Using Language Strategically in Context

Daniel Fried
Using Language in Context

**Grounding**

“Take me to the airport”

**Pragmatics**

“My neck hurts”
Using Language in Context

Grounding

“Stop at the second car”

Pragmatics

“Stop at the car”
Using Language Strategically

[Diagram: Generation and Interpretation with arrows indicating language and actions, such as "stop at the car" and "stop at the second car".]

References:
[e.g. Lewis 1969; Golland et al. 2010; Frank and Goodman 2012; Degen et al. 2013]
Using Language Strategically

**Generation**

**Speaker**

- stop at the car

**Listener**

- stop at the second car

**Interpretation**

**Speaker**

- stop at the car

**Listener**

- stop at the second car

- stop at the car

---

[e.g. Lewis 1969; Golland et al. 2010; Frank and Goodman 2012; Degen et al. 2013]
Using Language Strategically

**Generation**

Speaker

Listener

stop at the car

Listener

stop at the second car

**Interpretation**

Speaker

stop at the car

Listener

stop at the second car

Listener

stop at the car

Listener

Listener

stop at the car

[e.g. Lewis 1969; Golland et al. 2010; Frank and Goodman 2012; Degen et al. 2013]
Reasoning with Speakers and Listeners

stop at the second car
Strategic Language for...

Grounded Instructions

[Fried, Andreas, & Klein. NAACL 2018]

Negotiation Games

[FAIR Diplomacy Team. Science 2022]
Turn left and take a right at the table. Take a left at the painting and then take your first right. Wait next to the exercise equipment.

Discrete motion, but real images

[Chang et al. 2018]
Interpretation Task

Input instructions:

Go forward between the kitchen counters and then turn right into the living room. Walk forward onto the rug.

Output a route:
Go forward between the kitchen counters and then turn right into the living room. Walk forward onto the rug.
Models of Listeners and Speakers

Inputs

*Go forward between the kitchen counters...*

Instruction, \( i \)

**Listener**

\[ P_L(r \mid i) \]

**Speaker**

\[ P_S(i \mid r) \]

Outputs

*Route, \( r \)*

*Go forward between the kitchen counters...*

Instruction, \( i \)
LSTM Encoder

go past the couch

LSTM Decoder with Attention

[Anderson et al., 2018]
Base Speaker Model

LSTM Encoder

LSTM Decoder with Attention

\texttt{<START>} \quad go \quad past \quad the \quad couch \quad ...

[Fried et al., 2018]
Training Models on Human Instructions

Speaker

Fit Model

$\argmax_\theta P_S(i \mid r; \theta)$

Instruction

Go forward between the kitchen counters...

Human annotators

Fit Model

$\argmax_\phi P_L(r \mid i; \phi)$

Listener

Route
walk past the dining room table and chairs and take a right into the living room.

walk past the dining room table and chairs and take a right into the living room. stop once you are on the rug.
walk past the dining room table and chairs and take a right into the living room.

walk past the dining room table and chairs and take a right into the living room.
stop once you are on the rug.
Speaker

walk past the dining room table and chairs and take a right into the living room.

Listener

0.4

walk past the dining room table and chairs and take a right into the living room. stop once you are on the rug.

Listener

0.8

$P_L(r \mid i, \ldots)$
Generation Results

Human accuracy at following instructions from:

- Other humans
- Base speaker
- Pragmatic speaker

<table>
<thead>
<tr>
<th>Task</th>
<th>Other humans</th>
<th>Base speaker</th>
<th>Pragmatic speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAIL</td>
<td>62.8</td>
<td>75.2</td>
<td>73.2</td>
</tr>
<tr>
<td>Alchemy</td>
<td>66.0</td>
<td>75.3</td>
<td>83.3</td>
</tr>
<tr>
<td>Scene</td>
<td>29.3</td>
<td>31.3</td>
<td>69.3</td>
</tr>
<tr>
<td>Tangrams</td>
<td></td>
<td>60.0</td>
<td>66.0</td>
</tr>
</tbody>
</table>

[Fried et al., NAACL 2018]
Simulating people’s interpretations makes language more informative.
Pragmatic Interpretation

Listener

walk past the dining room table and chairs and take a right into the living room.

