A system of pragmatic pressures

Christopher Potts

UMass Amherst Linguistics

COMPSCI 591N, April 13, 2006

The birth of Gricean pragmatics

In the early 1960s, Chomsky showed us how to give compact, general specifications of natural language syntax.

In the late 1960s, philosopher and linguist H. Paul Grice had the inspired idea to do the same for (rational) social interactions.

Rules and maxims

Rules

$$
\begin{array}{rcl}
S & \Rightarrow & NP\ \ VP \\
NP & \Rightarrow & N|PN \\
N & \Rightarrow & \textit{hippo}|\ldots \\
VP & \Rightarrow & V_s\ \ S \\
VP & \Rightarrow & V_{trans}\ \ NP \\
V_s & \Rightarrow & \textit{realize}|\ldots \\
\vdots & &
\end{array}
$$

Maxims

Quality Above all, be truthful!

Relevance And be relevant!

Quantity Within those bounds, be as informative as you can!

Manner And do it as clearly and concisely as possible!

Syntactic rules are like physical laws.

Breaking them should lead to nonsense (or falsification).

Pragmatic rules (maxims) are like laws of the land.

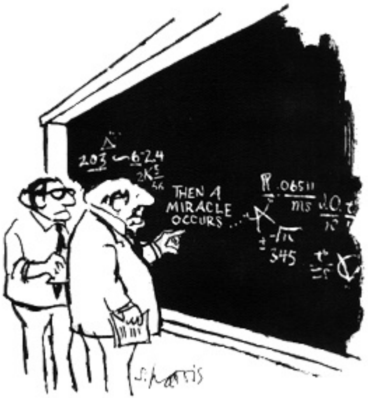
Breaking them can have noteworthy consequences.

“Then a miracle occurs”

The maxims do not yield easily to a treatment in the usual terms of semantic theory. One can usually be precise up to a point, but then . . .

"Then a miracle occurs"

The maxims do not yield easily to a treatment in the usual terms of semantic theory. One can usually be precise up to a point, but then . . .



"I think you should be more explicit here in step two."

The probability of formalizing the maxims

Some are skeptical:

- Beaver (2001:29) calls formalization in this area "notoriously problematic".
- Bach (1999) is more decisive, offering various reasons why "it seems futile for linguists to seek a formal pragmatics".
- Devitt and Sterelny (1987:§7.4) strike a similar chord.

It's a harsh verdict. Maxims (at least one) are the main engine behind all pragmatic theories.

A probable breakthrough

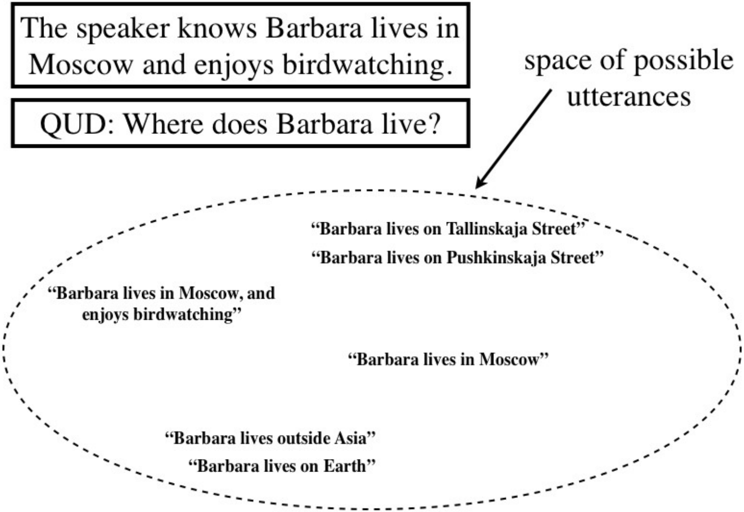
Things are looking up.

Reinhard Blutner, Gerhard Jäger, Arthur Merin, Craige Roberts, Robert van Rooij, and others have shed new light on the situation.

