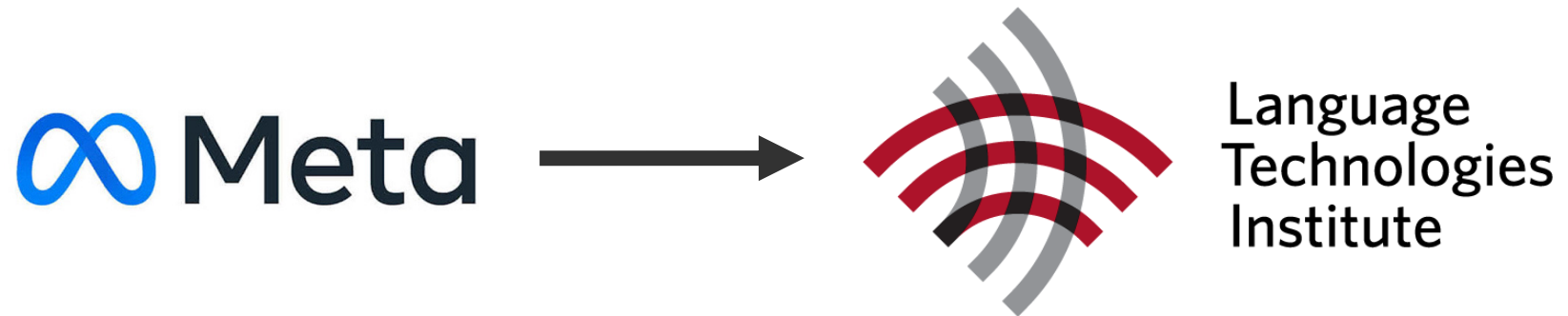


Using Language Strategically in Context

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”

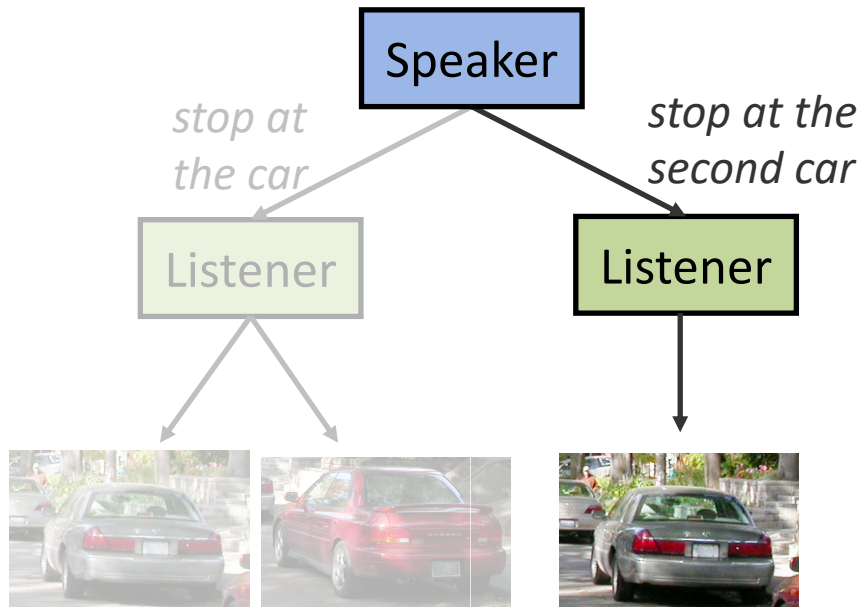


Using Language Strategically

Generation



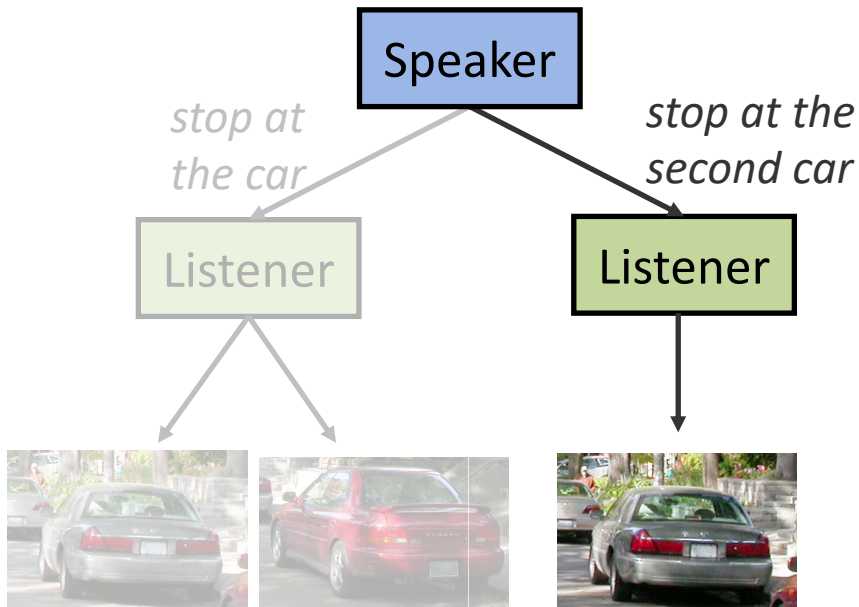
Interpretation



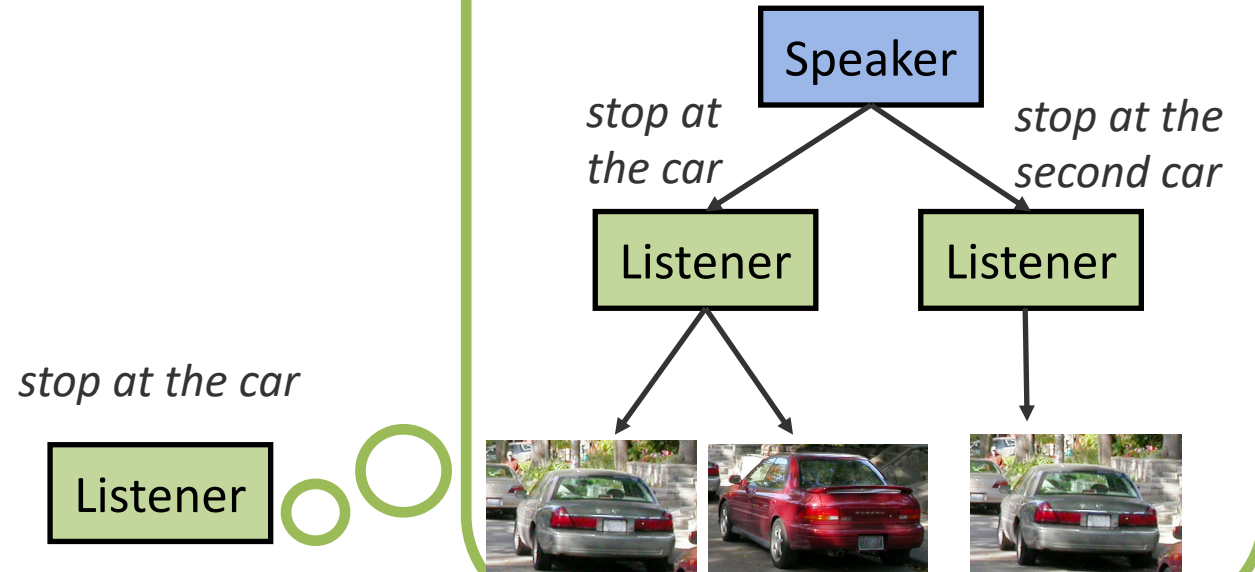
[e.g. Lewis 1969; Golland et al. 2010;
Frank and Goodman 2012; Degen et al. 2013]

