” in a logical context
- Alignment of cognitive and logical theories

89 / 148

## What phases of cognitive modeling exist?

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

90 / 148

## What phases of cognitive modeling exist?

### 3. Model implementation

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

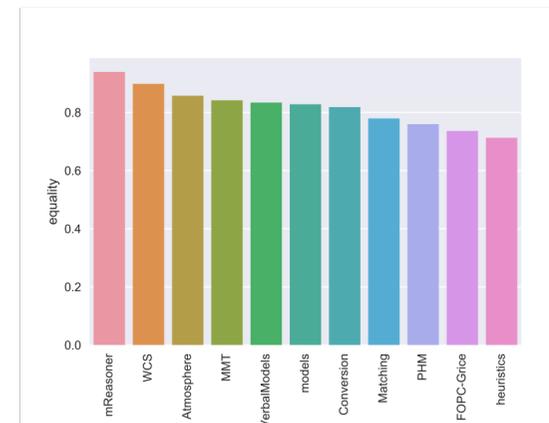
### 4. Model validation

- Parameter uncertainty
- Model comparison
- Model interpretation

⇒ Mental representation (→ conditionals) and the inference mechanism are core issues

91 / 148

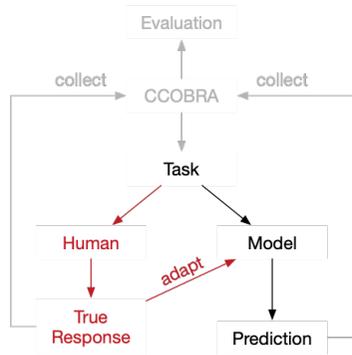
## Syllogistic Reasoning: Aggregate Data



- Existing cognitive theories do reach a high predictive power for aggregate data, i.e., predicting distribution of answers
- But, if we want to have an AI assistant that can adapt to our reasoning capabilities, does modeling “group answers” really helps us?

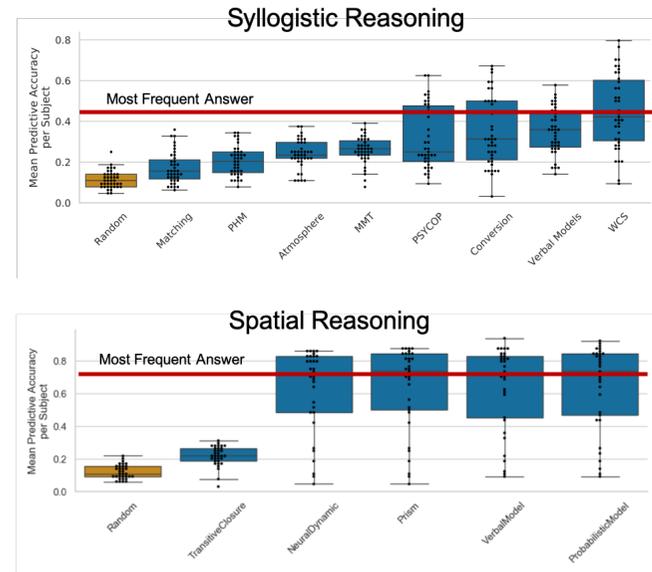
92 / 148

# Predict the Individual Reasoner



- Model receives general training problems
- Framework presents task
- Model generates predictions
- Prediction compared with true response
- Model adapts to the human response
- Framework presents next task

# Predict the Individual Reasoner



# Predict the Individual Reasoner: Propositional



# 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  $\neg B$

## Logical invalidity in the Suppression Task

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

### Suppression Task

“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%).

Applying AC and DA is usually deemed to be **irrational**, i.e., **rationality is usually assessed according to classical logic**.

97 / 148

## Inference patterns – examples

- **Suppression Task**: (MP (38%), MT (33%), AC (63%), DA (54%)) yields the inference pattern  $\varrho_{B89} = (\neg MP, \neg MT, AC, DA)$ .
- **Counterfactuals** [Thompson & Byrne 2002]: “If the car had been out of gas, then it would have stalled.”  
Overall inferences: (MP (78%), MT (85%), AC (41%), DA (50%)), yielding the inference pattern  $\varrho_{TB02} = (MP, MT, \neg AC, DA)$ .