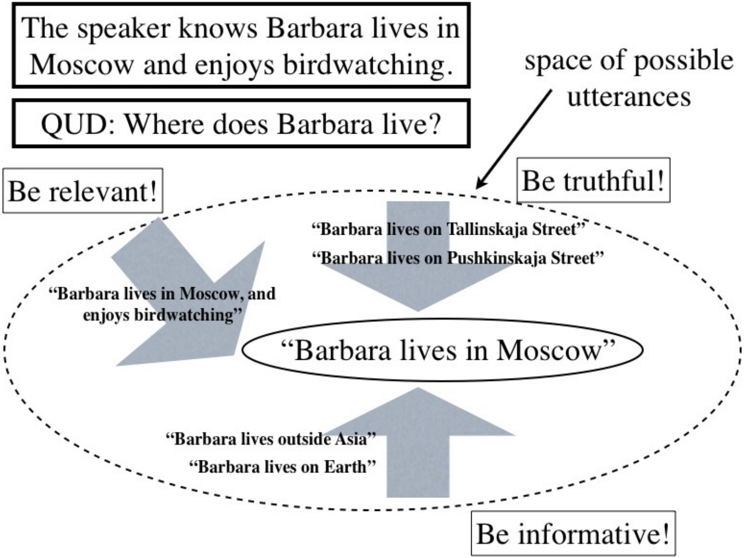
The chief innovation

A shift in emphasis from truth-conditions to probabilities.

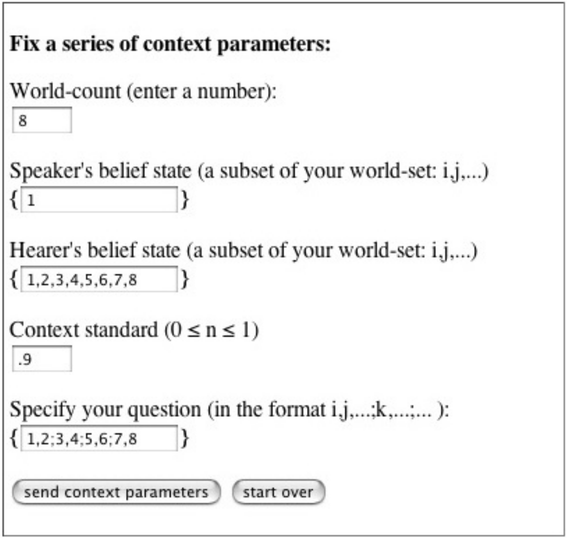
This talk in a picture



This talk in a picture



This talk as an interactive script



Fix a series of context parameters:

World-count (enter a number):

8

Speaker's belief state (a subset of your world-set: i,j,...)

{ 1 }

Hearer's belief state (a subset of your world-set: i,j,...)

{ 1,2,3,4,5,6,7,8 }

Context standard ($0 \leq n \leq 1$)

.9

Specify your question (in the format i,j,...;k,...;...):

{ 1,2;3,4;5,6;7,8 }

send context parameters | start over

This talk as an interactive script

OVERALL VALUES
[restricted to relevance-ranked propositions with quality ratings above the context standard]

Utterance	Quality	Quantity	Relevance
{1 2}	1	2.000	1
{1 2 7 8}	1	1.000	2
{1 2 5 6}	1	1.000	2
{1 2 3 4}	1	1.000	2
{1 2 5 6 7 8}	1	0.415	3
{1 2 3 4 5 6}	1	0.415	3
{1 2 3 4 7 8}	1	0.415	3
{1 2 3 4 5 6 7 8}	1	-0.000	4

PROPOSITIONS

POWER-SET({ w1 ... w8 })

THE CONTEXT STANDARD

.9 [lenient]

QUD

{ {1 2} {3 4} {5 6} {7 8} }

SPEAKER'S BELIEFS

w1 ==> 1.000
w2 ==> 0.000
w3 ==> 0.000
w4 ==> 0.000
w5 ==> 0.000
w6 ==> 0.000
w7 ==> 0.000
w8 ==> 0.000

HEARER'S BELIEFS

w1 ==> 0.125
w2 ==> 0.125

This talk in very general outline

1. Show you what the pragmatic theory looks like.
2. Show you some of the results it delivers.
3. Show you what I have done by way of computational implementation.
4. Highlight some drawbacks to the current implementation.

My primary sources

- Informativity measures relevant for quantity: Blutner 1998, 2000; van Rooy 2003c,a; Krifka 2003
- Probability distributions in linguistics: Merin 1997, 2005
- Relevance in general: van Rooy 2004a
- Relevance in interrogative contexts: van Rooy 2003b
- Overview on probabilistic and game-theoretic methods: Benz et al. 2005
- The related conceptual link with Bidirectional OT: Jäger 2002; van Rooy 2004b

Propositions as sets of possible worlds

. . . or sets of states of affairs, or sets of structured indices.

Sets, anyway.

We like the boolean structure and the off-the-shelf results from set theory.

The usefulness of propositions as sets

If φ is an expression of the language under study, then $[\![\varphi]\!]$ is its denotation (meaning).