Speaker

walk past the dining room table and chairs and take a right into the living room.

walk past the dining room table and chairs and take a right into the living room. stop once you are on the rug.
Pragmatic Interpretation

Listener
walk past the dining room table and chairs and take a right into the living room.

Speaker
walk past the dining room table and chairs and take a right into the living room.
walk past the dining room table and chairs and take a right into the living room. stop once you are on the rug.
Pragmatic Interpretation

Speaker

walk past the dining room table and chairs and take a right into the living room.

Listener

walk past the dining room table and chairs and take a right into the living room.

walk past the dining room table and chairs and take a right into the living room. stop once you are on the rug.
walk past the dining room table and chairs and take a right into the living room.

walk past the dining room table and chairs and take a right into the living room.

walk past the dining room table and chairs and take a right into the living room.

stop once you are on the rug.
Pragmatic Interpretation

Success rate at following human directions

Human performance

Wang et al. 2018: 25
Our Base Listener: 36
+Pragmatics: 54

[Fried*, Hu*, Cirik* et al., NeurIPS 2018]
Walk past hall table. Walk into bedroom. Make left at table clock. 
Wait at bathroom door threshold.
Walk past hall table. Walk into bedroom. Make left at table clock. Wait at bathroom door threshold.
Simulating why a speaker said what they did helps resolve ambiguity.
Strategic Language for...

Grounded Instructions

[Fried, Andreas, & Klein. NAACL 2018]

Negotiation Games

[Cicero; FAIR Diplomacy Team. Science 2022]
Diplomacy is a popular strategy board game from the 50s (JFK’s favorite game!). You’re one of 7 powers trying to conquer Europe in WW1.

Each turn has two parts:
- 1) Private natural language negotiation
- 2) Simultaneous actions

Winning requires both strategic play and the ability to persuade and cooperate.

Long-standing challenge for AI; past work has all been on language-free variants.
Challenges for Language in Diplomacy

(1) Hold **human-like** dialogue
   ▶ Hold long conversations of **hundreds of messages** with multiple parties
   ▶ Messages must be **grounded** in both the environment and dialogue history

(2) **Win the game** through strategic negotiation
   ▶ Interface with **external reasoning systems**
   ▶ Use language **intentionally** to achieve goal

Tension between these goals!
Imitation Learning

- Fine tuning language models works great for generating human-like dialogue
  - *Let’s do as much with it as possible*

- Use data from **50,000 human games** of Diplomacy acquired through a partnership with webdiplomacy.net

- **Fine-tune language models** on human dialogue
Sampling from even a perfect imitation learning model only reproduces *average play*

- There are many *weak players* in our dataset
- If our messages describe bad actions, no-one will cooperate with us

*We can’t beat the best humans by copying them*
Beyond Imitation Learning

- Decouple action planning and language generation
  - Another talk!

- Use a language-informed planner
  - Condition agent’s language on agent’s plans
  - Reason about the effect of agent’s language on other players’ plans
ENGLAND: Do you want NTH to support BEL?
FRANCE: No, BEL is moving to HOL.
ENGLAND: Alright i’ll support you in.

Model-Based Annotation:
- FRA: BEL -> HOL
- ENG: NTH S BEL -> HOL

Dialogue Conditioning:
- FRA: BEL -> HOL
- ENG: NTH S BEL -> HOL

Dialogue Model:
- ENGLAND: Alright i’ll support you in.

During Play:
- Planner
- FRA: BRE -> MAO
- ENG: LON -> NTH

Dialogue Model:
- ENGLAND: Okay, I’ll move to North Sea then.
**Actions**

- **France** helps **Cicero (England)** to Belgium
- **Cicero** supports **France** to Belgium

**Message**

- **ENG → FRA**: Mind supporting Edi - Bel?
- **ENG → FRA**: Let me know if you want me to support you to Bel, otherwise I'll probably poke Hol.
Honesty is (normally) the best policy

- Hard to account for the impact of lying
- Conversations after deceit are difficult
- Strong players use lying sparingly