99 / 148

## Inference patterns

However, people deviate so systematically from (MP) and (MT) and apply so frequently (AC) and (DA) that commonsense logics have to find a model for this.

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

Using a (nonmonotonic) conditional logic as normative theory to evaluate human inferences eliminates (basically) all irrationality!

**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**  $\varrho$  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.,  $\neg MP$ ) in an inference scenario.

98 / 148

## 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**

100 / 148

## → Basics of nonmonotonic logics

Remember the basics of nonmonotonic logics and plausibility:

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

- $\omega_1 \preceq \omega_2$   $\omega_1$  is deemed at least as plausible as  $\omega_2$
- $A \preceq B$  iff minimal models of  $A$  are at least as plausible as all models of  $B$
- $A \sim B$  iff  $AB \prec A\bar{B}$  – in the context of  $A$ ,  $B$  is more plausible than  $\bar{B}$
- $\Psi$  epistemic state equipped with a total preorder  $\preceq_\Psi$
- $Bel(\Psi) = Th(\min(\preceq_\Psi))$  most plausible beliefs in  $\Psi$

101 / 148

## Rationality in terms of nonmonotonic logics

reasoning pattern  $\varrho \rightarrow$  set of plausibility constraints  $\mathcal{C}(\varrho)$

$\mathcal{C}(\varrho)$  is **satisfiable** iff there is a plausibility relation  $\preceq$  on possible worlds that satisfies all constraints in  $\mathcal{C}(\varrho)$

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

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

Only 2 out of 16 patterns are irrational.

103 / 148

## Inference patterns → plausibility constraints

Rule	Inference	Plaus. constraint
MP	$A \vdash B$	$AB \prec A\bar{B}$
MT	$\bar{B} \vdash \bar{A}$	$\bar{A}\bar{B} \prec A\bar{B}$
AC	$B \vdash A$	$AB \prec \bar{A}B$
DA	$\bar{A} \vdash \bar{B}$	$\bar{A}\bar{B} \prec \bar{A}B$

Negated inference rules (e.g.,  $\neg$ MP) are implemented simply by negating the constraint (e.g.,  $A\bar{B} \preceq AB$ ).

102 / 148

## Empirical study – overview

22 studies with 35 experiments [Spiegel 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”
TTFE	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)”

(For more on this: see our talk on Thursday morning)

104 / 148

## Example counterfactuals (cont'd)

Constraints for the inference pattern  $\varrho_{TB02} = (\text{MP}, \text{MT}, \neg\text{AC}, \text{DA})$ :

$$\frac{\{AB \prec \overline{AB}, \overline{AB} \prec AB, \overline{AB} \preceq AB, \overline{AB} \prec \overline{AB}\}}{\equiv \overline{AB} \prec \overline{AB} \preceq AB \prec \overline{AB}}$$

In this example,  $\text{Bel}(\varrho_{TB02}) = \text{Ch}(\overline{AB})$ .

→ **Finding**: In the counterfactual case, people believe not only that the antecedent is false<sup>3</sup>, but also that the consequent is false!

<sup>3</sup>This is usually assumed in the counterfactual case

## Background knowledge in the Suppression task

Here we have the inference pattern

$$\varrho_{B89} = (\neg\text{MP}, \neg\text{MT}, \text{AC}, \text{DA}).$$

**Explanation generator** → two KBs can explain the inference pattern:

- $KB_{B89} = \{(e|l)\}$   
“If Lisa is in the library, then she (usually) has an essay to write”
- $KB'_{B89} = \{(\bar{l}|\bar{e})\}$   
“If Lisa does not have an essay to write, then she (usually) is not in the library”

This also 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.

