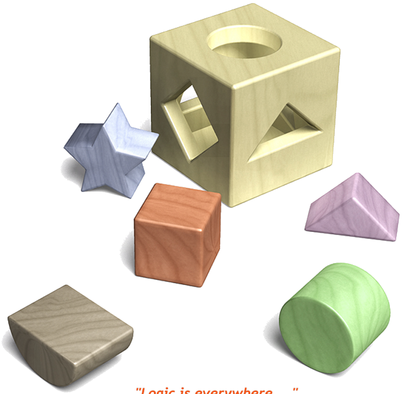
A New Computational Model for Human Reasoning

Steffen Hölldobler
International Center for Computational Logic
Technische Universität Dresden
Germany

- The Suppression Task
- The New Computational Model
- Weak Completion Semantics
- Abduction
- Well-Founded Semantics
- Discussion



"Logic is everywhere ..."

The Suppression Task – Part I

- Byrne: Suppressing Valid Inferences with Conditionals.
  Cognition 31, 61-83: 1989

- **Conditionals**

  LE If she has an essay to write then she will study late in the library.
  LT If she has a textbook to read then she will study late in the library.
  LO If the library stays open then she will study late in the library.

- **Facts** E She has an essay to write.
  ¬E She does not have an essay to write.

- Will she study late in the library? ☐ yes ☐ no ☐ I don't know

Conditionals	Facts	Yes	No	Don't Know
LE	E	96%		
LE & LT	E	96%		
LE & LO	E	38%		
LE	¬E		46%	
LE & LT	¬E		4%	
LE & LO	¬E		63%	

The Suppression Task – Part II

- **Conditionals**
  LE If she has an essay to write then she will study late in the library.
  LT If she has a textbook to read then she will study late in the library.
  LO If the library stays open then she will study late in the library.

- **Facts** L She will study late in the library.
  ¬L She will not study late in the library.

- **Has she an essay to write?** ☐ yes ☐ no ☐ I don't know

Conditionals	Facts	Yes	No	Don't Know
LE	L	53%		
LE & LT	L	16%		
LE & LO	L	55%		
LE	¬L		69%	
LE & LT	¬L		69%	
LE & LO	¬L		44%	

- **Classical logic is inadequate!**

A New Computational Model – Prelude

- **Can we find a logic which adequately models human reasoning?**
- **Stenning, van Lambalgen:**
  Human Reasoning and Cognititve Science. MIT Press: 2008
  - **Main ideas**
  - **Technical flaws**
- **H., Kencana Ramli:**
  Logic Programs under Three-Valued Łukasiewicz's Semantics.
  In: Logic Programming. Hill, Warren (eds), LNCS 5649, 464-478: 2009
  - **Adapted main ideas**
  - **Rigorous proofs**

The New Approach

- Reasoning towards an appropriate logical form
  - Logic programs
- Weak completion semantics
  - Non-monotonicity
- Three-valued Łukasiewicz logic
  - Least models
- An appropriate semantic operator
  - Least fixed points are least models
  - Least fixed points can be computed by iterating the operator
- Reasoning with respect to the least models
- Abduction and sceptical reasoning
- A connectionist realization

Logic Programs

- **Preliminaries**
  - An **atom** is an atomic propositions.
  - A **literal** is either an atom or its negation.
  - $\top$ and $\bot$ denote truth and falsehood, respectively.
- A **(logic) program** is a finite set of rules.
  - A **rule** is an expression of the form $A \leftarrow B_1 \wedge \cdots \wedge B_n$, where $n \geq 1$, A is an atom, and each B_i, $1 \leq i \leq n$, is either a literal, $\top$ or $\bot$.
  - A is called **head** and $B_1 \wedge \cdots \wedge B_n$ **body** of the rule.
  - Rules of the form $A \leftarrow \top$ are called **positive facts**.
  - Rules of the form $A \leftarrow \bot$ are called **negative facts**.