Using Language Strategically

Generation



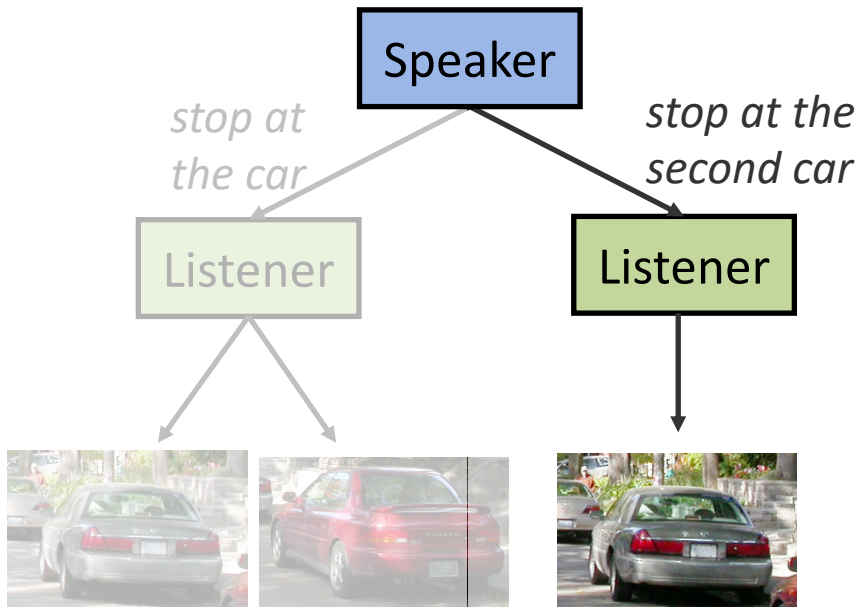
Interpretation



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

Using Language Strategically

Generation

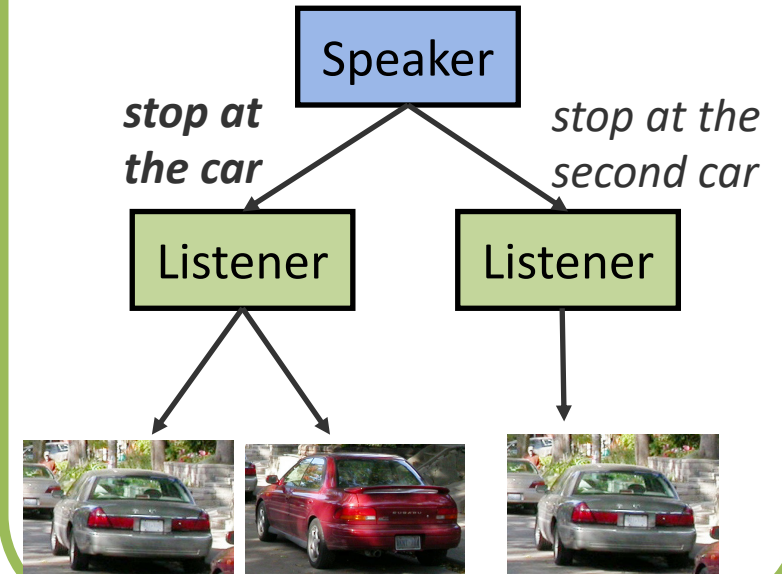


Interpretation



stop at the car

Listener

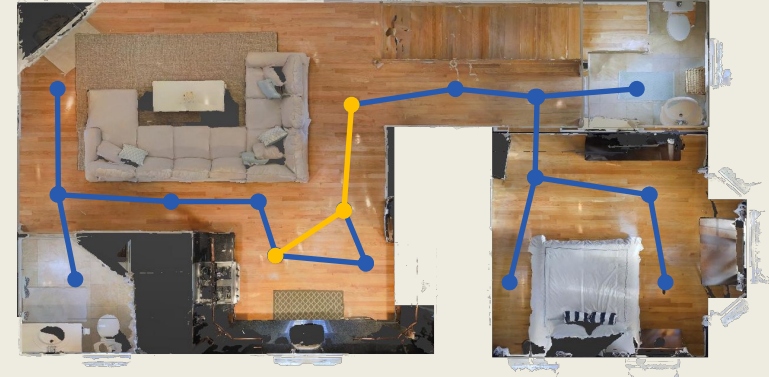


[e.g. Lewis 1969; Golland et al. 2010; Frank and Goodman 2012; Degen et al. 2013]

Reasoning with Speakers and Listeners



Grounded Instructions



[Fried*, Hu*, Cirik* et al. NeurIPS 2018]

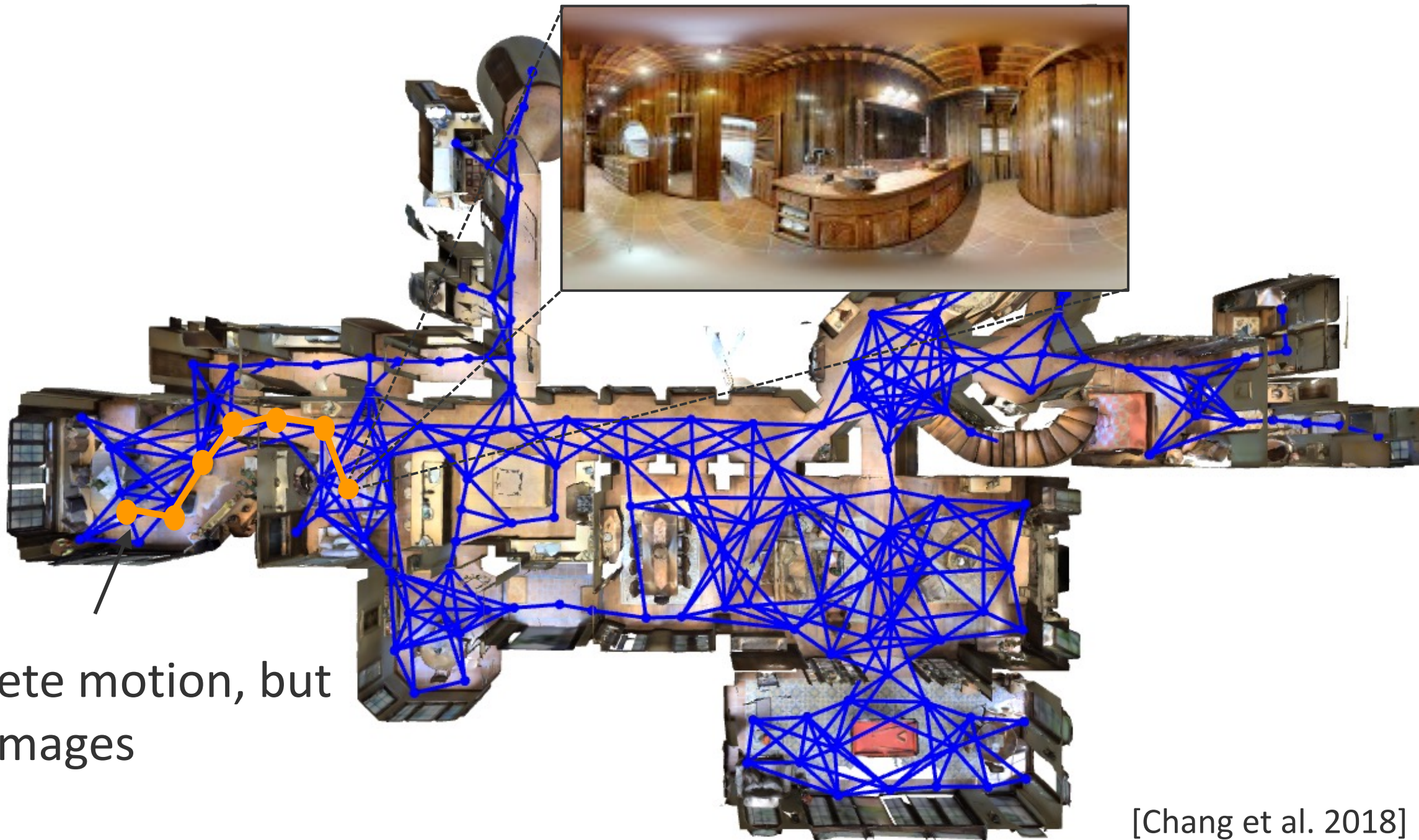
A map of Europe with various countries labeled with three-letter codes. The map is color-coded: pink for the British Isles, purple for Scandinavia and parts of Central Europe, brown for Central Europe, red for Eastern Europe, green for Southern Europe, and yellow for the Balkans and Turkey. Symbols include stars, circles, and crosses.

Acting in a Visual Environment



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.

[Vision-and-Language Navigation Task. Anderson et al., 2018]



Discrete motion, but
real images

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:



Generation Task

Input a route:



Output instructions:

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

$$P_L(r | i)$$



Outputs



Route, r



Route, r



Speaker

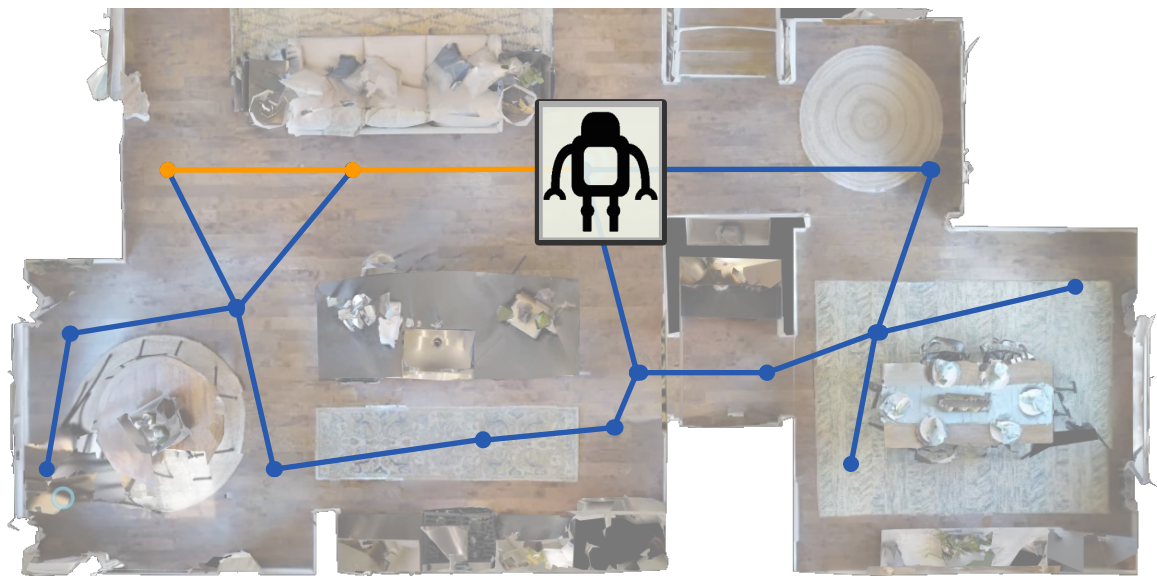
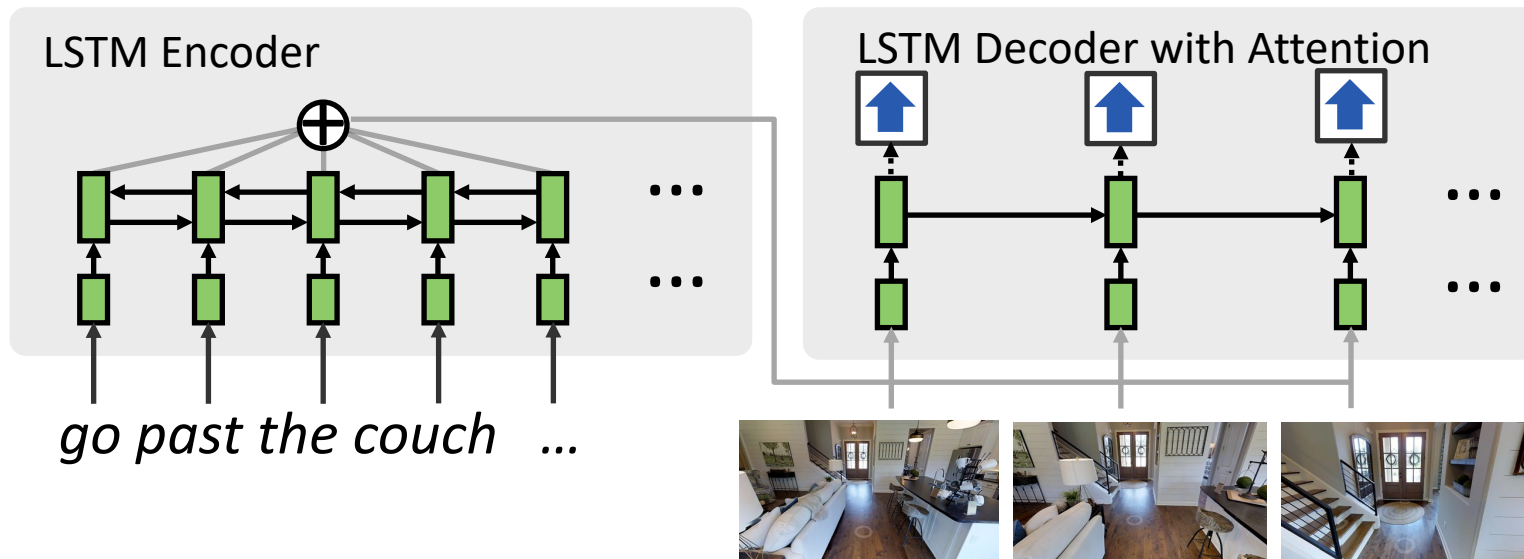
$$P_S(i | r)$$