## Background knowledge

Using c-representations and their parameters  $\kappa_i^-$ , we can further elaborate on the background knowledge that people (may) have used for reasoning:

With the algorithm **Explanation generator** [Eichhorn, Kern-Isberner, Ragni 2018] we're able to extract **basic conditionals** from inference patterns according to the following schema:

Rule	Conditional	Rule	Conditional
MP	$(B A)$	$\neg\text{MP}$	$(\overline{B} A)$
MT	$(\overline{A} \overline{B})$	$\neg\text{MT}$	$(A \overline{B})$
AC	$(A B)$	$\neg\text{AC}$	$(\overline{A} B)$
DA	$(\overline{B} \overline{A})$	$\neg\text{DA}$	$(B \overline{A})$

### Strong and weak conditionals for the inference rules

## Preliminary summary

Logics based on conditionals and rankings/total preorders can provide/ensure

- Explainability ✓
- Reverse engineering of human inferences ✓
- Implementability ✓
- Product correspondence, i.e., same inferences ✓
- Embedding “fallacies” in a logical context ✓
- **Future work**: Alignment of cognitive and logical theories
- **Future work**: system 1 and system 2 – prediction of intermediate steps

## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

109 / 148

## NMR and BR

Basically, belief revision deals with the problem of revising a (set or state) of beliefs  $K$  by new information  $A$  by applying a change operator  $*$ , obtaining a new (set or state) of beliefs  $K'$ :

$$K' = K * A$$

Belief revision is nonmonotonic:

- We can have  $K_1 \subseteq K_2$  but  $K_1 * A \not\subseteq K_2 * A$ ;
- We can have  $Ch(A) \subseteq Ch(B)$  but  $K * A \not\subseteq K * B$ .

Linking BR and NMR (and conditionals) via the Ramsey test

$$B \in K * A \text{ iff } A \sim_{(K)} B \text{ iff } \Psi_K \models (B|A)$$

111 / 148

## Section 5

## From nonmonotonic reasoning to belief revision

110 / 148

## The core ideas of AGM theory

The AGM postulates are recommendations for rational belief change:

- The beliefs of the agent should be deductively closed, i.e., the agent should apply logical reasoning whenever possible.
- The change operation should be successful. (This does not mean that the agent should believe everything!)
- In case of consistency, belief change should be performed via expansion.
- The result of belief change should only depend upon the semantical content of the new information.
- and more ...

112 / 148

## Rankings – also a semantics for (iterated) belief revision

## Theorem

A revision operator  $*$  satisfies the basic axioms of AGM belief revision iff there is a total preorder  $\leq_K$  (based on  $K$ ) on the set of possible worlds such that

$$\text{Mod}(K * A) = \min(\text{Mod}(A), \leq_K),$$

i.e.,

$$K * A = \text{Th}(\min(\text{Mod}(A), \leq_K))$$

Ranking functions  $\kappa$  can also be conveniently used to implement such total preorders  $\leq_K$  with  $\text{Bel}(\kappa) = K$ .

113 / 148

## Advanced belief revision for ranking functions

## Belief revision task for OCF

Given a prior OCF  $\kappa$  and some new information consisting of a set of conditionals  $\Delta = \{(B_1|A_1), \dots, (B_n|A_n)\}$ , find a posterior OCF  $\kappa^* = \kappa * \Delta$  such that  $\kappa^* \models \Delta$  and the revision complies with the core ideas of AGM.

This task involves

- iterated revision, since an epistemic state  $\kappa$  is changed;
- conditional revision, since the prior is revised by conditional information;
- multiple revision, since  $\Delta$  can be a set of plausible propositions by setting  $A \equiv (A|T)$ .

115 / 148

## Problems with AGM

- **Narrow logical framework:** Classical propositional logic, no room for uncertainty  
→ Richer epistemic frameworks?
- **One-step revision:** AGM belief revision does not consider changes of epistemic states (i.e., total preorders) nor revision strategies  
→ Iterated revision
- **New information:** Only one proposition – what about sets of propositions, conditional statements, sets of conditionals?  
→ Conditional and multiple belief revision

114 / 148

## A principle of conditional preservation for ranking functions

## OCF principle of conditional preservation (OCF-PCP)

Let  $\Omega = \{\omega_1, \dots, \omega_m\}$  and  $\Omega' = \{\omega'_1, \dots, \omega'_m\}$  be two sets of possible worlds (not necessarily different).

