Optimal Reasoning About Referential Expressions

Judith Degen\textsuperscript{1}  Michael Franke\textsuperscript{2}

\textsuperscript{1}Department of Brain and Cognitive Sciences
University of Rochester

\textsuperscript{2}Institute for Logic, Language and Computation
Universiteit van Amsterdam

September 19, 2012
Reference to objects
Reference to objects
Reference to objects
## Production (audience design)
Clark & Murphy, 1982; Horton & Keysar, 1996; Brown-Schmidt et al., 2008

Choose a **message** to convey a given intended meaning with sufficiently high probability.

## Comprehension (perspective-taking)
Keysar et al., 2000; Hanna et al., 2003; Heller et al., 2008

Infer the most likely intended **interpretation** upon observing an utterance.
Questions

1. How much strategic back-and-forth reasoning is involved in the production and comprehension of referential expressions?
2. How well do current game-theoretic models based on rational back-and-forth reasoning about interlocutors (Franke, 2009) account for behavioral data?
Outline

1. Game-theoretic pragmatics & IBR
2. Experiment 1 - comprehension
3. Experiment 2 - production
4. Discussion
An example
An example
An example
An example
An example
Signaling games

- sequential game:
  1. the sender/speaker S wants to convey an intended meaning $t$ out of a set of possible meanings $T$ according to a certain probability distribution $p^*$
  2. S chooses a message $m$ out of a set of possible messages $M$
  3. S transmits $m$ to the receiver/hearer R
  4. R guesses an interpretation/type $t'$, based on the sent message

- if $t = t'$, both players score a point, otherwise not
Exogeneous meaning

- we assume messages have conventional or iconic meaning

\[
\begin{align*}
[[ \text{red frog} ]] & = \{ \text{red frog} \} \\
[[ \text{green frog} ]] & = \{ \text{green frog} \} \\
[[ \text{red hat} ]] & = \{ \text{red hat} \} \\
[[ \text{blue hat} ]] & = \{ \text{green frog} \text{, red hat} \}
\end{align*}
\]
**Literal receiver**

<table>
<thead>
<tr>
<th></th>
<th><img src="image1" alt="Literal 1" /></th>
<th><img src="image2" alt="Literal 2" /></th>
<th><img src="image3" alt="Literal 3" /></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0$</td>
<td><img src="image4" alt="Receiver 1" /></td>
<td><img src="image5" alt="Receiver 2" /></td>
<td><img src="image6" alt="Receiver 3" /></td>
</tr>
<tr>
<td><img src="image1" alt="Literal 1" /></td>
<td>1 0 0</td>
<td>0 0 1</td>
<td>0 1/2 1/2</td>
</tr>
<tr>
<td><img src="image2" alt="Literal 2" /></td>
<td>0 0 1</td>
<td>0 1/2 1/2</td>
<td>1 0 0</td>
</tr>
<tr>
<td><img src="image3" alt="Literal 3" /></td>
<td>1 0 0</td>
<td>0 0 1</td>
<td>0 1/2 1/2</td>
</tr>
</tbody>
</table>

![Table Diagram](image7)
### Literal sender

<table>
<thead>
<tr>
<th>$S_0$</th>
<th>![Character 1]</th>
<th>![Character 2]</th>
<th>![Character 3]</th>
<th>![Character 4]</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Character 1]</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>![Character 2]</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>![Character 3]</td>
<td>0</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>![Character 4]</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
The Iterated Best Response sequence

- $S_0$ sends any true message
- $S_1$ best response to $S_0$
- $S_2$ best response to $S_1$
- $S_i$ (for $i > 2$)
- $R_0$ interprets messages literally
- $R_1$ best response to $R_0$
- $R_2$ best response to $R_1$
- $R_i$ (for $i > 2$)
Computing best responses

- Sender: choose only messages that maximize the expected utility of $S_k$, given $R_{k-1}$
- Receiver: choose only messages that maximize the expected utility of $R_k$, given $S_{k-1}$
- expected utility is a function of
  - outcome utility
  - the players’ probabilistic beliefs about interlocutor behavior
### Iterated Best Response

<table>
<thead>
<tr>
<th>$R_1$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>![Character 1]</td>
<td>![Character 2]</td>
<td>![Character 3]</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$S_1$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>![Character 1]</td>
<td>![Character 2]</td>
<td>![Character 3]</td>
<td>![Character 4]</td>
</tr>
<tr>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Experiment 1 - comprehension

- test participants’ behavior in a comprehension task implementing previously described signaling games
- 30 participants on Amazon’s Mechanical Turk
- initially 4 trials as senders
- 36 experimental trials
  - 6 *simple* (one-step) implicature trials
  - 6 *complex* (two-step) implicature trials
  - 24 filler trials (entirely unambiguous/ entirely ambiguous target)
Simple implicature trial

The previous participant said:

Click on the creature you think the previous participant intended you to pick.

Remember the participant could only say one of these things:
Simple implicature trial - predictions

- IBR predictions for distribution of responses over target and competitor:

![Graph showing proportion of choices for target and competitor responses.](graph.png)
The previous participant said:

Click on the creature you think the previous participant intended you to pick.