=> Agent honestly describes **actions it intends to take**

RUSSIA -> TURKEY: You stabbed me. I will let France solo.
RUSSIA -> TURKEY: Every turn you continue to stab me, I will keep helping him with all my might.
RUSSIA -> TURKEY: Repent and we can continue as we were.
TURKEY -> RUSSIA: You said I could take Rumania. Are you saying that was a lie?
RUSSIA -> TURKEY: You said Black Sea would take it
RUSSIA -> TURKEY: You lied
RUSSIA -> TURKEY: And you know it
RUSSIA -> TURKEY: Don't play me for a fool
RUSSIA -> TURKEY: You lose until you repent
Encouraging Strategic Honesty

**Planner**
- **FRA**: BRE -> MAO
- **ENG**: LON -> NTH …

**Dialogue Model**
- **ENGLAND**: Okay, I’ll move to North Sea then.
- Russia’s help.

**Intent Predictor**
- **FRA**: BRE -> MAO
- **ENG**: LON -> NTH …

**Filtering Ensemble**
- Δ𝔼[value] non-negative?
- Free of grounding errors?

✓

Non-grounding errors?
Cicero’s Dialogue (as FRANCE)

FRANCE: I’ll work with you but I need Tunis for now.

TURKEY: Nope, you gotta let me have it

FRANCE: No, I need it.

TURKEY: You have Serbia and Rome to take.

FRANCE: They're impossible targets

TURKEY: Greece - Ionian Ionian - Tyr

FRANCE: Then in fall you take Rome and Austria collapses.
This summer, we entered Cicero anonymously in an online Diplomacy league.

Cicero placed in the top 10% of players, and 2nd of 19 players who played ≥5 games. It achieved more than double the average human score.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Avg Score</th>
<th># Games</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35.0%</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>25.8%</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>24.5%</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>22.7%</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>21.0%</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>3.0%</td>
<td>6</td>
</tr>
<tr>
<td>20</td>
<td>2.6%</td>
<td>7</td>
</tr>
</tbody>
</table>
Takeaways from Cicero

- Decoupling strategic planning and language generation
  - Outperform imitation learning while still cooperating with people

- Strategy informs language

- Language is used strategically
Overall Takeaways

Language is a cooperative, multiagent process.

Language systems improve when they treat communication strategically.
Grounding Collaborators

Jacob Andreas
Taylor Berg-Kirkpatrick
Justin Chiu
Volkan Cirik
Trevor Darrell
Ronghang Hu

Dan Klein
Louis-Philippe Morency
Anna Rohrbach
Kate Saenko
Sheng Shen
Thanks!

dfried@cs.cmu.edu
https://dpfried.github.io

Cicero: https://ai.facebook.com/research/cicero/diplomacy/
Pragmatics and Communicative Success

- **Base Speaker**: throw out the purple chemical
- **Pragmatic Speaker**: throw out the first purple chemical
- **Human**: remove all the purple chemical from the beaker on the far left

The correct responses are marked with a checkmark.
Instruction Quality: Alchemy

Averaged from 3 or 5 point Likert scales [Daniele et al. 2017]. Differences between base and pragmatic all statistically significant by $\chi^2$ on counts.

### Amount of Information
- **Too Little**
  - Base speaker
  - Pragmatic speaker
  - Human instructions
- **Too Much**

### Difficulty of the Task
- **Very Hard**
  - Base speaker
  - Pragmatic speaker
  - Human instructions
- **Very Easy**

### Confidence in Reaching End State
- **Not Confident**
  - Base speaker
  - Pragmatic speaker
  - Human instructions
- **Confident**

---

Averaged from 3 or 5 point Likert scales [Daniele et al. 2017]. Differences between base and pragmatic all statistically significant by $\chi^2$ on counts.
Pragmatics and Communicative Success

Base Speaker
- remove the last figure
- add it back

Pragmatic Speaker
- remove the last figure
- add it back in the 3rd position

Human
- take away the last item
- undo the last step
Outperforming Training Data (Toy Example)

