On the Interplay Between Subjective Plausibility and Default Reasoning in the Context of Interpreting Natural Language Utterances

Erich Rast
erich@snafu.de

IFL, Universidade Nova de Lisboa

July 7, 2011

Overview

Interpretation and Background Beliefs

A Small Dialog

(1) Peter: Where is John?
(2) Lisa: He's over there. He's ready.
(3) Peter: Okay, let's go.

Interpretation and Background Beliefs

A Small Dialog

(1) Peter: Where is John?
(2) Lisa: He's over there. He's ready.
(3) Peter: Okay, let's go.

John is ready to go.

Interpretation and Background Beliefs

A Small Dialog

(1) Peter: Where is John?
(2) Lisa: He's over there. He's ready.
(3) Peter: Okay, let's go.

John is ready to shoot.

Interpretation and Background Beliefs

A Small Dialog

(1) Peter: Where is John?
(2) Lisa: He's over there. He's ready.
(3) Peter: Okay, let's go.

John is ready to jump.

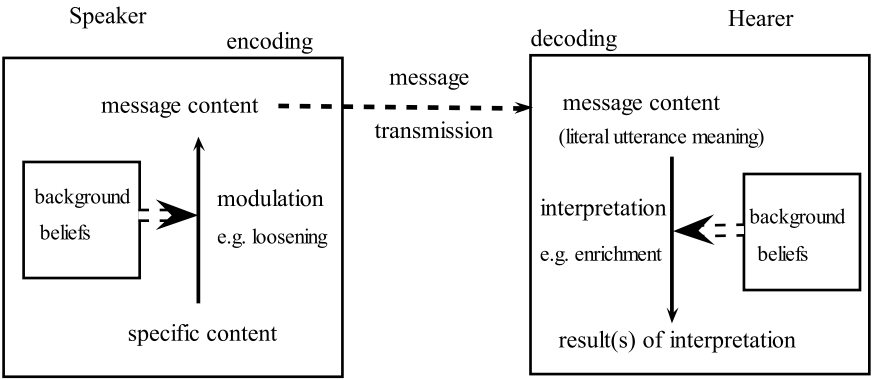
Interpretation and Background Beliefs

A Small Dialog

(1) Peter: Where is John?
(2) Lisa: He's over there. He's ready.
(3) Peter: Okay, let's go.

A well-known phenomenon:
The same dialog may be meant and interpreted in radically different ways.

General Model



The Central Ingredients of a Theory of Interpretation

A theory of interpretation requires at least the following ingredients:

1. A sufficiently rich and adequate representation of the literal meaning of utterances.
2. A representation of the individual beliefs and assumptions of the discourse participants.
3. A sufficiently rich and adequate representation of general background beliefs ('world knowledge') of discourse participants.
4. A mechanism that on the basis of these factors provides a model of how discourse participants arrive at an interpretation, where factors like the utterance context and the question under discussion are taken into account.

The Central Ingredients of a Theory of Interpretation

A theory of interpretation requires at least the following ingredients:

1. A sufficiently rich and adequate representation of the literal meaning of utterances.
2. A representation of the individual beliefs and assumptions of the discourse participants.
3. A sufficiently rich and adequate representation of general background beliefs ('world knowledge') of discourse participants.
4. A mechanism that on the basis of these factors provides a model of how discourse participants arrive at an interpretation, where factors like the utterance context and the question under discussion are taken into account.

The Central Ingredients of a Theory of Interpretation

A theory of interpretation requires at least the following ingredients:

1. A sufficiently rich and adequate representation of the literal meaning of utterances.
2. A representation of the individual beliefs and assumptions of the discourse participants.
3. A sufficiently rich and adequate representation of general background beliefs ('world knowledge') of discourse participants.
4. A mechanism that on the basis of these factors provides a model of how discourse participants arrive at an interpretation, where factors like the utterance context and the question under discussion are taken into account.