Remember the participant could only say one of these things:
Complex implicature trial - predictions

IBR predictions for distribution of responses over target and competitor:

- *The previous participant said:* 
  - Click on the creature you think the previous participant intended you to pick. Remember the participant could only say one of these things:

- ![Diagram of creatures](image)

- Proportion of choices
  - Target
  - Competitor
The previous participant said:

Click on the creature you think the previous participant intended you to pick.

Remember the participant could only say one of these things:
The previous participant said:

Click on the creature you think the previous participant intended you to pick.

Remember the participant could only say one of these things:
Results - proportion of responses by condition

Proportion of choices

Response
- target
- distractor
- competitor

Degen & Franke
Reasoning About Referential Expressions
September 19, 2012
Results - proportion of responses by condition

Proportion of choices

<table>
<thead>
<tr>
<th></th>
<th>target</th>
<th>distractor</th>
<th>competitor</th>
</tr>
</thead>
<tbody>
<tr>
<td>target</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>distractor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>competitor</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Degen & Franke
Reasoning About Referential Expressions
September 19, 2012
Results - proportion of responses by condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Proportion of choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>ambiguous filler</td>
<td></td>
</tr>
<tr>
<td>complex implicature</td>
<td></td>
</tr>
<tr>
<td>simple implicature</td>
<td></td>
</tr>
<tr>
<td>unambiguous filler</td>
<td></td>
</tr>
</tbody>
</table>

Response categories:
- **Target**
- **Distractor**
- **Competitor**
Results - distribution of subjects over target choices

Not predicted by standard IBR
Experiment 2 - production

- test participants’ behavior in the analogous production task
- 30 participants on Amazon’s Mechanical Turk
- 36 experimental trials
  - 6 *simple* (one-step) implicature trials
  - 6 *complex* (two-step) implicature trials
  - 24 filler trials (entirely unambiguous/entirely ambiguous target)
Your task is to get another worker to pick out the highlighted creature. It's not highlighted on their display.

Click on one of the following four messages to send it to the other worker and get them to pick out the right creature. The other worker knows you can only send these messages.
Your task is to get another worker to pick out the highlighted creature. It's not highlighted on their display. **Click on one of the following four messages to send it to the other worker and get them to pick out the right creature.** The other worker knows you can only send these messages.
Your task is to get another worker to pick out the highlighted creature. It's not highlighted on their display. **Click on one of the following four messages to send it to the other worker and get them to pick out the right creature.** The other worker knows you can only send these messages.
Your task is to get another worker to pick out the highlighted creature. It's not highlighted on their display. **Click on one of the following four messages to send it to the other worker and get them to pick out the right creature.** The other worker knows you can only send these messages.
Results - proportion of responses by condition

Proportion of choices

Response
- target
- distractors
- competitor

ambiguous filler
complex implicature
simple implicature
unambiguous filler
Results - proportion of responses by condition

Experiment 1 (comprehension)

Experiment 2 (production)
Results - distribution of subjects over target choices

Experiment 1
(comprehension)

Experiment 2
(production)
Interim summary

- asymmetry in production and comprehension: simple implicatures easier in production than comprehension and vice versa for complex implicatures
- not predicted by standard IBR
Success expectations are given in order for

- \( R \): target, competitor, distractor object
- \( S \): target, competitor, distractor\(_1\), distractor\(_2\) message

<table>
<thead>
<tr>
<th>level</th>
<th>simple</th>
<th>complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( R )</td>
<td>( S )</td>
</tr>
<tr>
<td>1</td>
<td>( \langle 2/3, 1/3, 0 \rangle )</td>
<td>( \langle 1, 1/2, 0, 0 \rangle )</td>
</tr>
<tr>
<td>2</td>
<td>( \langle 1, 0, 0 \rangle )</td>
<td>( \langle 1, 0, 0, 0 \rangle )</td>
</tr>
<tr>
<td>3</td>
<td>( \langle 1, 0, 0 \rangle )</td>
<td>( \langle 1, 0, 0, 0 \rangle )</td>
</tr>
</tbody>
</table>
Conclusion

- Interlocutors do take perspective and simulate each others' beliefs
  - but not always optimally
  - and less so as the number of required reasoning steps increases
- IBR requires updating to allow for probabilistic rather than categorical choice rule
Future directions

- moving into the realm of actual language: manipulating message costs
- manipulating utility of communicative success / failure
- interactive experiments with feedback $\rightsquigarrow$ learning
Thanks to

- EURO-XPRAG
- Tanenhaus lab
- Mike Tanenhaus & the NIH
- Gerhard Jäger
- Florian Jaeger
Results - Exp. 1 learning effects

**Simple Implicature**

- Target: 0.8
- Distractor: 0.2
- Competitor: 0

**Complex Implicature**

- Target: 0.6
- Distractor: 0.4
- Competitor: 0.2

Relative trial number vs. proportion of choices.