Reasoning Towards an Appropriate Logical Form

- Stenning, van Lambalgen 2008
- Represent conditionals as licences for implications

LE & E $\{\ell \leftarrow e \wedge \neg ab_1,\ ab_1 \leftarrow \bot,\ e \leftarrow \top\}$

LE & LT & E $\{\ell \leftarrow e \wedge \neg ab_1,\ ab_1 \leftarrow \bot,\ \ell \leftarrow t \wedge \neg ab_2,\ ab_2 \leftarrow \bot,\ e \leftarrow \top\}$

- Reason about additional premises

LE & LO & E $\{\ell \leftarrow e \wedge \neg ab_1,\ ab_1 \leftarrow \neg o,\ \ell \leftarrow o \wedge \neg ab_2,\ ab_2 \leftarrow \neg e,\ e \leftarrow \top\}$

Weak Completion

- Let $\mathcal{P}$ be a program. Consider the following transformation:
  1. All rules with the same head $A \leftarrow Body_1,\ A \leftarrow Body_2, \ldots$ are replaced by $A \leftarrow Body_1 \vee Body_2 \vee \ldots$.
  2. If an atom A is not the head of any rule in $\mathcal{P}$ then add $A \leftarrow \bot$.
  3. All occurrences of $\leftarrow$ are replaced by $\leftrightarrow$.
     - The resulting set is called completion of $\mathcal{P}$ or c $\mathcal{P}$.
     - If step 2 is omitted then the resulting set is called weak completion of $\mathcal{P}$ or wc $\mathcal{P}$.
- Completion versus weak completion
  - $\text{c}\,\{p \leftarrow q\} = \{p \leftrightarrow q,\ q \leftrightarrow \bot\} \neq \{p \leftrightarrow q\} = \text{wc}\,\{p \leftarrow q\}$
  - $\text{c}\,\{p \leftarrow q\} = \text{wc}\,\{p \leftarrow q,\ q \leftarrow \bot\}$

Three-Valued Interpretations (1)

- A (three-valued) interpretation assigns to each formula a value from $\{\top, \bot, U\}$. It is represented by $\langle I^\top, I^\bot \rangle$, where
  - $I^\top$ contains all atoms which are mapped to $\top$,
  - $I^\bot$ contains all atoms which are mapped to $\bot$,
  - $I^\top \cap I^\bot = \emptyset$.
  - All atoms which occur neither in $I^\top$ nor $I^\bot$ are mapped to U.
- Łukasiewicz: O logice trójwartościowej. Ruch Filozoficzny 5, 169-171: 1920

$$U \leftarrow_{3Ł} U = \top$$

- Kleene: Introduction to Metamathematics. North-Holland: 1962

$$U \leftarrow_{3K} U = U$$

Three-Valued Interpretations

- **Let $\mathcal{I}$ denote the set of all three-valued interpretations.**
- **Knowledge ordering**
  - $\langle I^\top, I^\bot \rangle \preceq \langle J^\top, J^\bot \rangle$ **iff** $I^\top \subseteq J^\top$ **and** $I^\bot \subseteq J^\bot$
- **Fitting: A Kripke-Kleene Semantics for Logic Programs. Journal of Logic Programming 2, 295-312: 1985**
  - $(\mathcal{I}, \preceq)$ **is a complete semi-lattice.**
- $\langle I^\top, I^\bot \rangle \cap \langle J^\top, J^\bot \rangle = \langle I^\top \cap J^\top, I^\bot \cap J^\bot \rangle$

Logic Programs under Weak Completion Semantics (WCS)

- H., Kencana Ramli 2009
  - The model intersection property holds for each program $\mathcal{P}$, i.e., $\cap\{I \mid I \models_{3Ł} \mathcal{P}\} \models_{3Ł} \mathcal{P}$.
  - This does not hold under Kleene semantics:

$$\begin{array}{lll} \langle\{p,q\},\emptyset\rangle & \models_{3K} & \{p \leftarrow q\} \\ \langle\emptyset,\{p,q\}\rangle & \models_{3K} & \{p \leftarrow q\} \end{array} \quad \text{but} \quad \langle\emptyset,\emptyset\rangle \not\models_{3K} \{p \leftarrow q\}$$

  - The model intersection property extends to weakly completed programs.
  - Each weakly completed program has a least model.
- WCS = weak completion + three-valued Łukasiewicz logic
  - WCS adequately models part I of the suppression task.