The Central Ingredients of a Theory of Interpretation

A theory of interpretation requires at least the following ingredients:

1. A sufficiently rich and adequate representation of the literal meaning of utterances.
2. A representation of the individual beliefs and assumptions of the discourse participants.
3. A sufficiently rich and adequate representation of general background beliefs ('world knowledge') of discourse participants.
4. A mechanism that on the basis of these factors provides a model of how discourse participants arrive at an interpretation, where factors like the utterance context and the question under discussion are taken into account.

The Central Ingredients of a Theory of Interpretation

A theory of interpretation requires at least the following ingredients:

1. A sufficiently rich and adequate representation of the literal meaning of utterances.
2. A representation of the individual beliefs and assumptions of the discourse participants.
3. A sufficiently rich and adequate representation of general background beliefs ('world knowledge') of discourse participants.
4. A mechanism that on the basis of these factors provides a model of how discourse participants arrive at an interpretation, where factors like the utterance context and the question under discussion are taken into account.

More Specific Ingredients

Done:

- Implement qualitative graded belief and corresponding assumptions.
- Provide a model of enrichment by abductive inference on the basis of this graded belief.

In progress:

- Transfer the above account to quantitative graded belief (probability theory, Dempster-Shafer, etc.)
- Make the representation of background beliefs realistic and sufficiently rich:
  - situations
  - default reasoning / typicality

Question: What is the interplay between default reasoning and graded belief in the context of this project?

More Specific Ingredients

Done:

- Implement qualitative graded belief and corresponding assumptions.
- Provide a model of enrichment by abductive inference on the basis of this graded belief.

In progress:

- Transfer the above account to quantitative graded belief (probability theory, Dempster-Shafer, etc.)
- Make the representation of background beliefs realistic and sufficiently rich:
  - situations
  - default reasoning / typicality

Question: What is the interplay between default reasoning and graded belief in the context of this project?

More Specific Ingredients

Done:

- Implement qualitative graded belief and corresponding assumptions.
- Provide a model of enrichment by abductive inference on the basis of this graded belief.

In progress:

- Transfer the above account to quantitative graded belief (probability theory, Dempster-Shafer, etc.)
- Make the representation of background beliefs realistic and sufficiently rich:
  - situations
  - default reasoning / typicality

Question: What is the interplay between default reasoning and graded belief in the context of this project?

More Specific Ingredients

Done:

- Implement qualitative graded belief and corresponding assumptions.
- Provide a model of enrichment by abductive inference on the basis of this graded belief.

In progress:

- Transfer the above account to quantitative graded belief (probability theory, Dempster-Shafer, etc.)
- Make the representation of background beliefs realistic and sufficiently rich:
  - situations
  - default reasoning / typicality

Question: What is the interplay between default reasoning and graded belief in the context of this project?

Graded Belief

What is graded belief? One understanding of graded belief is that the degree to which you believe something is reciprocal to your willingness to give up your belief in face of counter-evidence. Other accounts explain the degree of belief by the amount of money one is willing to bet that the believed proposition is true.

1. Qualitative graded belief: based on a preference relation over possible worlds or situations, various ways to lift comparison from points to sets; set-based approaches also possible
2. Quantitative belief: probability theory, Dempster-Shafer belief, possibility theory, ranking theory (Spohn), plausibility measures (Halpern)

Graded Belief

What is graded belief? One understanding of graded belief is that the degree to which you believe something is reciprocal to your willingness to give up your belief in face of counter-evidence. Other accounts explain the degree of belief by the amount of money one is willing to bet that the believed proposition is true.

1. Qualitative graded belief: based on a preference relation over possible worlds or situations, various ways to lift comparison from points to sets; set-based approaches also possible
2. Quantitative belief: probability theory, Dempster-Shafer belief, possibility theory, ranking theory (Spohn), plausibility measures (Halpern)

Graded Belief

What is graded belief? One understanding of graded belief is that the degree to which you believe something is reciprocal to your willingness to give up your belief in face of counter-evidence. Other accounts explain the degree of belief by the amount of money one is willing to bet that the believed proposition is true.