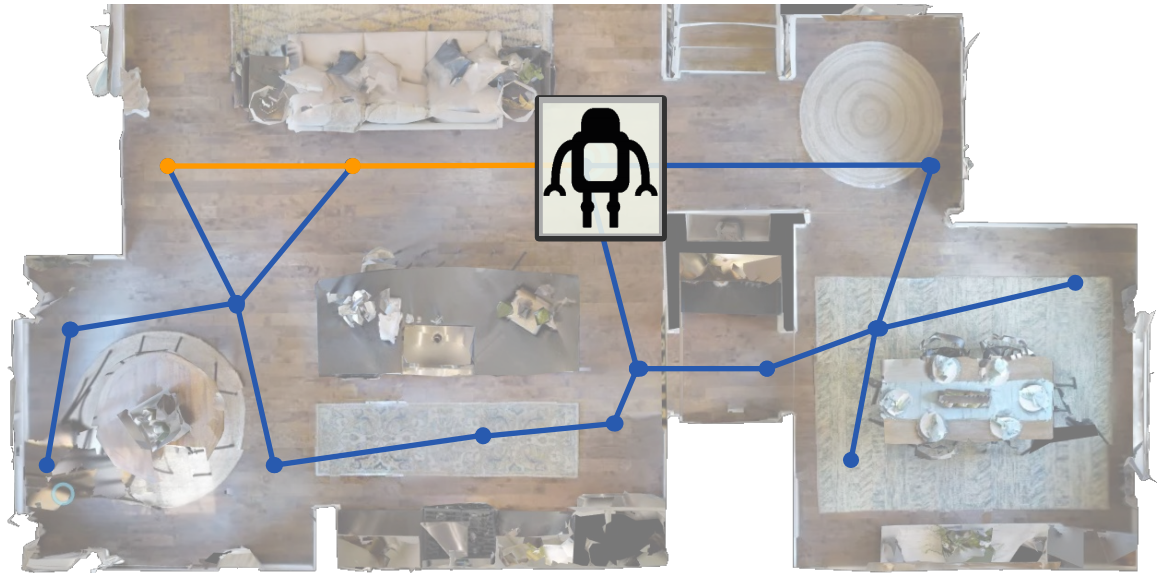
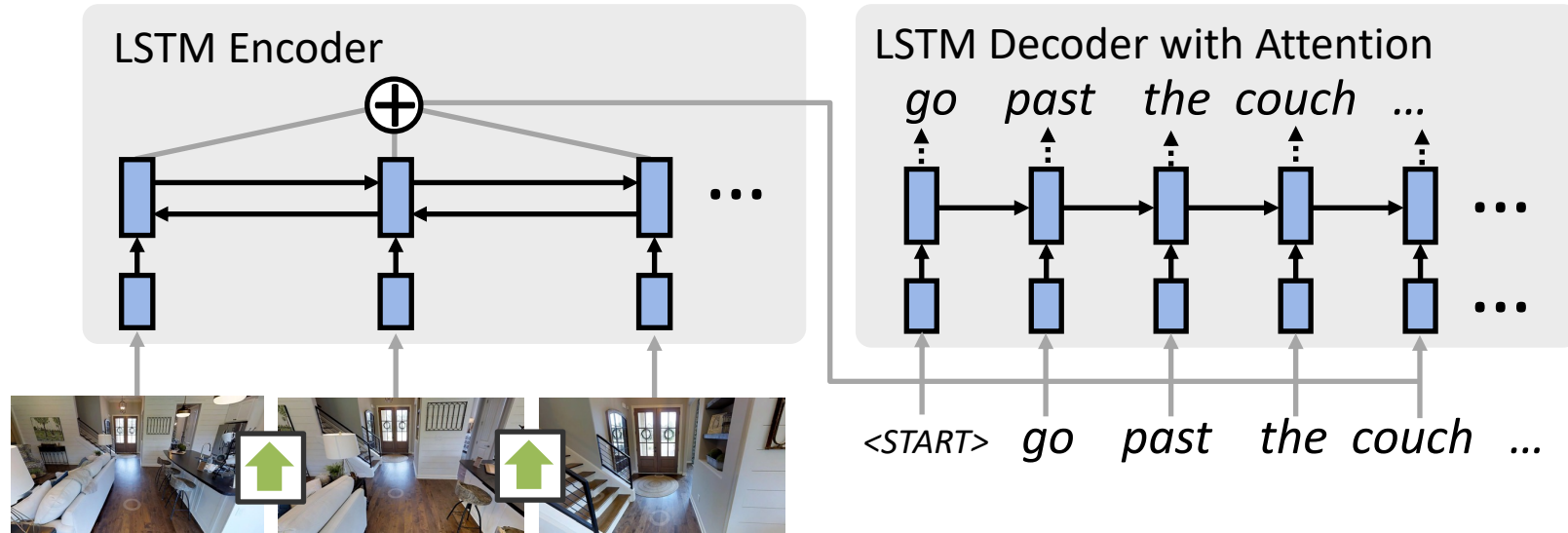
*Go forward between the
kitchen counters...*