Computing the Least Models of Weakly Completed Programs

- **How can we compute the least models of weakly completed programs?**
- Stenning, van Lambalgen 2008
  Consider the following immediate consequence operator:
  $\Phi_{\mathcal{P}}(I) = \langle J^{\top}, J^{\perp} \rangle$, **where**

$$
\begin{array}{lcl}
J^{\top} & = & \{A \mid \text{there exists } A \leftarrow Body \in \mathcal{P} \text{ with } I(Body) = \top\} \text{ and} \\
J^{\perp} & = & \{A \mid \text{there exists } A \leftarrow Body \in \mathcal{P} \text{ and} \\
 & & \quad \text{for all } A \leftarrow Body \in \mathcal{P} \text{ we find } I(Body) = \perp\}.
\end{array}
$$

- **Note** $\Phi_{\mathcal{P}}$ **'without the red condition' is the Fitting operator** (Fitting 1985).
- **Theorem** (H., Kencana Ramli 2009)
  (1) $\Phi_{\mathcal{P}}$ **is monotone on** $(\mathcal{I}, \subseteq)$.
  (2) $\Phi_{\mathcal{P}}$ **is continuous**
  and, hence, admits a least fixed point denoted by $\text{lfp}\, \Phi_{\mathcal{P}}$.
  (3) $\text{lfp}\, \Phi_{\mathcal{P}}$ **can be computed by iterating** $\Phi_{\mathcal{P}}$ **on** $\langle \emptyset, \emptyset \rangle$.
  (4) $\text{lm}_{3Ł}\, \text{wc}\, \mathcal{P} = \text{lfp}\, \Phi_{\mathcal{P}}$.

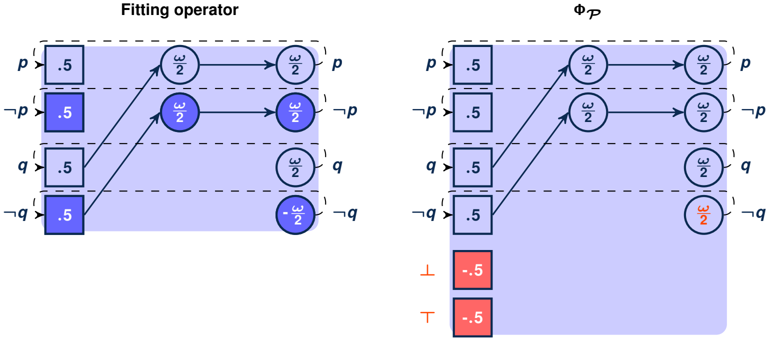
A Connectionist Realization – Some History

- Towell, Shavlik: Extracting Refined Rules from Knowledge–Based Neural Networks. Machine Learning 131, 71-101: 1993
  - **Feedforward networks for hierarchical logic programs.**
- H., Kalinke: Towards a New Massively Parallel Computational Model for Logic Programming In: Proceedings of the ECAI94 Workshop on Combining Symbolic and Connectionist Processing, 68-77: 1994
  - **Feedforward networks for the immediate consequent operator.**
  - **Additional recurrent connections to compute least fixed points.**
- Kalinke: Ein massiv paralleles Berechnungsmodell für normale logische Programme, TU Dresden, Fakultät Informatik: 1994
  - **Extension to three-valued programs under Kleene semantics.**
- **CORE Method**
  connectionist model generation using recurrent networks with feedforward core

Networks for Three-Valued Logic Programs

- Remember $\mathsf{c}\,\{p \leftarrow q\} = \{p \leftrightarrow q, q \leftarrow \bot\}$ and $\mathsf{wc}\,\{p \leftarrow q\} = \{p \leftrightarrow q\}$.
- A translation algorithm translates programs into feedforward network.
- Recurrent connections connect the output to the input layer.