- A proposition p is true at a world w iff $w \in p$.
- Sentence S entails sentence S' iff $[\![S]\!] \subseteq [\![S']\!]$.
- S and S' are synonymous iff $[\![S]\!] = [\![S']\!]$.
- If $[\![\textit{Lisa is smart}]\!] = A$ and Bart believes that Lisa is sick, then his belief state is a subset of A.
- The abstract goal of all discourse is to figure out which of the many possible worlds is the actual one.
- ...

Probability distributions

An additional (pragmatic) perspective on propositions.

Definition (Probability distributions)

A function $P : \wp(W) \mapsto [0,1]$ is a probability distribution iff

1. $P(W) = 1$
2. Probabilities are additive: if p and q are disjoint propositions, then $P(p \cup q) = P(p) + P(q)$

Mimicking propositions (Merin 1997, 2005)

P mimics *q*

The probability distribution P mimics the proposition q (a subset of W) iff

1. $P(\{w\}) = 0$ iff $w \notin q$
2. $P(\{w\}) = P(\{w'\})$ for all $w, w' \in q$

Mimicking propositions (Merin 1997, 2005)

P mimics *q*

The probability distribution P mimics the proposition q (a subset of W) iff

1. $P(\{w\}) = 0$ iff $w \notin q$
2. $P(\{w\}) = P(\{w'\})$ for all $w, w' \in q$

Intuitively

- By clause (i), 0 probability mimics non-membership.

Mimicking propositions (Merin 1997, 2005)

P mimics *q*

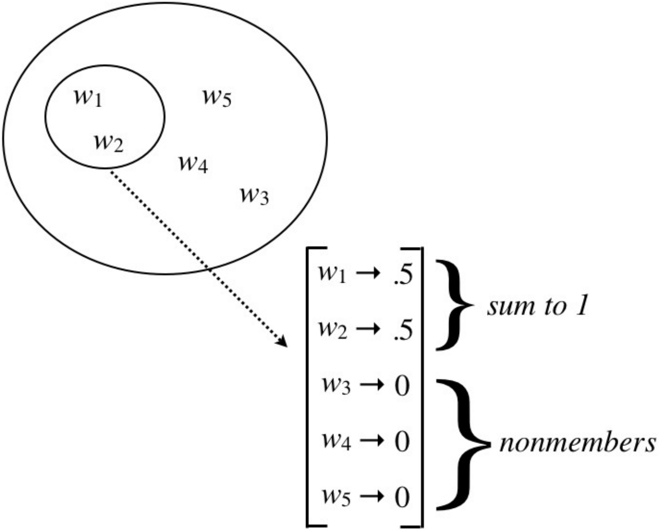
The probability distribution P mimics the proposition q (a subset of W) iff

1. $P(\{w\}) = 0$ iff $w \notin q$
2. $P(\{w\}) = P(\{w'\})$ for all $w, w' \in q$

Intuitively

- By clause (i), 0 probability mimics non-membership.
- By clause (ii), probabilities are evenly distributed across the worlds in the proposition.

Mimicking propositions (Merin 1997, 2005)



Quality (Grice 1975)

Contribute only what you know to be true. Do not say false things. Do not say things for which you lack evidence.

Quality (Grice 1975)

> Contribute only what you know to be true. Do not say false things. Do not say things for which you lack evidence.

Approximation

An utterance U by speaker s (in w) respects quality iff the semantic value of U is a superset of the set of belief worlds for s (in w).

A gradient interpretation

The above view of probability distributions is all we need for a fresh statement of Grice's quality maxim.

Our aim

A gradient view of quality, one that assigns relative values to utterances.

Quality ratings as probabilities

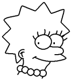
Denotations

- $[\![$*Barbara lives in Russia*$]\!] = \{w_1, w_2, w_3\}$
- $[\![$*Barbara lives in Moscow*$]\!] = \{w_1, w_2\}$
- $[\![$*Barbara lives on Tallinskaja St, Moscow*$]\!] = \{w_1\}$

$$P_{[\![\textbf{lisa}]\!]} = \begin{bmatrix} P(\{w_1\}) & \mapsto & .33 \\ P(\{w_2\}) & \mapsto & .33 \\ P(\{w_3\}) & \mapsto & .33 \\ P(\{w_4\}) & \mapsto & 0 \end{bmatrix}$$

Lisa's quality ratings

- For *Barbara lives in Russia*: 1
- For *Barbara lives in Moscow* denotes .66
- For *Barbara lives on Tallinskaja St, Moscow*: .33



