A Game-Theoretic Approach to Generating Spatial Descriptions

Dave Golland  Percy Liang  Dan Klein

Berkeley

NLP
Language is about Communication

Goal: refer to 01

[Grice, 1975]
Language is about Communication

Goal: refer to 01

**Strategy 1:** speak the truth
(Maxim of Quality)

[Grice, 1975]
Language is about Communication

Goal: refer to 01

**Strategy 1:** speak the truth
(Maxim of Quality)

*on* 03

[Grice, 1975]
Language is about Communication

Goal: refer to 01

Strategy 1: speak the truth
(Maxim of Quality)

on 03 right of 02

[Grice, 1975]
Language is about Communication

Goal: refer to 01

**Strategy 1:** speak the truth
(Maxim of Quality)

on 03  right of 02

Problem: ambiguity

[Grice, 1975]
Language is about Communication

Goal: refer to 01

**Strategy 1:** speak the truth  
(Maxim of Quality)

*on* 03  *right of* 02

Problem: ambiguity

**Strategy 2:** also be unambiguous  
(Maxims of Quality and Manner)

[Grice, 1975]
Language is about Communication

Goal: refer to O1

**Strategy 1:** speak the truth
(Maxim of Quality)

\[on \text{ O3 right of O2}\]

Problem: ambiguity

**Strategy 2:** also be unambiguous
(Maxims of Quality and Manner)

\[on \text{ O3}\]

[Grice, 1975]
Actual Example
Language Game

speaker
Language Game

speaker

listener
Language Game

speaker

listener

01

02

03
Language Game
Language Game

target $O$
speaker
utterance $w$
listener

right of 02

01
03
Language Game

\[
p_s(w | o)
\]

target \( o \) \hspace{1cm} \text{speaker} \hspace{1cm} \text{utterance} \hspace{1cm} \text{listener} \wedge

right of 02
Language Game

\[ p_s(w \mid o) \]

right of 02
Language Game

- **Target** $o$
- **Speaker** $w$
- **Listener** $g$
- **Guess** $p_s(w | o)$

**Right of 02**

- **01**
- **02**
- **03**
Language Game

$$o \xrightarrow{p_s(w \mid o)} w \xrightarrow{p_l(g \mid w)} g$$

right of 02
Language Game

\[
U(o, g) = \mathbb{I}[o = g] = \begin{cases} 
1 & \text{if } o = g \\
0 & \text{otherwise}
\end{cases}
\]
Language Game

\[
U(o, g) = \mathbb{I}[o = g] = \begin{cases} 
1 & \text{if } o = g \\
0 & \text{otherwise}
\end{cases}
\]

\[
EU(S, L) = \mathbb{E}_{S,L}[U(o, g)]
\]
Speaker Strategies

Assign scores to utterances via:

\[ p_s(w|o) \]
Speaker Strategies

Assign scores to utterances via:

\[ p_s(w|o) \]

Two speaker strategies:

1. semantics only (Maxim of Quality)
Speaker Strategies

Assign scores to utterances via:

\[ p_s(w|o) \]

Two speaker strategies:

1. semantics only \hspace{1cm} (Maxim of Quality)
2. semantics + pragmatics \hspace{1cm} (Maxims of Quality + Manner)
Game tree:

```
O1
  right of O2

O3
  on O3
```
Semantics Only

Game tree:

```
right of 02

01

on 03
```

$p_s(w|o)$ depends only on truth of utterance, does not need to take listener into account
Semantics Only

Game tree:

```
   right of 02
  /   \
01    on 03
```

$p_s(w|o)$ depends only on truth of utterance, does not need to take listener into account. Reflex speaker because it does not consider consequence of actions.
Maximize wrt. $p_s(w|o)$:

$$EU(s, L) = E_{s,L}[U(o, g)]$$
Maximize wrt. $p_S(w|o)$:

$$EU(s, l) = \mathbb{E}_{s,l}[U(o, g)]$$

$p_S(w|o)$ deterministically says:

$$\arg\max_{w'} p_L(o \mid w')$$
Maximize wrt. $p_S(w|o)$:

$$EU(S, L) = \mathbb{E}_{S,L}[U(o, g)]$$

$p_S(w|o)$ deterministically says:

$$\arg\max_{w'} p_L(o | w')$$

Needs embedded model of listener: $p_L(o|w)$
Maximize wrt. $p_S(w|o)$:

$$EU(s, l) = E_{s,l}[U(o, g)]$$

$p_S(w|o)$ deterministically says:

$$\arg\max_{w'} p_L(o | w')$$

Needs embedded model of listener: $p_L(o|w)$

01
\hspace{1cm} right
\hspace{1cm} of 02
\hspace{1cm} on 03
Maximize wrt. $p_s(w|o)$:

$$EU(s, L) = \mathbb{E}_{s, L}[U(o, g)]$$

$p_s(w|o)$ deterministically says:

$$\arg\max_{w'} p_l(o | w')$$

Needs embedded model of listener: $p_l(o|w)$
Semantics + Pragmatics

Maximize wrt. $p_s(w|o)$:

$$EU(s, L) = \mathbb{E}_{s,L}[U(o, g)]$$

$p_s(w|o)$ deterministically says:

$$\arg\max_{w'} p_L(o \mid w')$$

Needs embedded model of listener: $p_L(o|w)$

Rational speaker because it is optimal with respect to given listener.
Reflex vs. Rational

Reflex (semantics only)

Rational (semantics + pragmatics)
Reflex vs. Rational

\[ S \]  \hspace{1cm}  \[ S(L) \]

Reflex  
(semantics only)  

Rational  
(semantics + pragmatics)
Experimental Setup

Google sketchup: 43 rooms, average of 22 objects per room
Question: Where is the object outlined in red?
Answer: The object outlined in red is

Speaker: $o \rightarrow w$
Data Collection with Mechanical Turk

**Question:** Where is the object outlined in red?

**Answer:** The object outlined in red is left of

Yields annotated data:

\[ \{(o_1, w_1), \cdots, (o_n, w_n)\} \]
Question: What object is right of O2?

Listener: w → g
Evaluation with Mechanical Turk

Given $p_s(w|o)$, and $(o_1, \ldots, o_n)$,
Evaluation with Mechanical Turk

Given \( p_s(w|o) \), and \( (o_1, \ldots, o_n) \),

\( s \) generates:

\[
(w_1, \ldots, w_n)
\]
Evaluation with Mechanical Turk

Given \( p_s(w|o) \), and \((o_1, \ldots, o_n)\),

\( S \) generates:

\[(w_1, \ldots, w_n)\]

where:

\[w_i = \arg\max_w p_s(w|o_i)\]
Evaluation with Mechanical Turk

Given \( p_s(w|o) \), and \((o_1, \ldots, o_n)\),

\( s \) generates:

\[(w_1, \ldots, w_n)\]

where:

\[w_i = \arg\max_w p_s(w|o_i)\]

turkers generate:

\[(g_1, \ldots, g_n)\]
Evaluation with Mechanical Turk

Given \( p_S(w|o) \), and \((o_1, \ldots, o_n)\),

\( S \) generates:

\[
(w_1, \ldots, w_n)
\]

where:

\[
w_i = \arg\max_w p_S(w|o_i)
\]

turkers generate:

\[
(g_1, \ldots, g_n)
\]

compute success metric:

\[
\text{SUCCESS}(s) = \frac{1}{n} \sum_i \mathbb{I}[o_i = g_i]
\]
Evaluation with Mechanical Turk

Given $p_S(w|o)$, and $(o_1, \ldots, o_n)$, $s$ generates:

$$(w_1, \ldots, w_n)$$

where:

$$w_i = \arg\max_w p_S(w|o_i)$$

turkers generate:

$$(g_1, \ldots, g_n)$$

compute success metric:

$$\text{SUCCESS}(s) = \frac{1}{n} \sum_i \mathbb{I}[o_i = g_i]$$

Note: only collecting data and evaluating are done by humans.
Results

**LITERAL** agents put mass uniformly on true outputs

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflex</td>
<td>4.6%</td>
</tr>
<tr>
<td>Rational</td>
<td>33.7%</td>
</tr>
</tbody>
</table>

Rational speaker outperforms reflex speaker.
Key Points

Rational speaker $S(L)$ outperforms reflex speaker $S$. 
Key Points

Rational speaker $S(L)$ outperforms reflex speaker $S()$

Rational speaker $S(L)$ is optimal with respect to listener $L$
Key Points

Rational speaker $S(L)$ outperforms reflex speaker $S()$

Rational speaker $S(L)$ is optimal with respect to listener $L$

To improve rational speaker $S(L)$ must improve embedded listener $L$
Key Points

Rational speaker $S(L)$ outperforms reflex speaker $S()$. 