Instruction, i

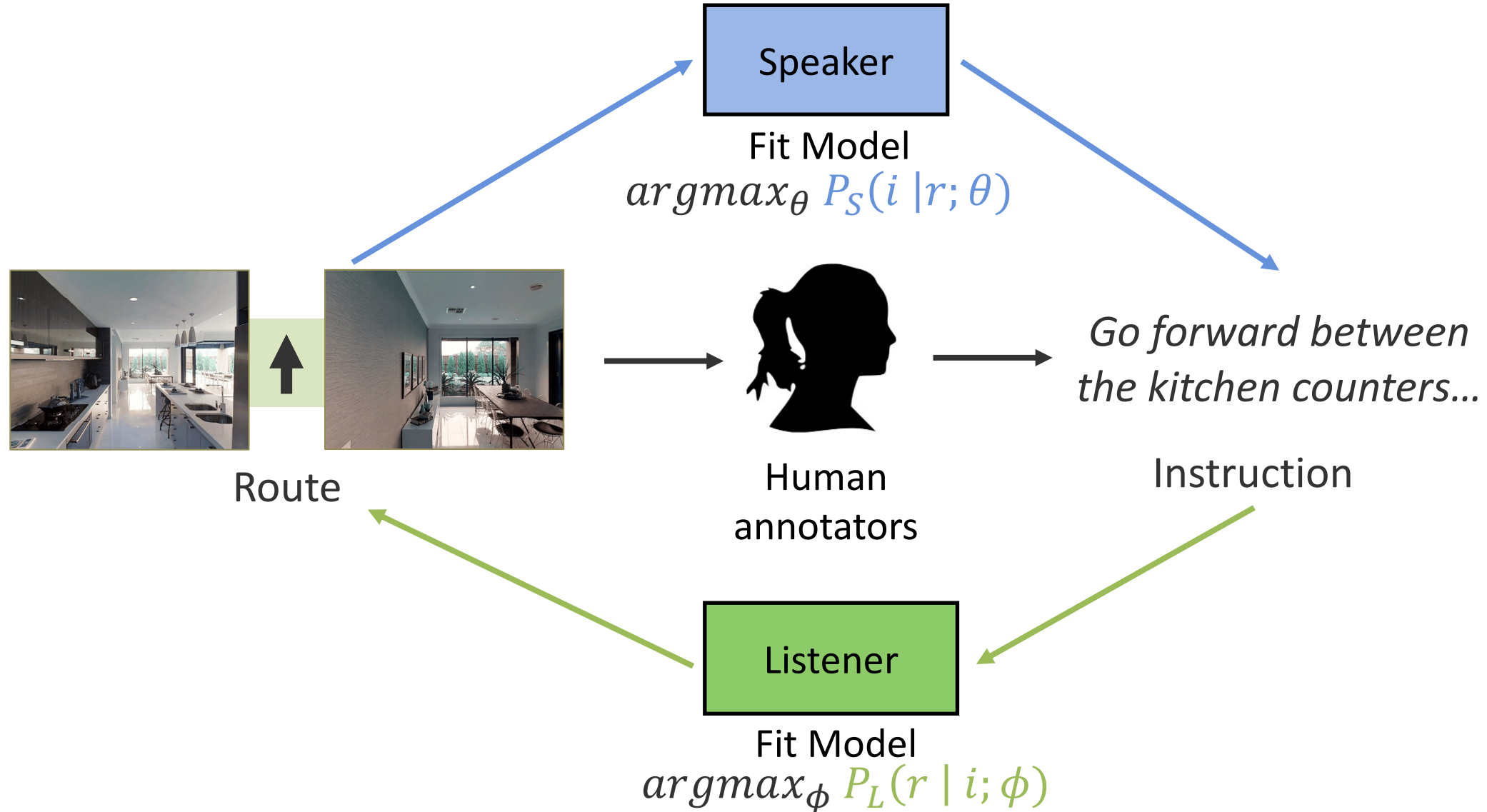
Base Listener Model



Base Speaker Model



Training Models on Human Instructions



Pragmatic Generation



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

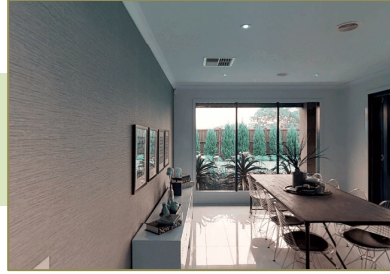
Speaker

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

Listener

Listener

Pragmatic Generation



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. stop once you are on the rug.

Listener

Listener

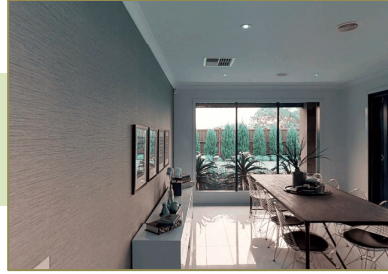
...



...



Pragmatic Generation



*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.
stop once you are on the rug.*

Listener

0.4

$P_L(r \mid i, \dots)$

Listener

0.8

...

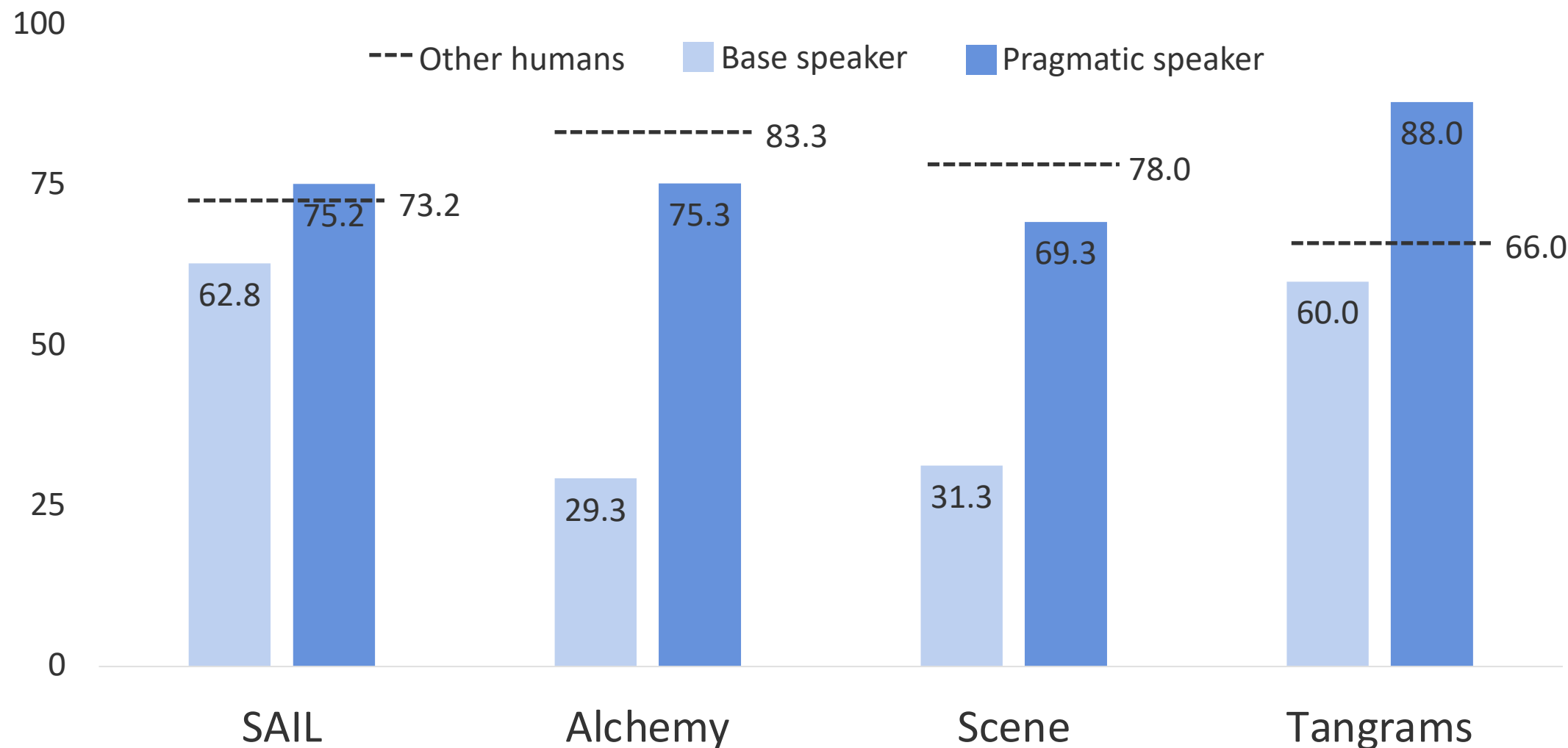


...

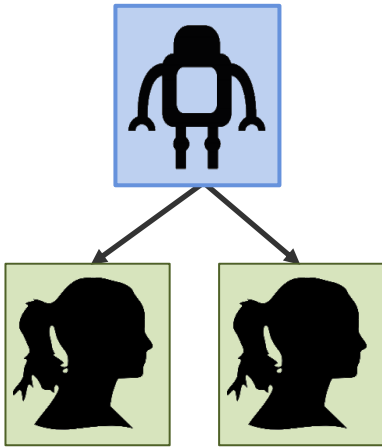


Generation Results

Human accuracy at following instructions from:

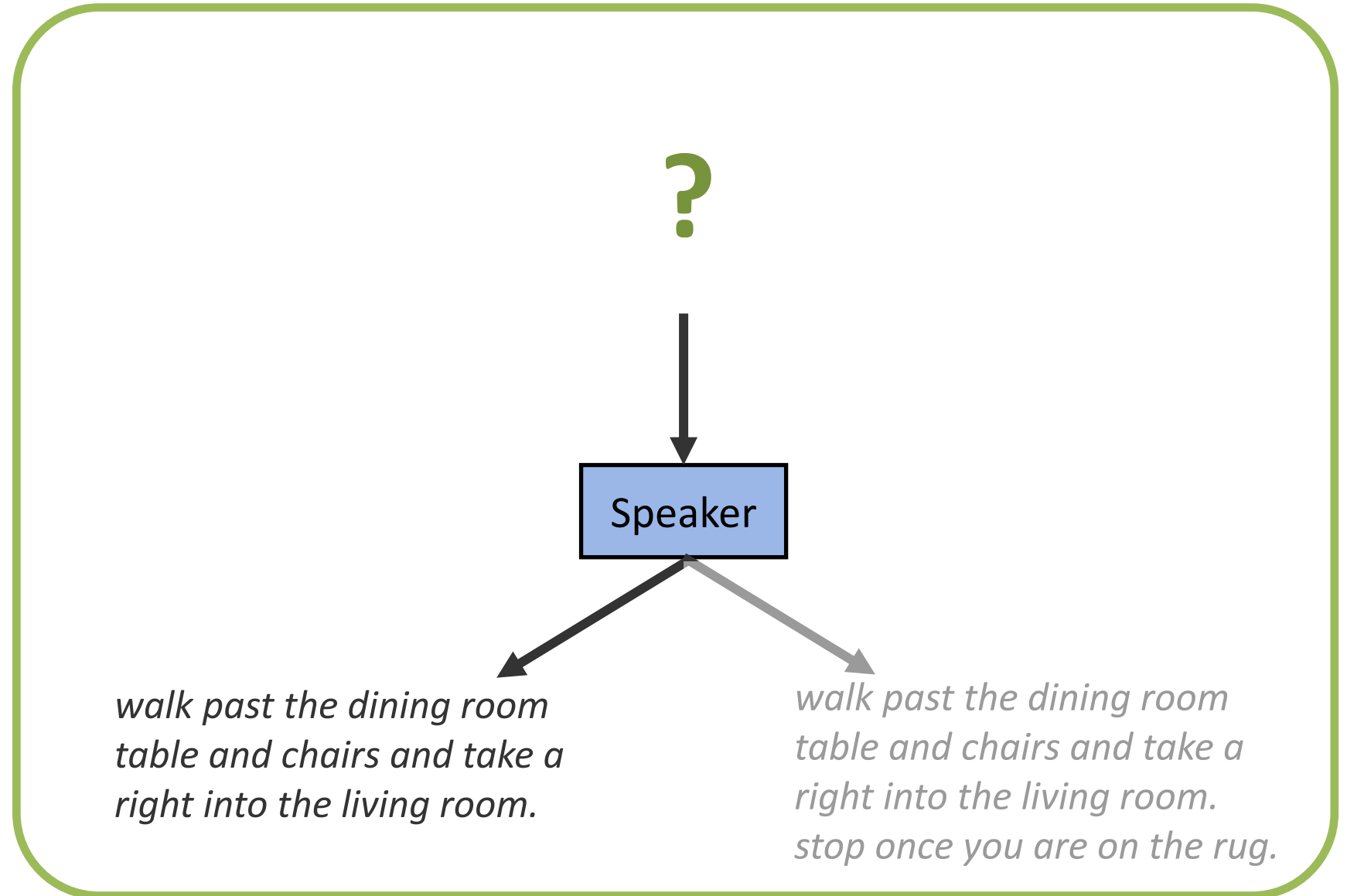
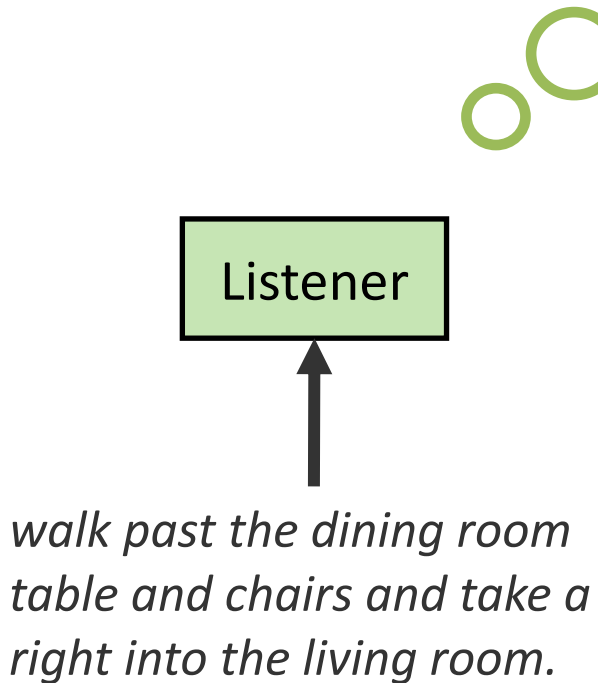