1. Qualitative graded belief: based on a preference relation over possible worlds or situations, various ways to lift comparison from points to sets; set-based approaches also possible
2. Quantitative belief: probability theory, Dempster-Shafer belief, possibility theory, ranking theory (Spohn), plausibility measures (Halpern)

Graded Belief

What is graded belief? One understanding of graded belief is that the degree to which you believe something is reciprocal to your willingness to give up your belief in face of counter-evidence. Other accounts explain the degree of belief by the amount of money one is willing to bet that the believed proposition is true.

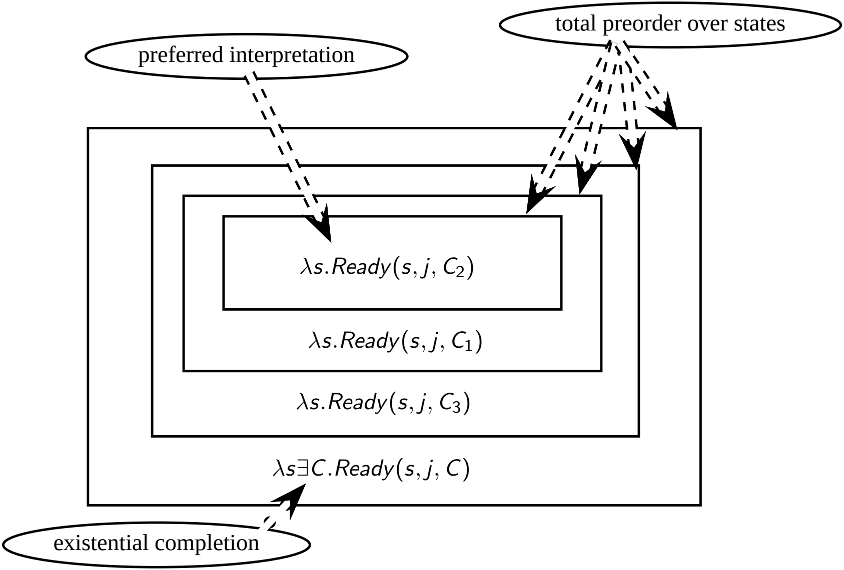
1. Qualitative graded belief: based on a preference relation over possible worlds or situations, various ways to lift comparison from points to sets; set-based approaches also possible
2. Quantitative belief: probability theory, Dempster-Shafer belief, possibility theory, ranking theory (Spohn), plausibility measures (Halpern)

Graded Belief

What is graded belief? One understanding of graded belief is that the degree to which you believe something is reciprocal to your willingness to give up your belief in face of counter-evidence. Other accounts explain the degree of belief by the amount of money one is willing to bet that the believed proposition is true.

1. Qualitative graded belief: based on a preference relation over possible worlds or situations, various ways to lift comparison from points to sets; set-based approaches also possible
2. Quantitative belief: probability theory, Dempster-Shafer belief, possibility theory, ranking theory (Spohn), plausibility measures (Halpern)

Qualitative Account



Default Reasoning / Typicality

There are many similar or closely related theories such as defeasible reasoning, non-monotonic reasoning, or prototype theory. For the purpose of modeling background beliefs we need two things:

- a way to express rules that specify what is typically the case
- a way to draw inferences from an agent's beliefs and these rules

Example

(4) Typically birds can fly.
(5) Typically penguins cannot fly.
(6) Tweety is a penguin.
$\rightsquigarrow$ Tweety cannot fly (unless he is an atypical penguin).

Default Reasoning / Typicality

There are many similar or closely related theories such as defeasible reasoning, non-monotonic reasoning, or prototype theory. For the purpose of modeling background beliefs we need two things:

- a way to express rules that specify what is typically the case
- a way to draw inferences from an agent's beliefs and these rules

Example

(4) Typically birds can fly.
(5) Typically penguins cannot fly.
(6) Tweety is a penguin.
$\rightsquigarrow$ Tweety cannot fly (unless he is an atypical penguin).