Training Data

Context → “Language”
- AX → x
- AX → x
- BX → x
- BX → x
- AX → ax
- BX → bx

True, but under-informative
True and informative

Base Speaker
- \( P_S(x \mid AX) = \frac{2}{3} \)
- \( P_S(ax \mid AX) = \frac{1}{3} \)

Base Listener
- \( P_L(AX \mid x) = \frac{1}{2} \)
- \( P_L(AX \mid ax) = 1 \)

Pragmatics as best response [Franke 2009; Jäger 2014]

Other formalisms:
- Recursive Bayesian agents [Frank and Goodman 2012; Jeon et al. 2020]
- Optimal transport of beliefs [Wang et al. 2020]
- Rate-distortion communication [Zaslavsky et al. 2020]
Speaker Also Helps In Training

Synthetic Instruction:
Walk past the dining room table and chairs and ...

Human Instruction:
Go down the stairs, go slight left at the bottom and go through door ...

Generate Instructions

Speaker

Synthetic routes through training environments

Human-annotated routes

Listener

Train
Speaker Also Helps In Training

![Bar chart showing navigation accuracy in unseen environments.]

- **Base listener**: 31%
- **+Augmentation**: 36%
- **+Pragmatics**: 50%
- **+Both**: 55%

The chart illustrates the improvement in navigation accuracy when using different techniques in unseen environments.
Why not Self-Play RL?

- **Natural Language Action Spaces are challenging**
  - Huge space, slow to generate, almost all strings are invalid, credit assignment...

- **Divergence from human language**

- **Language models aren’t causal**
  - For example, can be fooled by asserting that actions were agreed ("Thanks for agreeing to convoy your army to Bel this turn!")
  - Even nonsensical messages help convince an imitation-trained model that you are friendly with it
Cicero is Largely Able to Pass for Human

Eliiaa  Yesterday at 2:11 AM
I got the email and like
Holy shit what, I played with an AI? I don't ever remember playing with someone that didn't feel human like
How far is AI going holy shit
(1) Label Training Messages with Intents

1. Zero-shot classify turns where people were truthful

\[ P(\text{"You lied to me"}|s_t, a_t) \]

2. Train an intent model that predicts what actions people will take at the end of truthful turns.
(2) Use Intents as Control Tokens

Use intents as **control tokens** in dialogue model

Every message conditions on its intent
During play, Cicero conditions the dialogue model on **truthful** and **mutually beneficial** actions generated by *superhuman* planning engine.
Challenges for Language in Diplomacy

(1) Hold **human-like** dialogue
   - Hold long conversations of **hundreds of messages** with multiple parties
   - Messages must be **grounded** in both the environment and dialogue history

(2) **Win the game** through strategic negotiation
   - Interface with **external reasoning systems**
   - Use language **intentionally** to achieve goal

Tension between these goals!
Training Language Models to Negotiate

1. Adapt pre-trained LMs to generate dialogue in *grounded, out of domain* setting

2. Use language intentionally by *conditioning on plans*

3. Aggressively *filter messages* that may contain mistakes
Intent-Controlled Dialogue

Great plan, I’ll help you defend Belgium!
Intent-Controlled Dialogue

I’ll only support you if you help me against Germany.
Define intent of a message as the actions honest players would take if they agree to this proposal.

How can we infer these automatically?

(1) Label Training Messages with Intents
(1) Label Training Messages with Intents

(a) Intent Model Training

Intent Model
Trained to predict actual action players took after their last message in each turn
(1) Label Training Messages with Intents

(b) Intent Annotation

Intent Model
Trained to predict actual action players took after their last message in each turn

Inference over all messages
(1) Label Training Messages with Intents

(b) Intent Annotation

<table>
<thead>
<tr>
<th>Method</th>
<th>% of predictions containing labeled orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model</td>
<td>77</td>
</tr>
<tr>
<td>+ Initialized from dialogue model</td>
<td>87</td>
</tr>
<tr>
<td>+ Injected agreement messages</td>
<td>93</td>
</tr>
<tr>
<td>+ Restriction to truthful subset</td>
<td>97</td>
</tr>
</tbody>
</table>
Intent-Controlled Dialogue

I'll only support you if you help me against Germany.