Takeaway

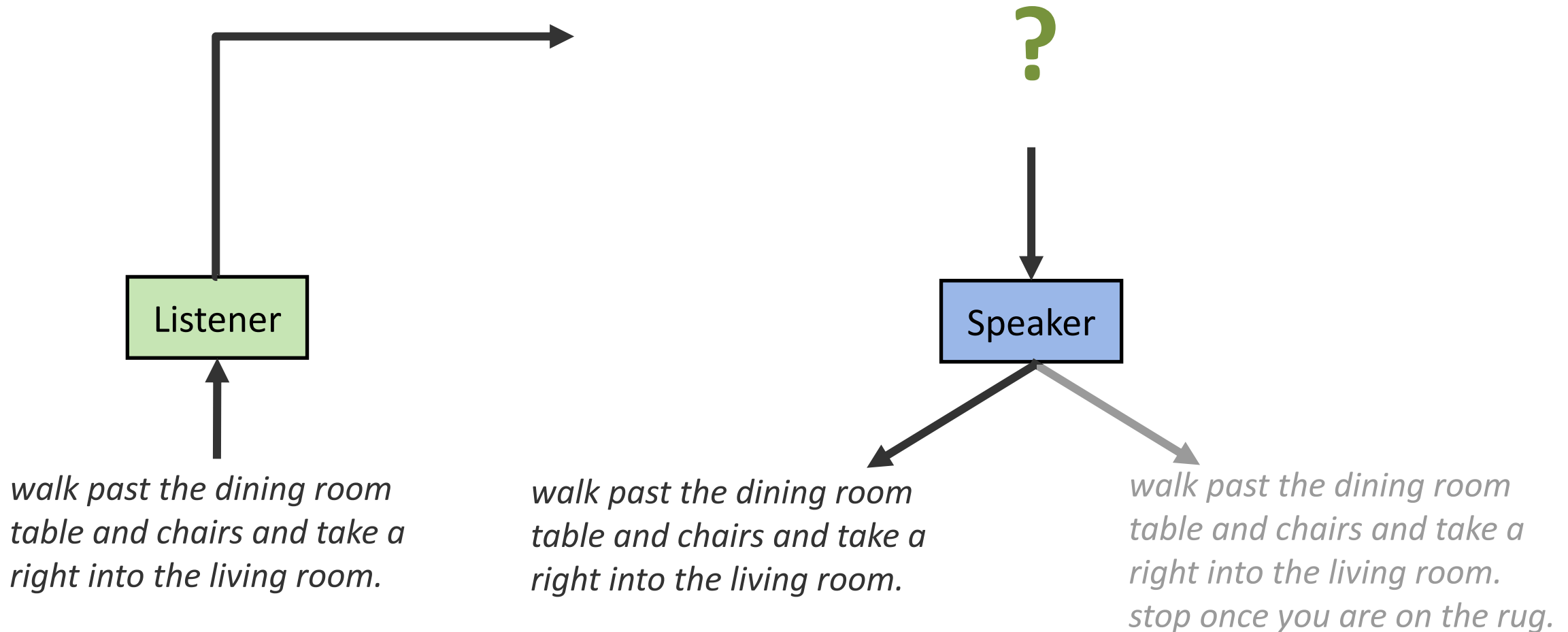


*Simulating people's interpretations
makes language more informative.*

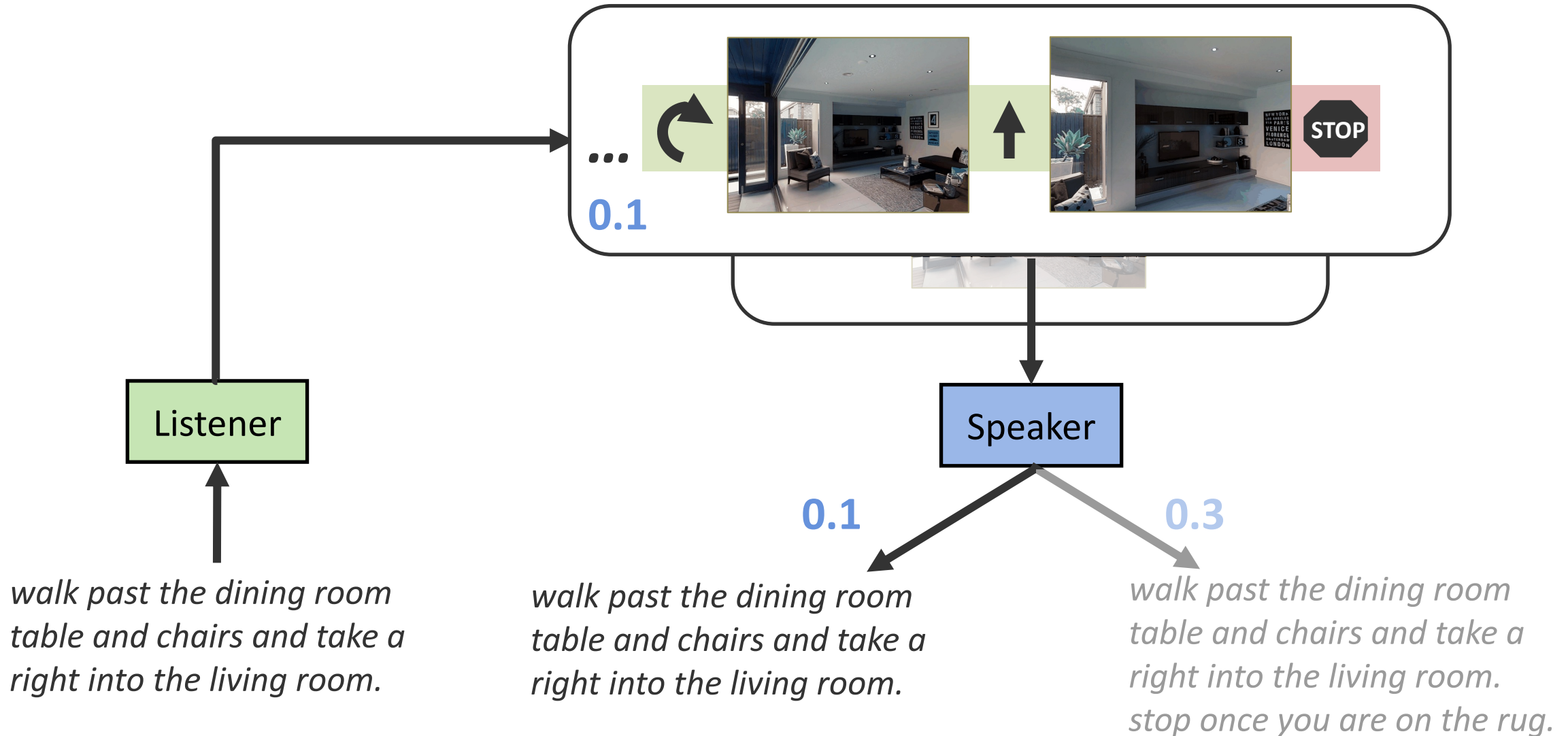
Pragmatic Interpretation



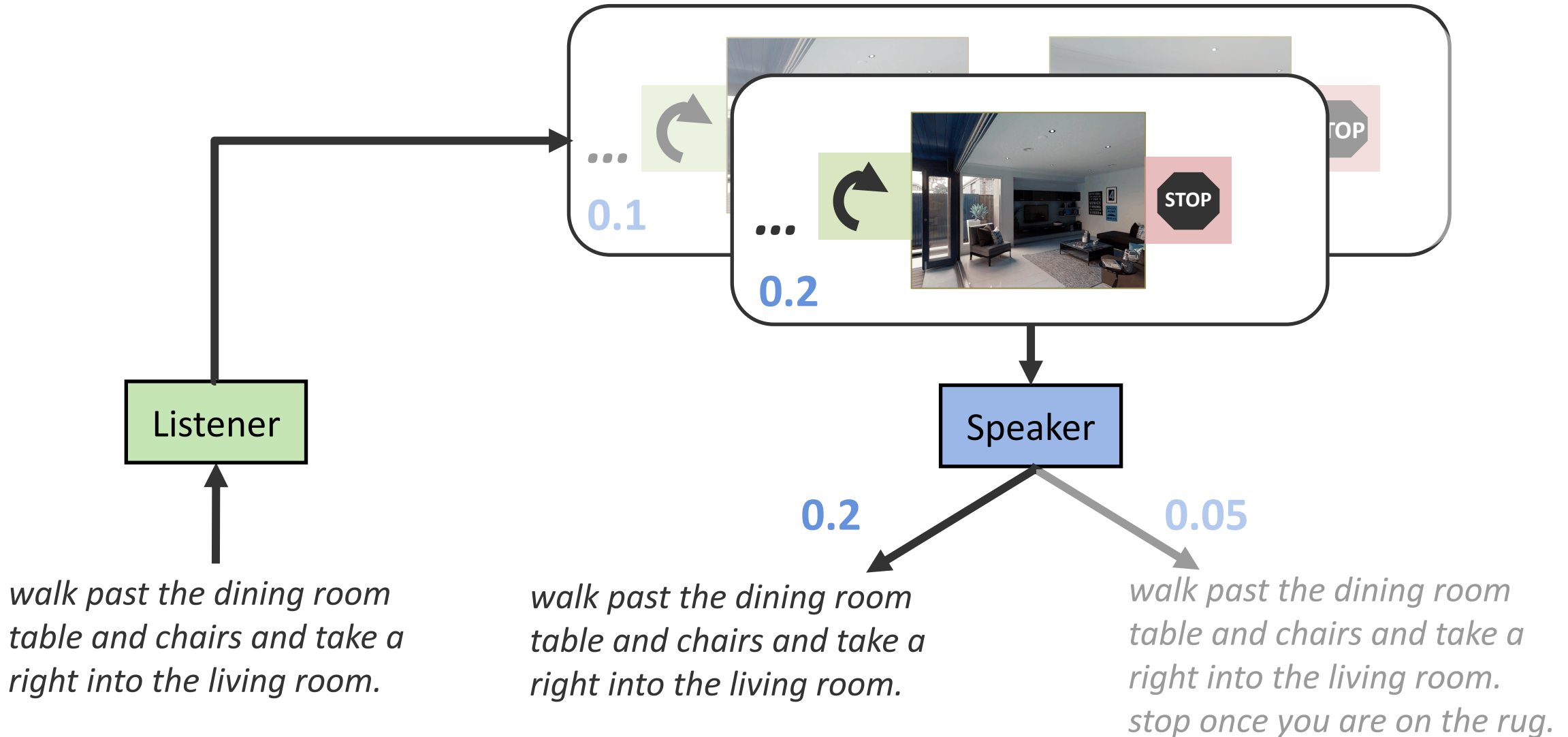
Pragmatic Interpretation



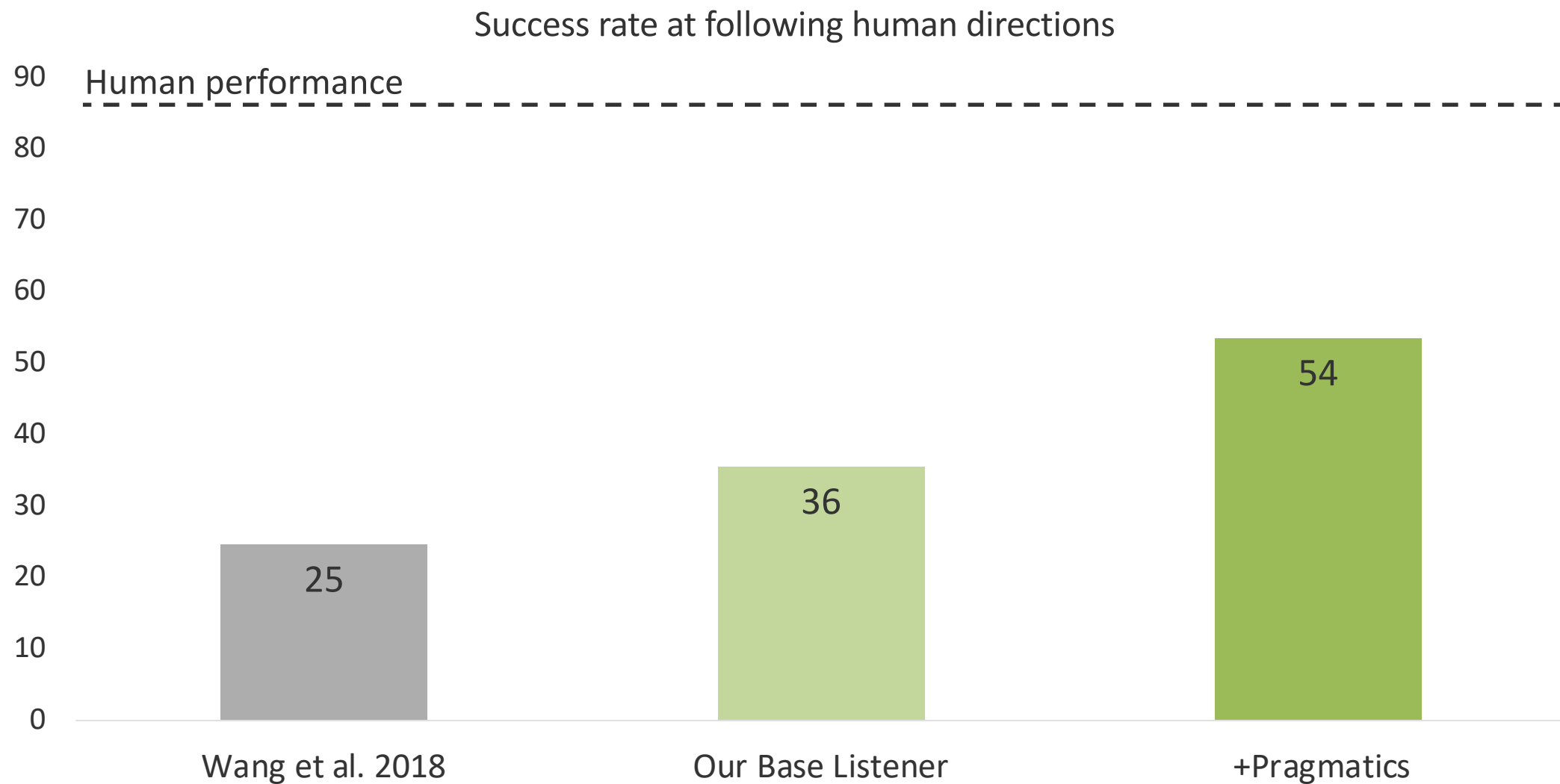
Pragmatic Interpretation



Pragmatic Interpretation



Pragmatic Interpretation



[Fried*, Hu*, Cirik* et al., NeurIPS 2018]

*Walk past hall table. Walk into bedroom. Make left at table clock.
Wait at bathroom door threshold.*



Base listener

*Walk past hall table. Walk into bedroom. Make left at table clock.
Wait at bathroom door threshold.*

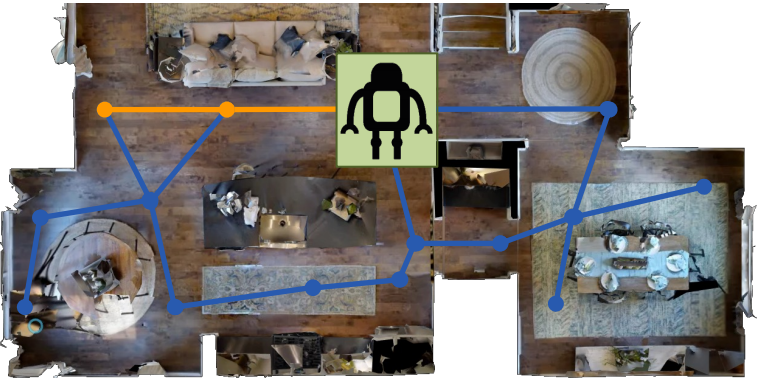


Pragmatic listener

Takeaway

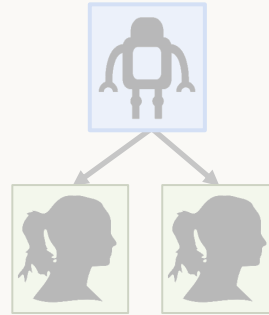


Simulating why a speaker said what they did helps resolve ambiguity.