If for each conditional  $(B_i|A_i)$  in  $\Delta$ ,  $\Omega$  and  $\Omega'$  behave the same, i.e., they show the same number of verifications resp. falsifications, then prior  $\kappa$  and posterior  $\kappa^*$  are balanced by

$$\begin{aligned} & (\kappa(\omega_1) + \dots + \kappa(\omega_m)) - (\kappa(\omega'_1) + \dots + \kappa(\omega'_m)) \\ & = (\kappa^*(\omega_1) + \dots + \kappa^*(\omega_m)) - (\kappa^*(\omega'_1) + \dots + \kappa^*(\omega'_m)) \end{aligned}$$

[Kern-Isberner 2001]

If  $\Omega$  and  $\Omega'$  behave the same with respect to  $\Delta$ , then their differences are the same in prior and posterior OCF.

116 / 148

## A simple principle of conditional preservation

The general principle of conditional preservation yields a simple, straightforward consequence:

### Simple PCP

**(SCondPres)** If two possible worlds  $\omega_1, \omega_2 \in \Omega$  verify resp. falsify exactly the same conditionals in  $\Delta$ , then

$$\kappa^*(\omega_1) - \kappa(\omega_1) = \kappa^*(\omega_2) - \kappa(\omega_2).$$

(SCondPres) claims that the amount of change between prior and posterior epistemic state depends only on the conditionals in the new information set, more precisely, on the so-called conditional structure of the respective world.

117 / 148

## Revision techniques for inductive reasoning

Nonmonotonic inductive reasoning based on belief revision techniques is possible by taking a uniform epistemic state  $\Psi_u (= \kappa_u)$ <sup>4</sup> as prior epistemic state:

Let  $\Delta = \{(B_1|A_1), \dots, (B_n|A_n)\}$  be a finite set of conditionals;

$$\kappa_u * \Delta$$

allows model-based inductive inference.

C-representations of  $\Delta \equiv$  c-revisions of  $\kappa_u$  by  $\Delta$

- improving system  $Z$  [Pearl 1990]
- generalizing system  $Z^*$  [Goldszmidt, Morris & Pearl 1993]

This allows for a seamless integration of reasoning and revision.

<sup>4</sup>  $\kappa_u(\omega) = 0$  for all  $\omega \in \Omega$

119 / 148

## C-revisions

... are revisions that satisfy the principle of conditional preservation.

New information  $\Delta = \{(B_1|A_1), \dots, (B_n|A_n)\}$

### OCF c-revision

$$\kappa^* = \kappa * \Delta : \kappa^*(\omega) = \kappa_0 + \kappa(\omega) + \sum_{\substack{1 \leq i \leq n \\ \omega \models A_i \bar{B}_i}} \kappa_i^-,$$

$\kappa_i^-$ 's have to be chosen appropriately to ensure  $\kappa^* \models \mathcal{R}$  (Success).

$\kappa_i^-$  is the impact that conditional  $(B_i|A_i)$  has in the change process. (Success) is satisfied iff for all  $i, 1 \leq i \leq n$ ,

$$\kappa_i^- > \min_{\omega \models A_i B_i} (\kappa(\omega) + \sum_{\substack{j \neq i \\ \omega \models A_j \bar{B}_j}} \kappa_j^-) - \min_{\omega \models A_i \bar{B}_i} (\kappa(\omega) + \sum_{\substack{j \neq i \\ \omega \models A_j \bar{B}_j}} \kappa_j^-).$$

118 / 148

## Extended example – birds' scenario

*f - flying      b - birds*  
*p - penguins   w - winged animal*  
*k - kiwis      d - doves.*

$\Delta$      $r_1 : (f|b)$     *birds fly*  
 $r_2 : (b|p)$     *penguins are birds*  
 $r_3 : (\bar{f}|p)$     *penguins do not fly*  
 $r_4 : (w|b)$     *birds have wings*  
 $r_5 : (b|k)$     *kiwis are birds*  
 $r_6 : (b|d)$     *doves are birds*

Strict knowledge: *Penguins, kiwis, and doves are pairwise exclusive.*

120 / 148

## Birds' scenario (cont'd)

Let  $*$  be a c-revision for OCF.