Quality thresholds

Basic principle for quality-ratings

An utterance U's quality rating is $P_S([\![U]\!])$, the probability of its content for the speaker S.

$$\boxed{\text{Every context } C \text{ has a quality threshold, } C_\tau.}$$

Quality elimination

No utterance with a quality rating below the threshold can be uttered.

$$\underbrace{0 \cdots\cdots\cdots\cdots\cdots .8}_{\text{bull session}} \; \underbrace{\cdots\cdots .95 \cdots\cdots}_{\text{lenient}} \; \underbrace{\cdots .98 \cdots}_{\text{normal}} \quad \underbrace{\cdots 1}_{\text{handling explosives}}$$

Evidence for thresholds: Bullshit (Frankfurt 1986)

Frankfurt (1986)

"One of the most salient features of our culture is that there is so much bullshit."

Evidence for thresholds: Bullshit (Frankfurt 1986)

Bullshitting differs from lying

- The liar cares about the truth. He intends to convey the opposite.
- The bullshitter says things for which he has limited or no evidence, but not with the aim of deceiving. He might hope that what he is saying is true. It's just that he is not justified in his assertions (and doesn't inform you of this fact).

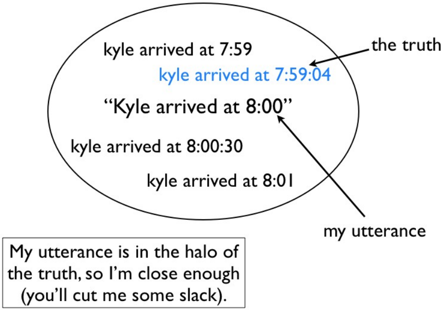
Suggestion

In bull sessions (while bullshitting) quality ratings hardly matter. The quality-rating threshold is set very low.

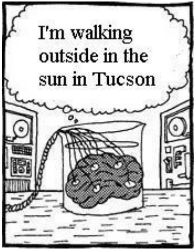
Additional sources of evidence

Epistemic particles

Pragmatic halos



Brains in vats



If quality reigned supreme

- If quality were the only pressure, speakers would restrict themselves to things with probability of 1.
- But speakers are rarely this confident about informative propositions.
- Speakers take risks with contingent truths. Why?
- They want their utterances to be relevant and contentful.
- The primary function of quality-ratings is to keep the forces of relevance and informativity from growing so powerful that they overwhelm belief.

Quantity

Make your contribution as informative as is required. Do not say more than is required.

Quantity

Make your contribution as informative as is required. Do not say more than is required.

Factoring out "required"

- The "required" portions of this maxim are duplicated by the relevance maxim.
- So quantity is a call for speakers to maximize information content.
- (This injunction is mitigated by quality and relevance.)

Measuring information content

There are many such measures on the market (van Rooy 2004a; Benz et al. 2005). I adopt a version of Blutner's (1998) proposal to derive information content from probabilities using a logarithmic function, in the mode of Carnap (1950).

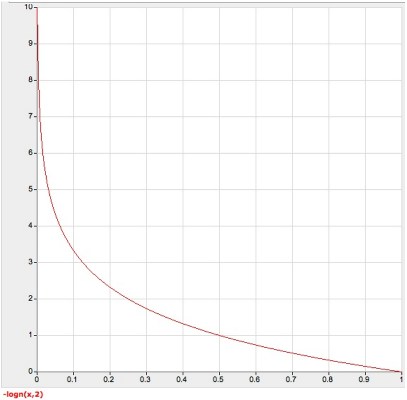
Definition (Information value of *p* for *a*)

$$\mathrm{inf}_a(p) = -\log_2 P_a(p)$$

The logarithm/probability inverse relationship

By this measure, informativity values rise as probabilities fall (with the probability of 0 assigned the pathological value ∞):

$-\log_2$



probability

Quantity ratings

- As with quality, we assign utterances *quantity ratings*.
- It won't do to identify these with the inf values for propositions relative to the speaker's probability distribution.
- On that approach, the more strongly a speaker believed a proposition *p*, the lower *p*'s information content would be.

The hearer's viewpoint

We should instead use the *hearer's* probability function in calculating the inf values relevant for pragmatics.

The hearer's viewpoint

We should instead use the *hearer's* probability function in calculating the inf values relevant for pragmatics.

To be sure, the speaker can only guess at what this probability distribution is like:

- I might accidentally tell you something you already know, on the mistaken assumption that it is new to you.
- In such cases, the information value of what I said is very low.
- You might also be insulted by my supposition about your belief state.