Rational speaker $S(L)$ is optimal with respect to listener $S(L)$. 

To improve rational speaker $S(L)$ must improve embedded listener $L$. 

Up next: extensions for improving the listener model.
Listener Extensions

- Training a Listener
- Generating Complex Utterances
- Modeling Listener Confusion
Listener Extensions

• Training a Listener
• Generating Complex Utterances
• Modeling Listener Confusion
Learned Listener Model

Before: listener was LITERAL
Learned Listener Model

**Before:** listener was **LITERAL**

**Now:** learn from mturk data:

\[\{(o_1, w_1), \ldots, (o_n, w_n)\}\]
Learned Listener Model

Before: listener was LITERAL
Now: learn from mturk data:
\[
\{(o_1, w_1), \ldots, (o_n, w_n)\}
\]

Train a log-linear model:
\[
p_{\text{LEARNED}}(g|w; \theta_L) \propto \exp\{\theta_L^\top \phi(g, w)\}
\]
Learned Listener Model

Before: listener was LITERAL

Now: learn from mturk data:
\[
\{(o_1, w_1), \ldots, (o_n, w_n)\}
\]

Train a log-linear model:
\[
p_{\text{LEARNED}}(g|w; \theta_L) \propto \exp\{\theta_L^\top \phi(g, w)\}
\]

Use it to define the rational $S(L)$ LEARNED speaker
Learned Listener Model

Before: listener was LITERAL
Now: learn from mturk data:
\[ \{(o_1, w_1), \ldots, (o_n, w_n)\} \]

Train a log-linear model:
\[ p_{\text{LEARNED}}(g|w; \theta_L) \propto \exp\{\theta_L^\top \phi(g, w)\} \]

Use it to define the rational \( S(L) \) \( \text{LEARNED} \) speaker

We also train \( p_S(w|o; \theta_S) \) using the same data and features
to get the reflex \( S(L) \) \( \text{LEARNED} \) speaker
Features

The features $\phi(g, w)$ are defined between:
Features

The features $\phi(g, w)$ are defined between:

- guess object $g$

Features inspired by [Regier, 2001; Tellex, 2009; Landau, 1993]
Features

The features $\phi(g, w)$ are defined between:

- guess object $g$
- $w = \text{right of } 02$

Features inspired by [Regier, 2001; Tellex, 2009; Landau, 1993]
Features

The features $\phi(g, w)$ are defined between:

- guess object $g$
- $w = \text{right of } 02$

Features inspired by [Regier, 2001; Tellex, 2009; Landau, 1993]
Features

The features $\phi(g, w)$ are defined between:

- guess object $g$
- $w = \text{right of } 02$

$w.r$ $w.o$

- $g$ and $w.o$ are bounding boxes

Features inspired by [Regier, 2001; Tellex, 2009; Landau, 1993]
Distance Features
Distance Features

\[ \phi_{\text{dist}} = \text{value of shortest distance between } g \text{ and } w.o \]
\[ \phi_{\text{top}1} = \mathbb{I}[g \text{ is closest to } w.o] \]
\[ \phi_{\text{top}5} = \mathbb{I}[g \text{ is among top 5 closest to } w.o] \]
\[ \phi_{\text{top}10} = \mathbb{I}[g \text{ is among top 10 closest to } w.o] \]
Containment Features

\[ g \]

\[ w.o \]
\[ \phi_{cont_2} = \frac{\text{vol}(w.o \cap g)}{\text{vol}(g)} \]

\[ \phi_{cont_1} = \frac{\text{vol}(w.o \cap g)}{\text{vol}(w.o)} \]
Projection Features

\[ g \]