Initial epistemic state:

$$\kappa = \kappa_u * \Delta,$$

where  $\kappa_u$  is the uniform ranking function.

Here we have:

Initial state obtained via  $\kappa_u * \Delta$

$$\kappa_u * \Delta \models (\bar{f}|p), (w|p), (w|d), (w|k), (f|d), (f|k)$$

Penguin-birds do not fly, but all birds – penguins, kiwis, and doves – inherit the property of having wings from their superclass *birds*; kiwis and doves are supposed to fly.

(Note that exactly the same beliefs hold for kiwis and doves!)

121 / 148

## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 References

123 / 148

## Birds' scenario (cont'd)

Current epistemic state:  $\kappa = \kappa_u * \Delta$

- **Revision:** Now, the agent gets to know that having wings is false for kiwis - kiwis do **not** possess wings:

$$\kappa_1^* = \kappa_u * (\Delta \cup \{(\bar{w}|k)\})$$

$\kappa_1^* \models (b|k), (f|k)$  – kiwis are birds, kiwis fly.

- **Update:** The agent learns from the news, that, due to some mysterious illness that has occurred recently among doves, the wings of newborn doves are nearly completely mutilated:

$$\kappa_2^* = \kappa * \{(\bar{w}|d)\} = (\kappa_u * \Delta) * \{(\bar{w}|d)\}$$

$\kappa_2^* \not\models (b|d), (f|d)$  – now it is unknown whether doves are birds or fly

122 / 148

## Section 6

## Probabilistic belief revision

124 / 148

## 200 years before AGM ...

Considering the task of belief change is not new: About 200 years before AGM theory, Bayes came up with his famous rule in probabilistics:

$$P(B|A) = \frac{P(A \wedge B)}{P(A)}$$

Actually, Bayesian conditioning fulfills the core ideas of AGM theory, but obviously, the contexts of the theories (changing a code of law for AGM vs. random experiments and chances – e.g., in gambling – for Bayes) seemed to be too diverse to realise a strong connection.

125 / 148

## An advanced probabilistic belief change task

The agent wants to adapt her probabilistic belief state  $P$  to a set of new conditional beliefs  $\mathcal{R} = \{(B_1|A_1)[x_1], \dots, (B_1|A_1)[x_1]\}$  – what is  $P * \{(B_1|A_1)[x_1], \dots, (B_1|A_1)[x_1]\}$ ?<sup>5</sup>

Use cross-entropy = information distance (= Kullback-Leibler-divergence)

$$R(Q, P) = \sum_{\omega \in \Omega} Q(\omega) \log \frac{Q(\omega)}{P(\omega)}$$

### ME belief change

Given some prior  $P$  and some new  $\mathcal{R}$ , choose the unique distribution

$$P^* = P *_{ME} \mathcal{R} = \arg \min_{Q \models \mathcal{R}} R(Q, P)$$

that satisfies  $\mathcal{R}$  and has minimal information distance to  $P$ .

The principle of minimum cross-entropy (MinREnt) generalizes the principle of maximum entropy (MaxEnt).

<sup>5</sup> $\mathcal{R}$  may contain probabilistic conditionals as well as probabilistic and logical facts.

127 / 148

## The general task of belief change

However, from a formal resp. epistemic point of view, the tasks are similar if not identical:

### General task of belief change

Given some (prior) epistemic state  $\Psi$  and some new information  $I$ , change beliefs rationally by applying a change operator  $*$  to obtain a (posterior) epistemic state  $\Psi'$ :

$$\Psi * I = \Psi'$$

AGM :  $\Psi = K$  set of propositional beliefs

Bayes :  $\Psi = P$  probability distribution

both :  $I = A$  propositional belief

126 / 148

## The big conditional picture

Probabilities		Ranking functions
Principle MaxEnt	$\longleftrightarrow$	c-representations
$\updownarrow$		$\updownarrow$
Principle MinREnt	$\longleftrightarrow$	c-revisions

### Principles of conditional preservation

underlie all these reasoning mechanisms – they emerged from a probabilistic principle of conditional preservation that is one of the main guidelines for the principles of optimum entropy.