Strategic Language for...

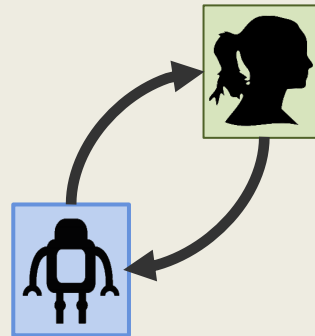
Grounded Instructions



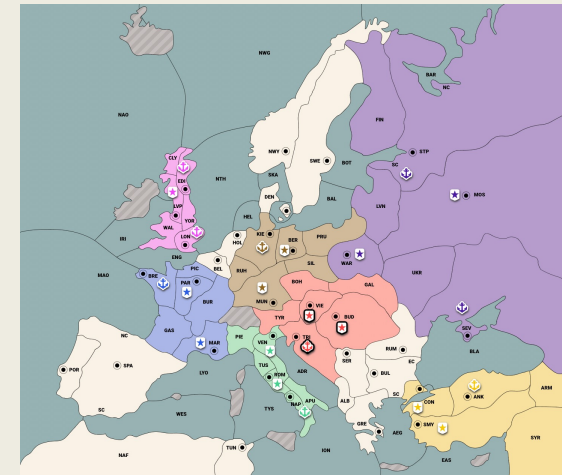
[Fried, Andreas, & Klein. NAACL 2018]

[Fried*, Hu*, Cirik* et al. NeurIPS 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

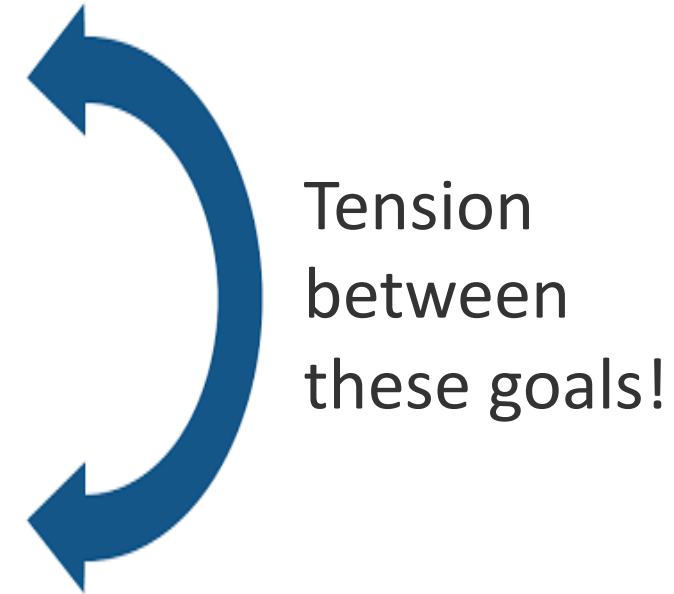
Challenges for Language in Diplomacy

(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



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

ENGLAND: Lonellllyyyyy

ENGLAND: I am so lonellllyyyyy

ENGLAND: Would really love to hear from you!

ENGLAND: I really think it's in your best interests to move to MAO

FRANCE: i am on the phone with my gf.

ENGLAND: Ah, I see.

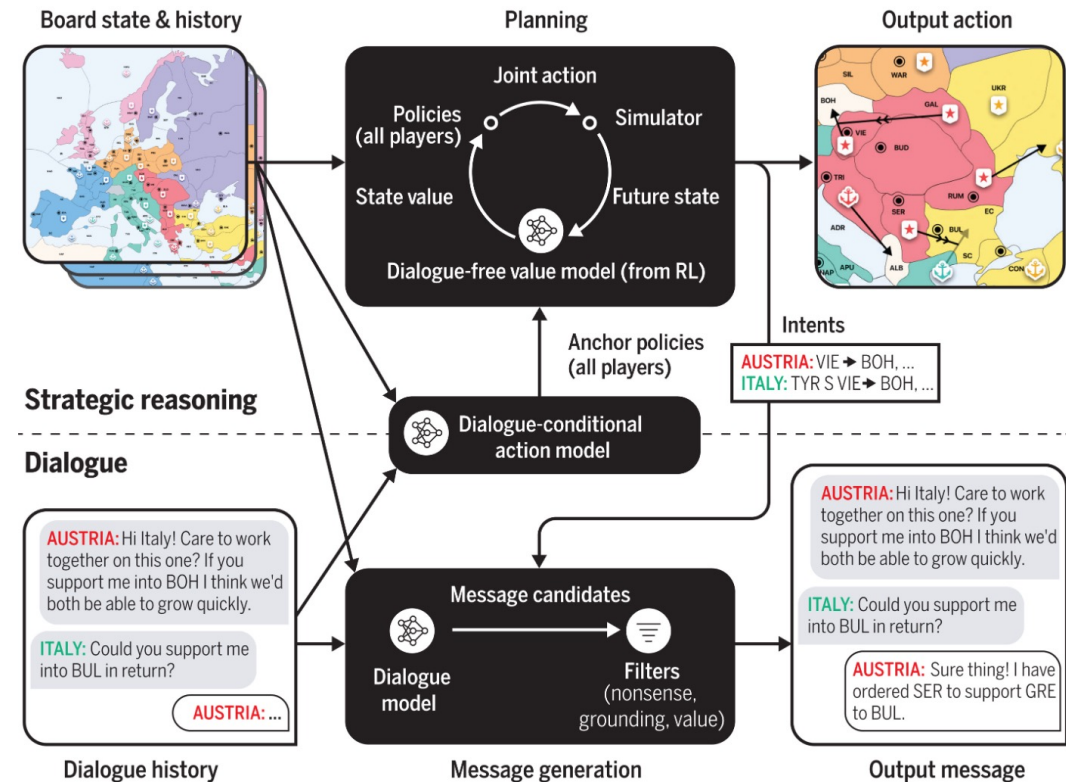
ENGLAND: can you go to MAO anyway? I'm sure your gf won't mind

Limits of Imitation Learning

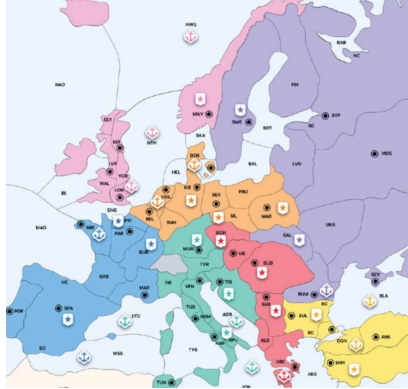
- ▶ Sampling from even a perfect imitation learning model only reproduces **average play**
 - ▷ There are many **weak players** in our dataset
 - ▷ If our messages describe bad actions, no-one will cooperate with us
- ▶ **We can't beat the best humans by copying them**

Beyond Imitation Learning

- ▶ Decouple action planning and language generation
- ▶ Use a language-informed planner
 - ▷ Another talk!
- ▶ Use planner-informed language
 - ▷ Condition agent's language on agent's plans
 - ▷ Reason about the effect of agent's language on other players' plans



Intended Action-Conditioned Dialogue



FRA: ...
ENG: NTH S BEL

ENGLAND: Do you want
NTH to support BEL?

FRA: BEL -> HOL
ENG: ...

FRANCE: No, BEL is
moving to HOL.

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

ENGLAND: Alright i'll
support you in.

Model-Based Annotation:

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

Intent
Predictor

ENGLAND: Alright i'll
support you in.

Dialogue Conditioning:

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

Dialogue
Model

ENGLAND: Alright i'll
support you in.

During Play:

Planner

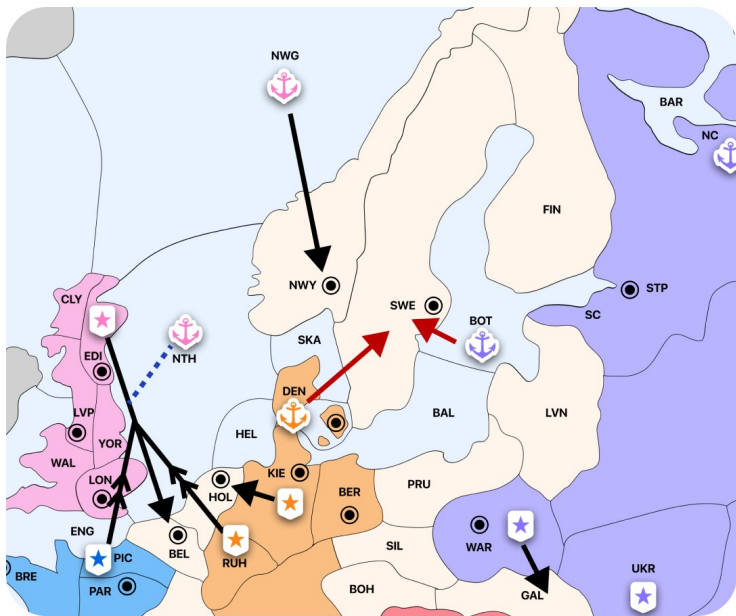
FRA: BRE -> MAO
ENG: LON -> NTH

Dialogue
Model

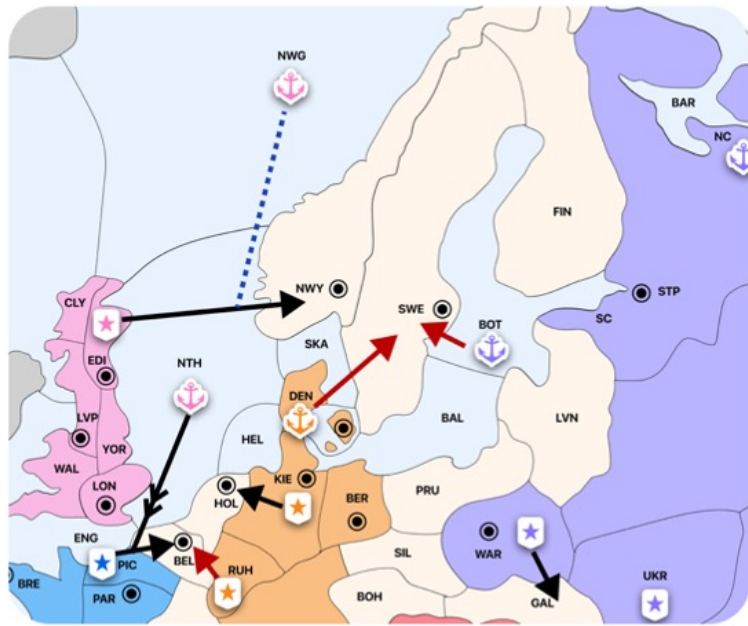
ENGLAND: Okay, I'll move
to North Sea then.