Kalinke 1994

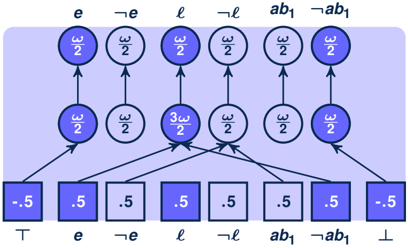


A CORE Method for the Suppression Task

- H., Kencana Ramli: Logics and Networks for Human Reasoning.
  In: Artificial Neural Networks, Alippi et.al. eds., LNCS 5769, 85-94: 2009
  - **For each program $\mathcal{P}$ there exists a feed-forward core computing $\Phi_{\mathcal{P}}$.**
  - **The recurrent network reaches a stable state representing lfp $\Phi_{\mathcal{P}}$ if initialized with $\langle\emptyset,\emptyset\rangle$.**

The Suppression Task – Modus Ponens

- **LE & E**
- $\mathcal{P} = \{\ell \leftarrow e \wedge \neg ab_1,\ e \leftarrow \top,\ ab_1 \leftarrow \bot\}$



- **lfp** $\Phi_{\mathcal{P}}$ = **Im**$_{3Ł}$ **wc** $\mathcal{P} = \langle\{\ell, e\}, \{ab_1\}\rangle$
- **From** $\langle\{\ell, e\}, \{ab\}\rangle$ **follows that she will study late in the library.**

Abduction

- **LE & LT & L**

Knowledge base	$\{\ell \leftarrow e \wedge \neg ab_1,\ ab_1 \leftarrow \bot,\ \ell \leftarrow t \wedge \neg ab_2,\ ab_2 \leftarrow \bot\}$
Observation	$\{\ell\}$
Abducibles	$\{e \leftarrow \top,\ e \leftarrow \bot,\ t \leftarrow \top,\ t \leftarrow \bot\}$
Minimal explanations	$\{e \leftarrow \top\}$ **and** $\{t \leftarrow \top\}$

  - **Reasoning credulously we conclude e.**
  - **Reasoning sceptically we cannot conclude e.**
  - Byrne 1989 **only 16% conclude e.**

- **WCS + sceptical abduction adequately models part II of the suppression task.**

- H., Philipp, Wernhard: An Abductive Model for Human Reasoning.
  In: Proceedings of the 10th International Symposium on Logical Formalizations of Commonsense Reasoning (CommonSense): 2011

- Dietz, H., Ragni: A Computational Logic Approach to the Suppression Task.
  In: Proceedings of the 34th Annual Conference of the Cognitive Science Society, Miyake et.al. eds., 1500-1505: 2012

Weak Completion versus Well-Founded Semantics (1)

- Dietz, H., Wernhard: Modeling the Suppression Task under Weak Completion and Well-Founded Semantics: Journal of Applied Non-Classical Logics (to appear)
  - A program is tight if it does not contain positive cycles.
  - All programs for the suppression (and the selection) task are tight.
  - Let $\mathcal{P}$ be a tight program and I an interpretation.
  - Theorem The following statements are equivalent:
    - I is a least model of the weak completion of $\mathcal{P}$.
    - I is a well-founded model of $\mathcal{P}'$,
      where $\mathcal{P}'$ is obtained from $\mathcal{P}$ by deleting all negative facts and adding for each undefined predicate symbol A occurring in $\mathcal{P}$ the rules $A \leftarrow \neg A'$ and $A' \leftarrow \neg A$, where A' is a new symbol.
- Well-founded semantics (WFS) appears to be adequate if conditionals do not contain positive cycles!

Weak Completion versus Well-Founded Semantics (2)

- How do humans reason with positive cycles?
  - If they open the window, then they open the window.
  - If they open the window, then it is cold.
    If it is cold, then they wear their jackets.
    If they wear their jackets, then they open the windows.
- Psychological study
  - We presented conditionals with positive cycles of length one, two and three, asked whether embedded propositions or their negations are entailed.
  - Preliminary results

length	positive	negative (WFS)	unknown (WCS)	response time
1	75 %	0 %	25 %	5257 msec
2	60 %	3 %	37 %	11516 msec
3	55 %	4 %	41 %	11680 msec

  - Humans consider positive cycles of length one as facts.
  - The longer the cycles, the more likely is the answer ’unknown’.
  - Almost nobody entailed negative propositions.