128 / 148

## Back to commonsense reasoning ...

Jeff Paris:

Common sense and maximum entropy.

Synthese 117, 75-93, 1999, Kluwer Academic Publishers

### Theorem

Each (model-based) probabilistic inference process  $N$  that satisfies 7 principles of commonsense reasoning coincides with MaxEnt inference.

→ MaxEnt reasoning

- satisfies commonsense principles of probabilistic reasoning, and
- is the only probabilistic inference process doing so.

129 / 148

## Translation of Syllogisms to Probabilistic Conditionals

(joint work with Marco Wilhelm)

Syllogism	Conditional
All A's are B's	$(B A)[1]$
No A's are B's	$(B A)[0]$
Some A's are B's	$(B A)[0.65]$
Some A's are not B's	$(B A)[0.15]$

131 / 148

## Principle of Maximum Entropy [Paris 2006]

- A knowledge base  $\mathcal{R} = \{(B_1|A_1)[p_1], \dots, (B_n|A_n)[p_n]\}$  is a set of conditional statements of the form:

“If  $A$  holds, then  $B$  follows with probability  $p$ .”

- A probability distribution  $\mathcal{P}$  can be seen as a formalization of the belief state of a reasoner with knowledge  $\mathcal{R}$  iff

$$\mathcal{P}(B_i|A_i) = p_i \quad \forall i = 1, \dots, n.$$

### Definition

The maximum entropy distribution  $\mathcal{P}_{\mathcal{R}}^{\text{ME}}$  is the unique probability distribution

$$\mathcal{P}_{\mathcal{R}}^{\text{ME}} = \arg \max_{\mathcal{P} \models \mathcal{R}} - \sum_{\omega \in \Omega} \mathcal{P}(\omega) \cdot \log_2(\mathcal{P}(\omega))$$

that satisfies  $\mathcal{R}$  and adds as few information as possible.

130 / 148

## How Does the MaxEnt Model Work?

Let  $\mathcal{R}$  be a set of conditionals derived from given syllogisms.

- 1 Calculate the maximum entropy distribution  $\mathcal{P}_{\mathcal{R}}^{\text{ME}}$ .

- 2 For every query

{All | No | Some | Some not} A's are B's?

calculate  $p = \mathcal{P}_{\mathcal{R}}^{\text{ME}}(B|A)$ .

- 3 If the quantifier is  $\left\{ \begin{array}{c} \text{All} \\ \text{No} \\ \text{Some} \\ \text{Some Not} \end{array} \right\}$  and  $\left\{ \begin{array}{c} p = 1 \\ p = 0 \\ p \in [0.65 \pm t] \\ p \in [0.15 \pm t] \end{array} \right\}$ ,

then accept the answer.

(here: threshold  $t = 0.1$ )

- 4 If no answer is accepted, return *NVC*. Otherwise, return any accepted answer.

132 / 148

## Performance of MaxEnt at the CogSci 2019 Challenge

- The [MaxEnt model for syllogisms](#) was evaluated on benchmark examples and proved to be comparable to best models.
- MaxEnt performs particularly well when no training data is available.
- When enough training data is available, a [MaxEnt-MFA hybrid model](#) performed best; prediction accuracy of MaxEnt MFA Hybrid: 44.03 %.