Default Reasoning / Typicality

There are many similar or closely related theories such as defeasible reasoning, non-monotonic reasoning, or prototype theory. For the purpose of modeling background beliefs we need two things:

- a way to express rules that specify what is typically the case
- a way to draw inferences from an agent's beliefs and these rules

Example

(4) Typically birds can fly.
(5) Typically penguins cannot fly.
(6) Tweety is a penguin.
$\rightsquigarrow$ Tweety cannot fly (unless he is an atypical penguin).

The Probabilistic Answer

- Question: What is the connection between typicality and graded belief?
- Answer: What is typically the case is what is (thought to be) the case with high probability.
- Using probability theory for graded belief, a high probability directly corresponds to a high degree of belief.

(5) Typically penguins cannot fly.

$$Pr(\neg\text{fly} \mid \text{penguin}) = 1 - \epsilon,$$

where ϵ is very small, close to 0.

Rhetorical question: Is there a problem with this point of view?

The Probabilistic Answer

- Question: What is the connection between typicality and graded belief?
- Answer: What is typically the case is what is (thought to be) the case with high probability.
- Using probability theory for graded belief, a high probability directly corresponds to a high degree of belief.

(5) Typically penguins cannot fly.

$$Pr(\neg \text{fly} \mid \text{penguin}) = 1 - \epsilon,$$

where ϵ is very small, close to 0.

Rhetorical question: Is there a problem with this point of view?

The Probabilistic Answer

- Question: What is the connection between typicality and graded belief?
- Answer: What is typically the case is what is (thought to be) the case with high probability.
- Using probability theory for graded belief, a high probability directly corresponds to a high degree of belief.

(5) Typically penguins cannot fly.

$$Pr(\neg\text{fly} \mid \text{penguin}) = 1 - \epsilon,$$

where ϵ is very small, close to 0.

Rhetorical question: Is there a problem with this point of view?

The Probabilistic Answer

- Question: What is the connection between typicality and graded belief?
- Answer: What is typically the case is what is (thought to be) the case with high probability.
- Using probability theory for graded belief, a high probability directly corresponds to a high degree of belief.

(5) Typically penguins cannot fly.

$$Pr(\neg\text{fly} \mid \text{penguin}) = 1 - \epsilon,$$

where ϵ is very small, close to 0.

Rhetorical question: Is there a problem with this point of view?

The Probabilistic Answer

- Question: What is the connection between typicality and graded belief?
- Answer: What is typically the case is what is (thought to be) the case with high probability.
- Using probability theory for graded belief, a high probability directly corresponds to a high degree of belief.

(5) Typically penguins cannot fly.

$$Pr(\neg\text{fly} \mid \text{penguin}) = 1 - \epsilon,$$

where ϵ is very small, close to 0.

Rhetorical question: Is there a problem with this point of view?

The Probabilistic Answer

- Question: What is the connection between typicality and graded belief?
- Answer: What is typically the case is what is (thought to be) the case with high probability.
- Using probability theory for graded belief, a high probability directly corresponds to a high degree of belief.

(5) Typically penguins cannot fly.

$$Pr(\neg \text{fly} \mid \text{penguin}) = 1 - \epsilon,$$

where ϵ is very small, close to 0.

Rhetorical question: Is there a problem with this point of view?

A Popular Counter-Argument

Typicality $\neq$ High Frequency

Counter-Scenario

Suppose that most birds are killed by a pandemic of avian flu with the exception of penguins which are immune to the disease. Now consider: (4) Typically birds can fly.

- Some AI researchers, particularly Pollock, have the 'intuition' that a statement like (4) would still be true in that scenario.
- The intuition could be explained further by pointing out that the state of the bird-population as a whole is atypical in the scenario. In the described scenario most typical birds have died.

A Popular Counter-Argument

Typicality $\neq$ High Frequency

Counter-Scenario

Suppose that most birds are killed by a pandemic of avian flu with the exception of penguins which are immune to the disease. Now consider: (4) Typically birds can fly.

- Some AI researchers, particularly Pollock, have the 'intuition' that a statement like (4) would still be true in that scenario.
- The intuition could be explained further by pointing out that the state of the bird-population as a whole is atypical in the scenario. In the described scenario most typical birds have died.