Quantity ratings defined

Definition (Quantity-rating)

The *quantity rating* of the proposition p in context C with addressee a is

$$\text{Quantity}_C(p) = \text{inf}_a(p)$$

Relevance

Make your contribution relevant.

Relevance

Make your contribution relevant.

Assessment

The concept is left unanalyzed.

Relevance

Make your contribution relevant.

Assessment

The concept is left unanalyzed.

We'll define relevance *to a question*

Partitions for questions

An ideal answer perfectly matches exactly one cell.

Street-level semantics

$$[\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cccc} \{w_1\} & \{w_2\} & \{w_3\} & \{w_4\} \\ \{w_5\} & \{w_6\} & \{w_7\} & \{w_8\} \end{array} \right\}$$

City-level semantics

$$[\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cc} \{w_1, w_2\} & \{w_3, w_4\} \\ \{w_5, w_6\} & \{w_7, w_8\} \end{array} \right\}$$

Country-level semantics

$$[\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{c} \{w_1, w_2, w_3, w_4\} \\ \{w_5, w_6, w_7, w_8\} \end{array} \right\}$$

Answers

Definition (Answers (van Rooy 2003b))

1. $p_Q = \{q \in Q \mid q \cap p \neq \emptyset\}$ (for p an answer to question Q)
2. $\mathrm{Ans}(p, Q) = |p_Q|$

Answers

Definition (Answers (van Rooy 2003b))

1. $p_Q = \{q \in Q \mid q \cap p \neq \emptyset\}$ (for p an answer to question Q)
2. $\text{Ans}(p, Q) = |p_Q|$

Answer to Q

- A complete answer to Q has cardinality 1 by this measure.
- Partial answers have cardinalities greater than 1.
- (Only the empty-set answer has a cardinality of 0.)

A sample calculation

$$\text{QUD} = [\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cc} \{w_1, w_2\} & \{w_3, w_4\} \\ \{w_5, w_6\} & \{w_7, w_8\} \end{array} \right\}$$

A sample calculation

$$\text{QUD} = [\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cc} \{w_1, w_2\} & \{w_3, w_4\} \\ \{w_5, w_6\} & \{w_7, w_8\} \end{array} \right\}$$

1. $[\![\textit{Barbara lives on earth}]\!] = \{w_1 \ldots w_8\}$

$$\text{Ans}([\![\textit{Barbara lives on earth}]\!], \text{QUD}) = 4$$

A sample calculation

$$\text{QUD} = [\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cc} \{w_1, w_2\} & \{w_3, w_4\} \\ \{w_5, w_6\} & \{w_7, w_8\} \end{array} \right\}$$

1. $[\![\textit{Barbara lives on earth}]\!] = \{w_1 \ldots w_8\}$

$$\text{Ans}([\![\textit{Barbara lives on earth}]\!], \text{QUD}) = 4$$

2. $[\![\textit{Barbara lives in Russia}]\!] = \{w_1 \ldots w_4\}$

$$\text{Ans}([\![\textit{Barbara lives in Russia}]\!], \text{QUD}) = 2$$

A sample calculation

$$\text{QUD} = [\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cc} \{w_1, w_2\} & \{w_3, w_4\} \\ \{w_5, w_6\} & \{w_7, w_8\} \end{array} \right\}$$

1. $[\![\textit{Barbara lives on earth}]\!] = \{w_1 \ldots w_8\}$

   $$\text{Ans}([\![\textit{Barbara lives on earth}]\!], \text{QUD}) = 4$$

2. $[\![\textit{Barbara lives in Russia}]\!] = \{w_1 \ldots w_4\}$

   $$\text{Ans}([\![\textit{Barbara lives in Russia}]\!], \text{QUD}) = 2$$

3. $[\![\textit{Barbara lives in Moscow}]\!] = \{w_1, w_2\}$

   $$\text{Ans}([\![\textit{Barbara lives in Moscow}]\!], \text{QUD}) = 1$$

A sample calculation

$$\text{QUD} = [\![\textit{Where does Barbara live?}]\!] = \left\{ \begin{array}{cc} \{w_1, w_2\} & \{w_3, w_4\} \\ \{w_5, w_6\} & \{w_7, w_8\} \end{array} \right\}$$