133 / 148

References

## Section 7

### References

135 / 148

## Talk Overview

- 1 How human reasoning deviates from classical logic
- 2 Cognitive perspective
- 3 Formal models of commonsense reasoning
- 4 Cognitive aspects of Cognitive Logics
- 5 From nonmonotonic reasoning to belief revision
- 6 Probabilistic belief revision
- 7 **References**

134 / 148

References

## References: Cognitive I

- [1] J. Beattie and J. Baron.  
Confirmation and matching biases in hypothesis testing.  
*The Quarterly Journal of Experimental Psychology*, 40(2):269–297, 1988.
- [2] R. Byrne.  
Suppressing valid inferences with conditionals.  
*Cognition*, 31:61–83, 1989.
- [3] J. R. Cox and R. A. Griggs.  
The effects of experience on performance in wason's selection task.  
*Memory & Cognition*, 10(5):496–502.
- [4] J. S. B. Evans.  
Toward a statistical theory of reasoning.  
*The Quarterly Journal of Experimental Psychology*, 29(4):621–635, 1977.

136 / 148

## References: Cognitive II

- [5] J. S. B. Evans, J. L. Barston, and P. Pollard.  
On the conflict between logic and belief in syllogistic reasoning.  
*Memory & cognition*, 11(3):295–306, 1983.
- [6] A. Feeney and S. J. Handley.  
The suppression of q card selections: Evidence for deductive inference in wason's selection task.  
*The Quarterly Journal of Experimental Psychology: Section A*, 53(4):1224–1242, 2000.
- [7] L. Fiddick, L. Cosmides, and J. Tooby.  
No interpretation without representation: the role of domain-specific representations and inferences in the wason selection task.  
*Cognition*, 77(1):1–79.

137 / 148

## References: Cognitive IV

- [12] K. C. Klauer, J. Musch, and B. Naumer.  
On belief bias in syllogistic reasoning.  
*Psychological Review*, 107(4):852–884, 2000.
- [13] M. Oaksford and N. Chater.  
A rational analysis of the selection task as optimal data selection.  
*Psychological Review*, 101(4):608, 1994.
- [14] P. Pollard.  
The effect of thematic content on the 'wason selection task'.  
*Current Psychological Research*, 1(1):21–29.
- [15] P. Pollard.  
Nonindependence of selections on the wason selection task.  
*Bulletin of the Psychonomic Society*, 23(4):317–320, 1985.

139 / 148

## References: Cognitive III

- [8] R. A. Griggs.  
Memory cueing and instructional effects on wason's selection task.  
*Current Psychology*, 3(4):3–10.
- [9] P. N. Johnson-Laird.  
*How we reason*.  
Oxford University Press, New York, 2006.
- [10] P. N. Johnson-Laird and P. C. Wason.  
A theoretical analysis of insight into a reasoning task.  
*Cognitive Psychology*, 1(2):134–148, 1970.
- [11] K. N. Kirby.  
Probabilities and utilities of fictional outcomes in wason's four-card selection task.  
*Cognition*, 51(1):1–28, 1994.

138 / 148

## References: Cognitive V

- [16] M. Ragni, E.-A. Dietz, I. Kola, and S. 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.
- [17] M. Ragni, I. Kola, and P. Johnson-Laird.  
On selecting evidence to test hypotheses: A theory of selection tasks.  
*Psychological Bulletin*, 144(8):779–796, 2018.
- [18] K. Stenning and M. Van Lambalgen.  
*Human reasoning and cognitive science*.  
MIT Press, 2008.

140 / 148

## References: Cognitive VI

- [19] D. Trippas, D. Kellen, H. Singmann, G. Pennycook, D. J. Koehler, J. A. Fugelsang, and C. Dubé.  
Characterizing belief bias in syllogistic reasoning: A hierarchical bayesian meta-analysis of roc data.  
*Psychonomic bulletin & review*, 25(6):2141–2174, 2018.
- [20] A. Tversky and D. Kahneman.  
Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment.  
*Psychological Review*, 90(4):293–315, 1983.
- [21] P. Wason.  
Reasoning about a rule.  
*Quarterly Journal of Experimental Psychology*, 20(3):273–281, 1968.

141 / 148

## References: Formal I

- [Ada75] E.W. Adams.  
*The Logic of Conditionals*.  
D. Reidel, Dordrecht, 1975.
- [Ant97] G. Antoniou.  
*Nonmonotonic reasoning*.  
MIT Press, Cambridge, Mass., 1997.
- [DHR12] E. Dietz, S. Hölldobler, and M. Ragni.  
A Computational Approach to the Suppression Task.  
In N. Miyake, D. Peebles, and R.P. Cooper, editors, *Proc. of the 34th Annual Conference of the Cognitive Science Society*, pages 1500–1505, Austin, TX, 2012. Cognitive Science Society.