A Popular Counter-Argument

Typicality $\neq$ High Frequency

Counter-Scenario

Suppose that most birds are killed by a pandemic of avian flu with the exception of penguins which are immune to the disease. Now consider: (4) Typically birds can fly.

- Some AI researchers, particularly Pollock, have the 'intuition' that a statement like (4) would still be true in that scenario.
- The intuition could be explained further by pointing out that the state of the bird-population as a whole is atypical in the scenario. In the described scenario most typical birds have died.

Is the Counter-Argument Convincing?

- Scenarios in which the frequency of typical individuals is low cannot be used to argue against using probability theory for representing typicality in general.
- They only speak against conflating typicality with high probability / high frequency.
- You could still opt to use to represent typicality by a probability measure.
- More importantly: Shouldn't we rather say that, given the scenario, birds typically *used to be able to fly* but now typically cannot fly? Wouldn't we revise the typicality statement in light of frequency data against it?

Is the Counter-Argument Convincing?

- Scenarios in which the frequency of typical individuals is low cannot be used to argue against using probability theory for representing typicality in general.
- They only speak against conflating typicality with high probability / high frequency.
- You could still opt to use to represent typicality by a probability measure.
- More importantly: Shouldn't we rather say that, given the scenario, birds typically *used to be able to fly* but now typically cannot fly? Wouldn't we revise the typicality statement in light of frequency data against it?

Is the Counter-Argument Convincing?

- Scenarios in which the frequency of typical individuals is low cannot be used to argue against using probability theory for representing typicality in general.
- They only speak against conflating typicality with high probability / high frequency.
- You could still opt to use to represent typicality by a probability measure.
- More importantly: Shouldn't we rather say that, given the scenario, birds typically *used to be able to fly* but now typically cannot fly? Wouldn't we revise the typicality statement in light of frequency data against it?

Is the Counter-Argument Convincing?

- Scenarios in which the frequency of typical individuals is low cannot be used to argue against using probability theory for representing typicality in general.
- They only speak against conflating typicality with high probability / high frequency.
- You could still opt to use to represent typicality by a probability measure.
- More importantly: Shouldn't we rather say that, given the scenario, birds typically *used to be able to fly* but now typically cannot fly? Wouldn't we revise the typicality statement in light of frequency data against it?

Is the Counter-Argument Convincing?

- Scenarios in which the frequency of typical individuals is low cannot be used to argue against using probability theory for representing typicality in general.
- They only speak against conflating typicality with high probability / high frequency.
- You could still opt to use to represent typicality by a probability measure.
- More importantly: Shouldn't we rather say that, given the scenario, birds typically *used to be able to fly* but now typically cannot fly? Wouldn't we revise the typicality statement in light of frequency data against it?

A More Practical Counter-Argument

In the context of modeling natural language interpretation, we'd like to ① determine the most plausible scenario (as dependent on the QUD) and ② possibly draw default inferences:

A More Practical Counter-Argument

In the context of modeling natural language interpretation, we'd like to ① determine the most plausible scenario (as dependent on the QUD) and ② possibly draw default inferences:

A More Practical Counter-Argument

In the context of modeling natural language interpretation, we'd like to ① determine the most plausible scenario (as dependent on the QUD) and ② possibly draw default inferences:

A More Practical Counter-Argument

In the context of modeling natural language interpretation, we'd like to ① determine the most plausible scenario (as dependent on the QUD) and ② possibly draw default inferences:

skydiving context

typically (in plane & wearing parachute $\rightarrow$ jump)

$\rightsquigarrow$ jump

squad context

typically (in squad & being armed $\rightarrow$ shoot)

$\rightsquigarrow$ shoot

party context — most plausible

typically (at party & late & party over $\rightarrow$ go home)

$\rightsquigarrow$ go home

Interpretation and Typicality: Overview

- Input: assumptions, background belief, literal meaning of utterance, default rules
- Output: the preferred interpretation
- Processing:

① Conditionalize the assumptions by the literal meaning to the effect that afterwards the literal meaning is believed to a certain degree $\sigma > 1/2$.