1. $[\![\textit{Barbara lives on earth}]\!] = \{w_1 \ldots w_8\}$

$$\text{Ans}([\![\textit{Barbara lives on earth}]\!], \text{QUD}) = 4$$

2. $[\![\textit{Barbara lives in Russia}]\!] = \{w_1 \ldots w_4\}$

$$\text{Ans}([\![\textit{Barbara lives in Russia}]\!], \text{QUD}) = 2$$

3. $[\![\textit{Barbara lives in Moscow}]\!] = \{w_1, w_2\}$

$$\text{Ans}([\![\textit{Barbara lives in Moscow}]\!], \text{QUD}) = 1$$

4. $[\![\textit{Barbara lives on Tallinskaja Street}]\!] = \{w_1\}$

$$\text{Ans}([\![\textit{Barbara lives on Tallinskaja Street}]\!], \text{QUD}) = 1$$

Relevance ranking

1. Sort the space of utterances with quality ratings above the threshold into equivalence classes based on Ans-values.
2. For each Ans-equivalence class, get the utterances with the lowest quantity ratings in that class. Keep them, and throw out the rest.
3. The Ans ordering of the remaining set is the relevance ranking.

Relevance ranking exemplified

Question Q

$$\left\{ \begin{array}{c} \{w_1, w_2\} \\ \{w_3, w_4\} \end{array} \right\} \Longrightarrow$$

Strikeouts indicate relevance-based eliminations.

Ans ordering for Q

$$\text{Ans}\,1 = \left\{ \begin{array}{cccc} \multicolumn{2}{c}{\{w_1, w_2\}} & \multicolumn{2}{c}{\{w_3, w_4\}} \\ \cancel{\{w_1\}} & \cancel{\{w_2\}} & \cancel{\{w_3\}} & \cancel{\{w_4\}} \end{array} \right\}$$

$$\text{Ans}\,2 = \left\{ \begin{array}{c} \{w_1, w_2, w_3, w_4\} \\ \cancel{\{w_1, w_2, w_3\}} \quad \cancel{\{w_1, w_2, w_4\}} \quad \cancel{\{w_2, w_3, w_4\}} \\ \cancel{\{w_1, w_3\}} \quad \cancel{\{w_1, w_4\}} \quad \cancel{\{w_2, w_3\}} \quad \cancel{\{w_2, w_4\}} \end{array} \right\} \Longrightarrow$$

Relevance ranking for Q

$$\begin{tikzcd} \{w_1, w_2\} \arrow[dr] & & \{w_3, w_4\} \arrow[dl] \\ & \{w_1, w_2, w_3, w_4\} & \end{tikzcd}$$

Felicitous utterances

We now pool the above values into a general definition, the heart of the theory.

Contexts

A *context* is a tuple

$$\left\langle P_S, P_H, Q, C_\tau, U \right\rangle$$

where P_S is the speaker's belief state (probability function), P_H is the hearer's belief state, Q is a question under discussion, C_τ is a quality threshold, and U is an utterance.

Felicitous utterances

The set of felicitous utterances for a partial context $C = \langle P_S, P_H, Q, C_\tau \rangle$ is obtained as follows:

1. From the set of all propositions, eliminate those that have quality ratings at or below C_τ.
2. With the resulting set, determine relevance rankings and throw out all utterances without such rankings.
3. From the resulting set of relevance-ranked utterances, extract the utterances with the lowest Ans values.
4. From the resulting set, select the utterances with the highest quantity ratings. These are the felicitous utterances for C.

Felicitous utterances, more concisely

$\boxed{P = \wp(W)\text{, the set of all propositions}}$

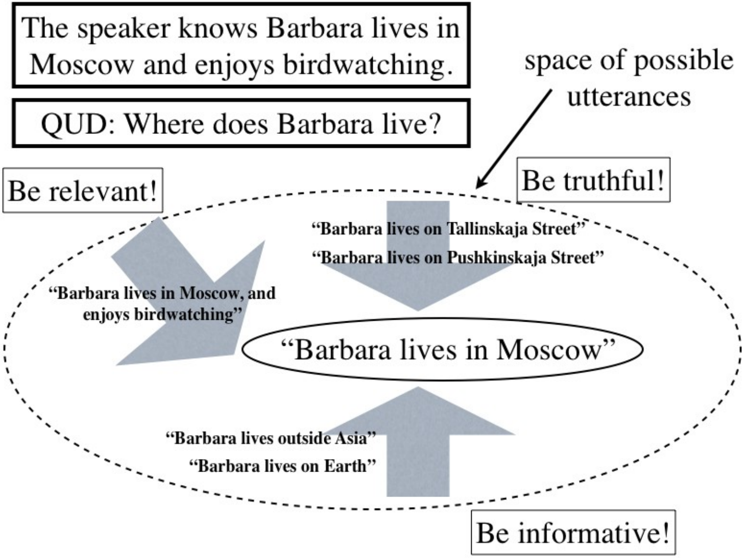
Finding the best U for $C = \langle P_S, P_H, Q, C_\tau \rangle$