The Selection Task – Abstract Case

- Wason: Reasoning about a Rule.
  The Quarterly Journal of Experimental Psychology 20, 273-281: 1968
- **Consider cards which have a letter on one side and a number on the other side.**

D	F	3	7

- **Consider the rule:**

  if there is a D on one side, then there is a 3 on the other side.

- **Which cards do you have to turn in order to show that the rule holds?**
  - **Only 10% of the subjects give the logically correct solutions.**

An Analysis

- Almost everyone (89%) correctly selects D.
  - Corresponds to modus ponens in classical logic.
- Almost everyone (84%) correctly does not select F.
  - *Because the condition does not mention F.*
- Many (62%) incorrectly select 3.
  - *If there is a 3 on one side, then there is a D on the other side*.
  - Converse of the given conditional.
- Only a small percentage of subjects (25%) correctly selects 7.
  - *If the number on one side is not 3, then the letter on the other side is not D.*
  - Contrapositive of the given conditional.

The Selection Task – Social Case

- Griggs, Cox: The elusive thematic materials effect in the Wason selection task. British Journal of Psychology 73, 407-420: 1982
- **Consider cards which have a person's age on the one side and a drink on the other side.**

beer	coke	22yrs	16yrs

- **Consider the rule:**

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

- **Which cards do you have to turn in order to show that the rule holds?**
  - **Most people solve this variant correctly.**

Formalizing the Social Case

- **The conditional is viewed as a social constraint.**
- **Let o and a be propositional variables denoting that the person is older than 19 years and is drinking alcohol, respectively.**
- **The rule is encoded by $\mathcal{C} = \{o \leftarrow a \wedge \neg ab\}$.**
- **Consider the four cases:**

case	$\mathcal{P}$	$\mathrm{lm}_{3Ł}\,\mathrm{wc}\,\mathcal{P}$				
beer	$\{a \leftarrow \top,\ ab \leftarrow \bot\}$	$\langle\{a\}, \{ab\}\rangle$	$\not\models_{3Ł}$	$\mathcal{C}$	$\leadsto$	check
22yrs	$\{o \leftarrow \top,\ ab \leftarrow \bot\}$	$\langle\{o\}, \{ab\}\rangle$	$\models_{3Ł}$	$\mathcal{C}$	$\leadsto$	no check
coke	$\{a \leftarrow \bot,\ ab \leftarrow \bot\}$	$\langle\emptyset, \{a, ab\}\rangle$	$\models_{3Ł}$	$\mathcal{C}$	$\leadsto$	no check
16yrs	$\{o \leftarrow \bot,\ ab \leftarrow \bot\}$	$\langle\emptyset, \{o, ab\}\rangle$	$\not\models_{3Ł}$	$\mathcal{C}$	$\leadsto$	check

Formalizing the Abstract Case

- The conditional is viewed as a belief.
- Let $D, F, 3, 7$ be propositional variables denoting that the corresponding symbol is on one side.
- Consider $\mathcal{P} = \{3 \leftarrow D \wedge \neg ab,\ ab \leftarrow \bot\}$ with $\mathrm{lm}_{3Ł}\,\mathrm{wc}\,\mathcal{P} = \langle \emptyset, \{ab\} \rangle$.
- $\langle \emptyset, \{ab\} \rangle$ does not explain any letter on a card.
- The set of abducibles is $\{D \leftarrow \top, D \leftarrow \bot, F \leftarrow \top, F \leftarrow \bot, 7 \leftarrow \top, 7 \leftarrow \bot\}$.
- Consider the four cases:

$\mathcal{O}$	$\mathcal{E}$	$\mathrm{lm}_{3Ł}\,\mathrm{wc}\,(\mathcal{P} \cup \mathcal{E})$		
D	$\{D \leftarrow \top\}$	$\langle \{D, 3\}, \{ab\} \rangle$	$\leadsto$	turn,
F	$\{F \leftarrow \top\}$	$\langle \{F\}, \{ab\} \rangle$	$\leadsto$	no turn,
3	$\{D \leftarrow \top\}$	$\langle \{D, 3\}, \{ab\} \rangle$	$\leadsto$	turn,
7	$\{7 \leftarrow \top\}$	$\langle \{7\}, \{ab\} \rangle$	$\leadsto$	no turn.