$\Downarrow$

② Find the most plausible situation.

$\Downarrow$

③ Draw inferences based on the typicality with regards to this situation.

The ϵ-based frequency approach to typicality would mix up steps ② and ③.

Interpretation and Typicality: Overview

- Input: assumptions, background belief, literal meaning of utterance, default rules
- Output: the preferred interpretation
- Processing:

① Conditionalize the assumptions by the literal meaning to the effect that afterwards the literal meaning is believed to a certain degree $\sigma > 1/2$.

⇓

② Find the most plausible situation.

⇓

③ Draw inferences based on the typicality with regards to this situation.

The ϵ-based frequency approach to typicality would mix up steps ② and ③.

Interpretation and Typicality: Overview

- Input: assumptions, background belief, literal meaning of utterance, default rules
- Output: the preferred interpretation
- Processing:

① Conditionalize the assumptions by the literal meaning to the effect that afterwards the literal meaning is believed to a certain degree $\sigma > 1/2$.

$\Downarrow$

② Find the most plausible situation.

$\Downarrow$

③ Draw inferences based on the typicality with regards to this situation.

The ϵ-based frequency approach to typicality would mix up steps ② and ③.

Interpretation and Typicality: Overview

- Input: assumptions, background belief, literal meaning of utterance, default rules
- Output: the preferred interpretation
- Processing:

① Conditionalize the assumptions by the literal meaning to the effect that afterwards the literal meaning is believed to a certain degree $\sigma > 1/2$.

$\Downarrow$

② Find the most plausible situation.

$\Downarrow$

③ Draw inferences based on the typicality with regards to this situation.

The ϵ-based frequency approach to typicality would mix up steps ② and ③.

Interpretation and Typicality: Overview

- Input: assumptions, background belief, literal meaning of utterance, default rules
- Output: the preferred interpretation
- Processing:

① Conditionalize the assumptions by the literal meaning to the effect that afterwards the literal meaning is believed to a certain degree $\sigma > 1/2$.

$\Downarrow$

② Find the most plausible situation.

$\Downarrow$

③ Draw inferences based on the typicality with regards to this situation.

The ϵ-based frequency approach to typicality would mix up steps ② and ③.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

Towards a Quantitative Model

Plans for a quantitative approach:

1. define a *probability measure* over a set of situations D_s for modeling assumptions
2. express default rules by constraints expressed in terms of a distinct *possibility measure*
3. *Jeffrey-condition* the hearer's assumptions by the literal meaning of the utterance to the effect that the literal meaning is afterwards believed to degree $f(hearer, speaker)$.
4. from the result abduce the state that is most likely to obtain
5. compute default inferences from this state

One problem with this approach is that the 'abduction' step is very limited. A more elaborate method is needed. Fortunately, there is plenty of literature on probabilistic abduction.

The End

Appendix

System P

- **LLE** left logical equivalence: If $\phi \equiv \phi'$ is a propositional tautology, then from $\phi \rightsquigarrow \psi$ infer $\phi' \rightsquigarrow \psi$
- **RW** right weakening: If $\psi \rightarrow \psi'$ is a propositional tautology, then from $\phi \rightsquigarrow \psi$ infer $\phi \rightsquigarrow \psi'$
- **REF** reflexivity: $\phi \rightsquigarrow \phi$
- **AND** right conjunction: From $\phi \rightsquigarrow \psi_1$ and $\phi \rightsquigarrow \phi_2$ infer $\phi \rightsquigarrow (\psi_1 \wedge \psi_2)$
- **OR** left disjunction: From $\phi_1 \rightsquigarrow \psi$ and $\phi_2 \rightsquigarrow \psi$ infer $(\phi_1 \vee \phi_2) \rightsquigarrow \psi$
- **CM** cautious monotonicity: From $\phi \rightsquigarrow \psi_1$ and $\phi \rightsquigarrow \psi_2$ infer $(\phi \wedge \psi_2) \rightsquigarrow \psi_1$