Quality elimination
$P_Q = P - \{p \mid P_S(p) \leqslant C_\tau\}$ $\Longrightarrow$

Relevance elimination and Ans**-minimization**
$P_R = \text{min}_{\text{Ans}}(P_Q - \{p \mid p \text{ is not relevance ranked}\})$ $\Longrightarrow$

Quantity maximization
$P_{\text{felicitous}} = \text{max}_{\text{Quantity}}(P_R)$

Felicitous utterances, in a picture



An overview note on conversational implicatures

The conversational implicature is the central meaning designation in Gricean pragmatics.

- Conversational implicatures are propositions.
- But they are not entailments of sentences.
- Rather, they arise via a complex interaction between the context, the maxims, and the denotation of the sentence uttered.
- Slight changes to the context can cause them to disappear.

A few examples (relevance example)

1. Does Smith have a girlfriend?
2. Is Smith interviewing for a new job?
3. Where is Smith these days?

He has been spending a lot of time in New York lately.

A few examples (quantity example)

"I haven't used drugs in the last seven years."

[cf. "I've never used drugs."]

A few examples (quantity and relevance)

Can you pass the salt?

1. . . . said at the dinner table to an adult with functioning arms
2. . . . said over the phone to a person with two broken arms
3. . . . said to someone who is visiting a society in which outsiders were not allowed to touch any food

A quantity-based implicature

A: Which city does Barbara live in?

B: Barbara lives in Russia.

We perceive the conversational implicature that B is not positioned to give a complete answer to the question. Why?

Inferences about the speaker's belief state

Suppose the speaker S uttered U in a context C with question Q. The set of potential belief states for S is the set of all P_S such that

1. the speaker's utterance is above the quality threshold according to P_S (i.e., $P_S(\llbracket U \rrbracket) > C_\tau$); and
2. the speaker could not have answered A more completely with P_S (i.e., there is no utterance U' such that $P_S(\llbracket U' \rrbracket) > C_\tau$ and $\mathrm{Ans}_Q\llbracket U \rrbracket > \mathrm{Ans}_Q\llbracket U' \rrbracket$).

Output of the CGI/Perl implementation

1 => 0.500 2 => 0.500 3 => 0 4 => 0	quality-rating: 1.000 relevance-ranking: 2

HEARER ASSUMPTIONS

1. The speaker's utterance has a quality-rating above the context standard
2. The speaker has uttered something with the lowest relevance-ranking possible given his belief state.

HEARER INFERENCE

In light of the above assumptions, the speaker's belief state must be represented by one of the probability distributions at left.

INFORMATION ABOUT THE CURRENT CONTEXT	
Context standard:	.9
Utterance content:	{ 1 2 }
Relevance-ranking of utterance:	2

Question under discussion:	{
	{1}
	{2}
	{3}
	{4}
	}

A missing quantity-based implicature

A: Which city does Barbara live in?

B: Barbara lives in Moscow.

We do not infer that B has exhausted his knowledge. Perhaps he can name a street address (perhaps not).

Output of the CGI/Perl implementation

1 => 0.500 2 => 0.500 3 => 0 4 => 0	quality-rating: 1.000 relevance-ranking: 1
1 => 0 2 => 1.000 3 => 0 4 => 0	quality-rating: 1.000 relevance-ranking: 1
1 => 1.000 2 => 0 3 => 0 4 => 0	quality-rating: 1.000 relevance-ranking: 1

HEARER ASSUMPTIONS

1. The speaker's utterance has a quality-rating above the context standard
2. The speaker has uttered something with the lowest relevance-ranking possible given his belief state.

HEARER INFERENCE

In light of the above assumptions, the speaker's belief state must be represented by one of the probability distributions at left.

INFORMATION ABOUT THE CURRENT CONTEXT	
Context standard:	.9
Utterance content:	{ 1 2 }
Relevance-ranking of utterance:	1

Question under discussion:	{
	{1 2}
	{3 4}
	}

In sum: We've solved for the speaker's belief state

$Q, C_\tau, U \Rightarrow$ a set of potential belief states for the speaker

Another useful equation

$$P_H + U \Rightarrow \text{ a set of potential questions under discussion}$$

"I'm from New Jersey."