\[ w.o \]

\[ 21 \]
Projection Features
Projection Features

\[ v \]
Projection Features
Projection Features
Projection Features
**Projection Features**

\[
\begin{align*}
\phi_{projx} &= f_x \\
\phi_{projy} &= f_y \\
\phi_{projz} &= f_z \\
\phi_{proj1} &= \mathbb{I}[f_x = \max\{f_x, f_y, f_z\}] \\
\phi_{proj2} &= \mathbb{I}[f_y = \max\{f_x, f_y, f_z\}] \\
\phi_{proj3} &= \mathbb{I}[f_z = \max\{f_x, f_y, f_z\}]
\end{align*}
\]
# Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflex S</td>
<td>4.6%</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>33.7%</td>
</tr>
</tbody>
</table>

**Literal**
## Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflex</td>
<td><img src="4.6%25" alt="Literal" /></td>
</tr>
<tr>
<td>Rational</td>
<td><img src="33.7%25" alt="Literal" /></td>
</tr>
<tr>
<td>Reflex</td>
<td><img src="38.4%25" alt="Learned" /></td>
</tr>
<tr>
<td>Rational</td>
<td><img src="52.6%25" alt="Learned" /></td>
</tr>
</tbody>
</table>
What’s missing?

Two things are missing from the setup so far.
What’s missing?

Two things are missing from the setup so far.
1. Arbitrary descriptors
What’s missing?

Two things are missing from the setup so far.

1. Arbitrary descriptors

![Diagram showing a table with a vase and a lamp, with labels O1, O2, and O3.]
What’s missing?

Two things are missing from the setup so far.

1. Arbitrary descriptors (not today)
What’s missing?

Two things are missing from the setup so far.
1. Arbitrary descriptors (not today)

We will not be seeing 100% in this talk.

2.
What’s missing?

Two things are missing from the setup so far.

1. Arbitrary descriptors (not today)

2. Complex utterances

We will not be seeing 100% in this talk.

2. Complex utterances
What’s missing?

Two things are missing from the setup so far.
1. Arbitrary descriptors (not today)
2. Complex utterances (coming up)

We will not be seeing 100% in this talk.
Listener Extensions

- Training a Listener
- Generating Complex Utterances
- Modeling Listener Confusion
Complex Utterances

Before: utterances were simple, such as:

right of O2 on O3
Complex Utterances

Before: utterances were simple, such as:

\[ \text{right of O2} \quad \text{on O3} \]

Now: utterances are from grammar:
Complex Utterances

Before: utterances were simple, such as:

\[ \text{right of O2 on O3} \]

Now: utterances are from grammar:

\[ \text{[noun]} \quad N \rightarrow \text{something} \mid 01 \mid 02 \mid \cdots \]
Complex Utterances

Before: utterances were simple, such as:

\[
\text{right of } O2 \quad \text{on } O3
\]

Now: utterances are from grammar:

\[
\begin{align*}
\text{[noun]} & \quad N \rightarrow \text{something} \mid 01 \mid 02 \mid \cdots \\
\text{[relation]} & \quad R \rightarrow \text{on} \mid \text{right of} \mid \cdots
\end{align*}
\]
Complex Utterances

Before: utterances were simple, such as:

\[\text{right of } 02 \quad \text{on } 03\]

Now: utterances are from grammar:

\[
\begin{align*}
\text{noun} & : \quad N \to something \mid 01 \mid 02 \mid \cdots \\
\text{relation} & : \quad R \to on \mid right of \mid \cdots \\
\text{relativization} & : \quad RP \to R \, NP
\end{align*}
\]
Complex Utterances

Before: utterances were simple, such as:

right of 02 on 03

Now: utterances are from grammar:

[noun] N → something | 01 | 02 | ···
[relation] R → on | right of | ···
[relativization] RP → R NP

right of 02
Complex Utterances

Before: utterances were simple, such as:

\[ \text{right of 02} \quad \text{on 03} \]

Now: utterances are from grammar:

- **noun**
  \[ N \rightarrow \text{something} \mid 01 \mid 02 \mid \cdots \]
- **relation**
  \[ R \rightarrow \text{on} \mid \text{right of} \mid \cdots \]
- **relativization**
  \[ \text{RP} \rightarrow \text{R NP} \]

\[ \text{on something right of 02} \]
Complex Utterances

Before: utterances were simple, such as:

\[ \text{right of 02} \quad \text{on 03} \]

Now: utterances are from grammar:

- [noun] \quad N \rightarrow \text{something} \mid 01 \mid 02 \mid \cdots
- [relation] \quad R \rightarrow \text{on} \mid \text{right of} \mid \cdots
- [relativization] \quad \text{RP} \rightarrow R \text{NP}
- [conjunction] \quad \text{NP} \rightarrow N \text{RP}^*
Complex Utterances

Before: utterances were simple, such as:

\[
\text{right of 02} \quad \text{on 03}
\]

Now: utterances are from grammar:

- [noun] \( N \rightarrow \text{something} \mid 01 \mid 02 \mid \cdots \)
- [relation] \( R \rightarrow \text{on} \mid \text{right of} \mid \cdots \)
- [relativization] \( RP \rightarrow R \ NP \)
- [conjunction] \( NP \rightarrow N \ RP^* \)

\[
\text{right of 02 and on 03}
\]
Example Interpretation

Computing: $p(g \mid on\ something\ right\ of\ 02)$
Example Interpretation

Computing: \( p(g \mid \text{on something right of } 02) \)

If \( w \) is rooted at \( N \), \( p_L(g \mid w) = \mathbb{I}[w = g] \).
Example Interpretation

Computing: $p(g \mid on\ something\ right\ of\ 02)$

If $w$ is rooted at $RP$, recurse on $NP$ subtree, use base listener.
Example Interpretation

Computing: $p(g \mid \text{on something right of 02})$

If $w$ is rooted at $N$ and $w = \text{something}$, $p_L(g \mid w)$ is uniform.
Example Interpretation

Computing: $p(g \mid \text{on something right of } 02)$

If $w$ is rooted at $\text{NP}$, recurse on children, multiply and renormalize.
Example Interpretation

Computing: \( p(g \mid on \ something \ right \ of \ O2) \)

If \( w \) is rooted at \( RP \), recurse on \( NP \) subtree, use base listener.
## Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflex S</td>
<td>38.4%</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>52.6%</td>
</tr>
</tbody>
</table>

LEARNED
## Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflex S(□)</td>
<td>LEARNED</td>
</tr>
<tr>
<td>Reflex S(□)</td>
<td>LEARNED</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>LEARNED</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>LEARNED</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>LEARNED</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>LEARNED</td>
</tr>
<tr>
<td>Success</td>
<td></td>
</tr>
<tr>
<td>Reflex S(□)</td>
<td>38.4%</td>
</tr>
<tr>
<td>Reflex S(□)</td>
<td>52.6%</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>52.6%</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>51.0%</td>
</tr>
<tr>
<td>Rational S(L)</td>
<td>51.0%</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflex</td>
<td>Learned</td>
</tr>
<tr>
<td>Rational</td>
<td>Learned</td>
</tr>
<tr>
<td>Rational</td>
<td>Learned compositional</td>
</tr>
</tbody>
</table>

Problem: introducing complex utterances hurts success
Listener Confusion

Observations: success is lower
Listener Confusion

Observations: success is lower & all utterances are longer
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

Right of the lamp
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

Right of the lamp and on the table
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

*Right of the lamp and on the table and below the ceiling*
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

*Right of the lamp and on the table and below the ceiling and in the room*
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

Right of the lamp and on the table and below the ceiling and in the room and etc.
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

*Right of the lamp and on the table and below the ceiling and in the room and etc.*

Maxim of manner: also be brief
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

*Right of the lamp and on the table and below the ceiling and in the room and etc.*

Maxim of manner: also be brief

- saves time
Listener Confusion

Observations: success is lower & all utterances are longer

Is longer always better?

Right of the lamp and on the table and below the ceiling and in the room and etc.