Lit. S. Kraus, D. Lehmann and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence 44 (1990): 167–207. This formulation was taken directly from slides by Jäger (see also Halpern (2003))

A Formal Argument Against the Probabilistic Approach

$$M \vDash A \rightsquigarrow B \quad (1)$$

$$\text{iff.} \quad Pr(B \mid A) = 1 - \epsilon \quad (2)$$

- satisfies LLE, RW, and REF
- does not satisfy AND, OR, and CM

Strict Conditioning

$$Bel(X \mid P) = \frac{Bel(X \wedge P)}{Bel(P)} \qquad (3)$$

- If we revise Bel to a new Bel' by conditioning on P, then $Bel'(P) = 1$.

Jeffrey-Conditioning

For new degree of belief α of P:

$$Bel'(X) = \alpha \frac{Bel(X \wedge P)}{Bel(P)} + (1-\alpha)\frac{Bel(X \wedge \neg P)}{Bel(\neg P)} \tag{4}$$

$$= \alpha Bel(X \mid P) + (1-\alpha) Bel(X \mid \neg P) \tag{5}$$

- Jeffrey conditioning is not commutative with respect to the order of its inputs.

Lit. Jeffrey, R.: *The Logic of Decision*. New York: McGraw-Hill 1965.

Jeffrey-Conditioning: Example

$Bel(.)$:	q	$\neg q$	
p	0.4	0.2	0.6
$\neg p$	0.3	0.1	0.4
	0.7	0.3	1

Conditioning to $Bel'(P) = 0.8$ results in:

$Bel'(.)$:	q	$\neg q$	
p	$0.5\overline{3}$	$0.2\overline{6}$	0.8
$\neg p$	0.15	0.05	0.2
	$0.68\overline{3}$	$0.31\overline{6}$	1

Field (1978) Conditioning

$$Bel'(X) = \frac{e^{\alpha} Bel(X \wedge P) + e^{-\alpha} Bel(X \wedge \neg P)}{e^{\alpha} Bel(P) + e^{-\alpha} Bel(\neg P)} \tag{6}$$

Lit. Field, H.: A Note on Jeffrey Conditionalization. *Philosophy of Science 45*, 361–367.

Probability Theory

For pairwise disjoint $X_i, X_j \subseteq W(1 \leq i, j \leq n)$:

$$Bel(X) \geq 0 \tag{7}$$

$$Bel(W) = 1 \tag{8}$$

$$Bel(\bigcup_{i=1}^{n} X_i) = \sum_{i=1}^{n} X_i \tag{9}$$

Possibility Measures

A possibility distribution is a function $\Pi : \mathcal{A} \to \mathbb{R}$, where $\mathcal{A}$ is a set of subsets of the total space W, including W, and closed under complementation and finite intersections, s.t. for every $A, B \in \mathcal{A}$:

$$\Pi(\emptyset) = 0 \tag{10}$$

$$\Pi(W) = 1 \tag{11}$$

$$\Pi(A \cup B) = max(\Pi(A), \Pi(B)) \tag{12}$$

- In infinite domains, *max* must be replaced with the supremum function.

Reference Huber (2009, 14). Lit. Huber, F.: Belief and Degrees of Belief. In *Degrees of Belief*, Springer 2009, 1–33. Zadeh, L. A.: Fuzzy Sets as a Basis for a Theory of Possibility. *Fuzzy Sets and Systems 1*, 3–28.

Possibility Measures and Default Reasoning

$$M \models A \rightsquigarrow B \tag{13}$$

$$\text{iff.} \quad \Pi(A \wedge B) > \Pi(A \wedge \neg B) \tag{14}$$

- satisfies all of system P: LLE, RW, REF, AND, OR, CM
- ‘auto-deduction principle’

Benferhat, S., Dubois, D., and Prade, H.: Nonmonotonic Reasoning, Conditional Objects and Possibility Theory, *Artificial Intelligence Journal 92*, 259–276.