A Computational Logic Approach to the Selection Task

- **The computational logic approach to model human reasoning can be extended to adequately handle the selection task**
  - **if the social case is understood as a social constraint and**
  - **if the abstract case is understood as a belief.**

- **Kowalski: Computational Logic and Human Life: How to be Artificially Intelligent. Cambridge University Press 2011**
- **Dietz, H., Ragni: A Computational Logic Approach to the Abstract and the Social Case of the Selection Task. In: Proceedings of the 11th International Symposium on Logic Formalizations of Commonsense Reasoning: 2013**

Contraction Mappings

- Banach: Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. Fundamenta Mathematicae 3, 133-181: 1922
  - A contraction f on a complete metric space $(\mathcal{X}, d)$ has a unique fixed point; the sequence $x, f(x), f(f(x)), \ldots$ converges to this fixed point for any $x \in \mathcal{X}$.
- A level mapping is a mapping $|\cdot|$ from the set of atoms to $\mathbb{N}$.
  It is extended to literals by defining $|\neg A| = |A|$ for each atom A.
- Let $\mathcal{I}$ be the set of all interpretations and $I, J \in \mathcal{I}$.

$$d_{|\cdot|}(I, J) = \begin{cases} \frac{1}{2^n} & \text{if } I \neq J, \\ & \text{for all } A \text{ with } |A| < n : I(A) = J(A) \neq \mathsf{U}, \text{ and} \\ & \text{for some } A \text{ with } |A| = n : I(A) \neq J(A) \text{ or } I(A) = J(A) = \mathsf{U}, \\ 0 & \text{otherwise.} \end{cases}$$

Contraction Mappings and Human Reasoning

- H., Kencana-Ramli: Contraction Properties of a Semantic Operator for Human Reasoning. In: Proceedings of the Fifth International Conference on Information, Li and Yen (eds.), 228-231: 2009
  - $(\mathcal{I}, d_{|\cdot|})$ is a complete metric space.
  - If $\mathcal{P}$ is acyclic then $\Phi_{\mathcal{P}}$ is a contraction.
- All programs for the suppression and the selection task are acyclic.
  - Computation of the least models can start with an arbitrary interpretation.

Discussion

- Logic appears to be adequate for human reasoning if
  - weak completion semantics,
  - Łukasiewicz logic,
  - the Stenning and van Lambalgen semantic operator, and
  - sceptical abduction are used.
- Human reasoning is modeled by
  - reasoning towards an appropriate logic program and, thereafter,
  - reasoning with respect to the least model of its weak completion.
- This approach matches data from studies in human reasoning.
- There is a connectionist encoding.
- There are many interesting and challenging open questions.

Some Open Problems (1)

- **Negation**
  - How is negation treated in human reasoning?
- **Errors**
  - How can frequently made errors be explained in the proposed approach?
- **Łukasiezicz logic**
  - Is the Łukasiezicz logic adequate?
- **Completion**
  - Under which conditions is human reasoning adequately modeled by completion and/or weak completion?

Some Open Problems (2)

- **Contractions**
  - **Do humans exhibit a behavior which can be adequately modeled by contractional semantic operators?**
- **Explanations**
  - **Do humans consider minimal explanations?**
  - **In which order are (minimal) explanations generated by humans if there are several?**
  - **Does attention play a role in the selection of (minimal) explanations?**
- **Reasoning**
  - **Do humans reason sceptically or credulously?**
  - **How does a connectionist realization of sceptical reasoning looks like?**