Maxim of manner: also be brief
• saves time
• prevents confusion
Listener Extensions

- Training a Listener
- Generating Complex Utterances
- Modeling Listener Confusion
Modeling Listener Confusion

Problem: our model does not match turkers

\[ \tilde{p}(g|w) = \alpha|w| p(g|w) \uparrow \downarrow \]
understand

+ \( (1 - \alpha|w|) p_{\text{rnd}}(g|w) \uparrow \downarrow \)

confused
Modeling Listener Confusion

Problem: our model does not match turkers
Confused turkers guess randomly.

\[ \tilde{p}(g|w) = \alpha |w| p(g|w) + (1-\alpha |w|) p_{rnd}(g|w) \]
Modeling Listener Confusion

**Problem:** our model does not match turkers

Confused turkers guess randomly.

\[ \tilde{p}(g|w) = \alpha |w| p_l(g|w) \]

\[ + (1-\alpha |w|) p_{rnd}(g|w) \]

\[ = \underbrace{\text{understand}}_{\text{confused}} \]

\[ 30 \]
Modeling Listener Confusion

Problem: our model does not match turkers
Confused turkers guess randomly.

\[ \tilde{p}(g|w) = \alpha |w| p(l(g|w)) + (1 - \alpha |w|) p_{\text{rnd}}(g|w) \]

\[ \text{understand} \]
\[ \text{confused} \]
Modeling Listener Confusion

Problem: our model does not match turkers
Confused turkers guess randomly.

\[ \tilde{p}_l(g|w) = \alpha |w| p_l(g|w) + (1 - \alpha |w|) p_{rnd}(g|w) \]
Modeling Listener Confusion

Problem: our model does not match turkers
Confused turkers guess randomly.

\[
\alpha|w| p_L(g|w) + (1 - \alpha|w|) p_{\text{rnd}}(g|w)
\]

\(\alpha\) understand
Modeling Listener Confusion

**Problem:** our model does not match turkers

Confused turkers guess randomly.

\[
\tilde{p}(g|w) = \alpha |w| p_L(g|w) + (1 - \alpha |w|) \]

\text{understand}
Modeling Listener Confusion

Problem: our model does not match turkers

Confused turkers guess randomly.

\[
\tilde{p}_l(g \mid w) = \alpha \left| w \right| p_l(g \mid w) + (1 - \alpha \left| w \right|) p_{rnd}(g \mid w)
\]

understand confused
Modeling Listener Confusion

Problem: our model does not match turkers
Confused turkers guess randomly.

\[
\tilde{p}_L(g \mid w) = \alpha^{|w|} p_L(g \mid w) + (1 - \alpha^{|w|}) p_{\text{rnd}}(g \mid w)
\]

understand confused
## Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational S(L) Learned</td>
<td>52.6%</td>
</tr>
<tr>
<td>Rational S(L) Learned compositional</td>
<td>51.0%</td>
</tr>
</tbody>
</table>
### Results

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational LEARNED</td>
<td>52.6%</td>
</tr>
<tr>
<td>Rational LEARNED compositional</td>
<td>51.0%</td>
</tr>
<tr>
<td>Rational LEARNED +confusion model</td>
<td>54.5%</td>
</tr>
</tbody>
</table>
Final Remarks

Observation: people avoid saying ambiguous utterances
Final Remarks

Observation: people avoid saying ambiguous utterances

- Grice (1975) expressed this observation in his maxims
Final Remarks

Observation: people avoid saying ambiguous utterances
- Grice (1975) expressed this observation in his maxims
- Parikh (2001), Stalnaker (2005), and Jäger (2008) explore relation between pragmatic models and game theory
Final Remarks

Observation: people avoid saying ambiguous utterances

• Grice (1975) expressed this observation in his maxims

• Parikh (2001), Stalnaker (2005), and Jäger (2008) explore relation between pragmatic models and game theory

Meanwhile:
Landau (1993), Regier (2001), and Tellex (2009) have been studying spatial language
Final Remarks

Observation: people avoid saying ambiguous utterances
• Grice (1975) expressed this observation in his maxims
• Parikh (2001), Stalnaker (2005), and Jäger (2008) explore relation between pragmatic models and game theory

Meanwhile:
Landau (1993), Regier (2001), and Tellex (2009) have been studying spatial language

Our contribution: we show how a game theoretic pragmatics model can be used to successfully generate spatial descriptions
Thank you!