Hearer inference: the question isn't

- what country are you from?
- what continent are you from?
- what planet are you from

Where to go for the tools

Linked from:

http://people.umass.edu/potts/computation/

The scripts are CGI, so you needn't download any software or run anything on your machine.

They are written in Perl, because that's what I know best (and it's one of the few things allowed by OIT).

My goal for the implementation

A better understanding of my theory.

- The ability to change parameters quickly.
- The ability to test new examples fast.
- The ability to help others see how it works (and indicate to them that the complicated descriptions do not mean that the theory is vague or intractable).

General approach

1. Generate a space of propositions (user-supplied maximum).
2. Cut down on this space using the maxims as defined above.
   - For the first script, this yields the space of felicitous utterances.
   - For the second script, this yields a set of potential belief states for the speaker.

The concepts involved are not overly demanding

1. Generate a power set.
2. Construct and conditionalize probability distributions.
3. Compare a set of sets Q with a set p, counting the number of sets in Q with which p has a nonempty intersection.
4. Some sorting based on the above counts.

Technical drawbacks

For world spaces greater than about 8, the server hangs. The space is too big, and it grows too fast.

It's a severe limitation; logical space is composed of an absurdly large number of possible worlds. To paraphrase my grad school advisor Geoff Pullum, to call it astronomical would only reveal your ignorance about the comparatively small size of the universe.

But it's a start

And the computational work has proved linguistically illuminating in a variety of ways.

> Any and all suggestions for improving the implementation (and the theory behind it) are extremely welcome!

References I

Bach, Kent. 1999. The semantics–pragmatics distinction: What it is and why it matters. In Ken Turner, ed., *The Semantics–Pragmatics Interface from Different Points of View*, 65–84. Oxford: Elsevier. URL `http://online.sfsu.edu/~kbach/spd.htm`.

Beaver, David Ian. 2001. *Presupposition and Assertion in Dynamic Semantics*. Stanford, MA: CSLI.

Benz, Anton, Gerhard Jäger, and Robert van Rooij. 2005. An introduction to game theory for linguists. In Anton Benz, Gerhard Jäger, and Robert van Rooij, eds., *Game Theory and Pragmatics*. Houndsmills, Basingstoke, Hampshire: Palgrave Macmillan.

Blutner, Reinhard. 1998. Lexical pragmatics. *Journal of Semantics* 15(2):115–162.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17(3):189–216.

Carnap, Rudolph. 1950. *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Devitt, Michael and Kim Sterelny. 1987. *Language and Reality: An Introduction to the Philosophy of Language*. Cambridge, MA: MIT Press.

Frankfurt, Harry G. 1986. On bullshit. *Raritan* 6(2):81–100. Reprinted in Frankfurt (1988), 117–133.

Frankfurt, Harry G. 1988. *The Importance of What we Care About*. Cambridge University Press.

References II

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole and Jerry Morgan, eds., *Syntax and Semantics*, Volume 3: Speech Acts, 43–58. New York: Academic Press.

Jäger, Gerhard. 2002. Some notes on the formal properties of bidirectional Optimality Theory. *Journal of Logic, Language and Information* 11(4):427–451.

Krifka, Manfred. 2003. Polarity items in questions, Invited talk, Korean Society of Logic and Information, Inha University, Incheon, Korean, June 28.

Merin, Arthur. 1997. If all our arguments had to be conclusive, there would be few of them. Arbeitspapiere SFB 340 101, University of Stuttgart, Stuttgart. URL `http://semanticsarchive.net/Archive/jVkZDI3M/`.

Merin, Arthur. 2005. Probabalistic and deterministic presuppositions. In Rainer Bäuerle, Uwe Reyle, and Thomas Ede Zimmermann, eds., *Presuppositions and Discourse: Essays Offered to Hans Kamp*. Amsterdam: Elsevier. URL `http://semanticsarchive.net/Archive/GE0MGM5N/`.

van Rooy, Robert. 2003a. Negative polarity items in questions: Strength as relevance. *Journal of Semantics* 20:239–273.

van Rooy, Robert. 2003b. Questioning to resolve decision problems. *Linguistics and Philosophy* 26(6):727–763.

van Rooy, Robert. 2003c. Relevance and bidirectional OT. In Reinhard Blutner and Henk Zeevat, eds., *Pragmatics in Optimality Theory*, 173–210. Oxford: Palgrave Macmillan.

References III

van Rooy, Robert. 2004a. Relevance of complex sentences, Ms., University of Amsterdam.

van Rooy, Robert. 2004b. Signalling games select Horn strategies. *Linguistics and Philosophy* 27(4):493–527.