143 / 148

## References: Cognitive VII

- [22] S. A. Yachanin.  
Facilitation in wason's selection task: Content and instructions.  
*Current Psychology*, 5(1):20–29.
- [23] S. A. Yachanin and R. D. Tweney.  
The effect of thematic content on cognitive strategies in the four-card selection task.  
*Bulletin of the Psychonomic Society*, 19(2):87–90, 2013.

142 / 148

## References: Formal II

- [EKIR18] Christian Eichhorn, Gabriele Kern-Isberner, and Marco Ragni.  
Rational inference patterns based on conditional logic.  
In S.A. McIlraith and K.Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18)*. AAAI Press, 2018.
- [Gär92] P. Gärdenfors.  
Belief revision and nonmonotonic logic: Two sides of the same coin?  
In *Proceedings European Conference on Artificial Intelligence, ECAI'92*, pages 768–773. Pitman Publishing, 1992.
- [GMP93] M. Goldszmidt, P. Morris, and J. Pearl.  
A maximum entropy approach to nonmonotonic reasoning.  
*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):220–232, 1993.

144 / 148

## References: Formal III

- [GP96] M. Goldszmidt and J. Pearl.  
Qualitative probabilities for default reasoning, belief revision, and causal modeling.  
*Artificial Intelligence*, 84:57–112, 1996.
- [HK09] S. Hölldobler and C. D. Kencana Ramli.  
Logic programs under three-valued Łukasiewicz semantics.  
In P. M. Hill and D. S. Warren, editors, *Logic Programming, 25th International Conference, ICLP 2009*, volume 5649 of *LNCS*, pages 464–478, Heidelberg, 2009. Springer.
- [KI98] G. Kern-Isberner.  
Characterizing the principle of minimum cross-entropy within a conditional-logical framework.  
*Artificial Intelligence*, 98:169–208, 1998.

145 / 148

## References: Formal V

- [Par99] Jeff Paris.  
Common sense and maximum entropy.  
*Synthese*, 117:75–93, 1999.
- [Pea90] J. Pearl.  
System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning.  
In R. Parikh, editor, *Proc. of the 3rd conference on Theoretical aspects of reasoning about knowledge (TARK1990)*, pages 121–135, San Francisco, CA, USA, 1990. Morgan Kaufmann Publishers Inc.
- [Rei80] R. Reiter.  
A logic for default reasoning.  
*Artificial Intelligence*, 13:81–132, 1980.

147 / 148

## References: Formal IV

- [KI01] G. Kern-Isberner.  
*Conditionals in Nonmonotonic Reasoning and Belief Revision*.  
Number 2087 in *LNCS*. Springer, Berlin, DE, 2001.
- [KI04] G. Kern-Isberner.  
A thorough axiomatization of a principle of conditional preservation in belief revision.  
*Annals of Mathematics and Artificial Intelligence*, 40:127–164, 2004.
- [Mak89] D. Makinson.  
General theory of cumulative inference.  
In M. Reinfrank et al., editors, *Non-monotonic Reasoning*, pages 1–18. Springer Lecture Notes on Artificial Intelligence 346, Berlin, 1989.

146 / 148

## References: Formal VI

- [REKI16] Marco Ragni, Christian Eichhorn, and Gabriele Kern-Isberner.  
Simulating human inferences in the light of new information: a formal analysis.  
In S. Kambhampati, editor, *Proceedings of the 25th International Joint Conference on Artificial Intelligence, IJCAI 2016*, pages 2604–2610, Palo Alto, CA, USA, 2016. AAAI Press.
- [Spo12] W. Spohn.  
*The Laws of Belief: Ranking Theory and Its Philosophical Applications*.  
Oxford University Press, Oxford, UK, 2012.

148 / 148