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

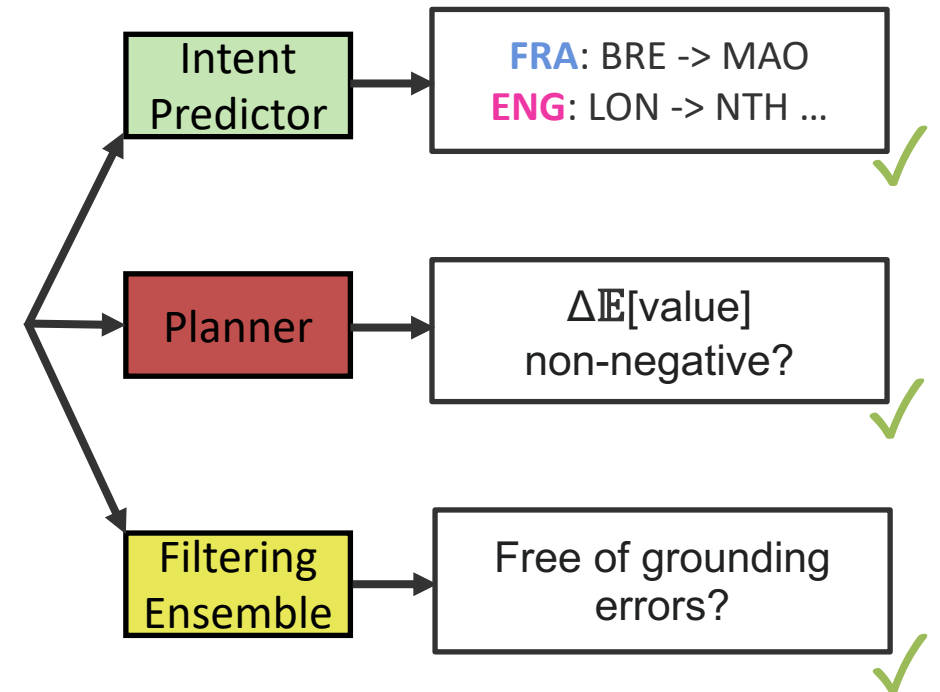
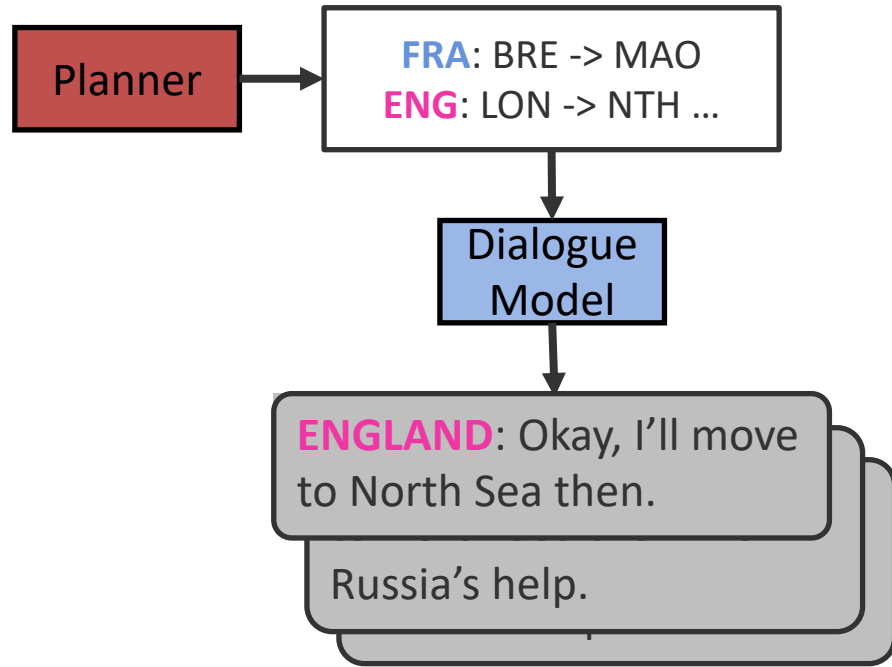
RUSSIA -> **TURKEY** : You lied

RUSSIA -> **TURKEY** : And you know it

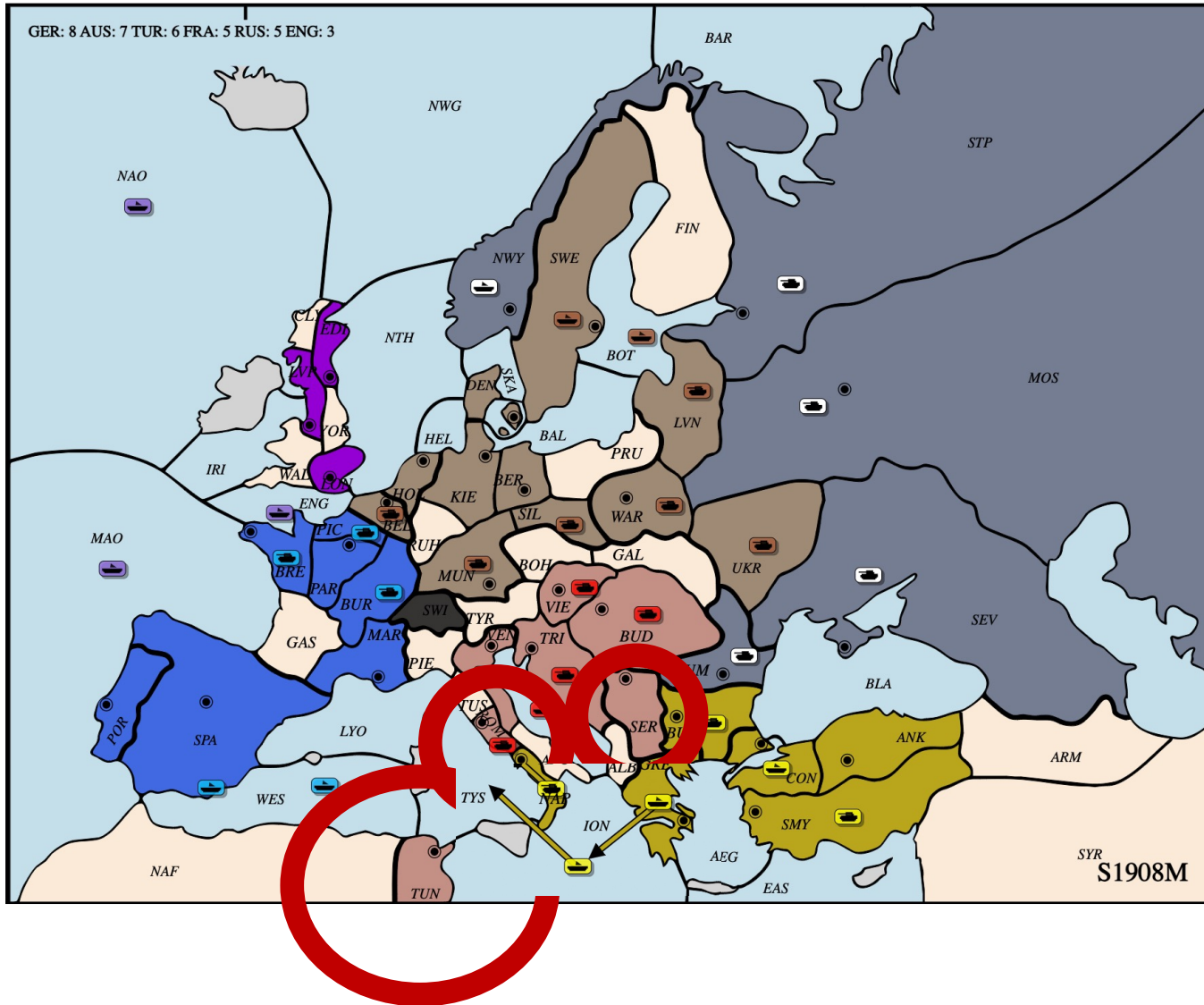
RUSSIA -> **TURKEY** : Don't play me for a fool

RUSSIA -> **TURKEY** : You lose until you repent

Encouraging Strategic Honesty



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.

FRANCE: You have Serbia and Rome to take.

TURKEY: they're impossible targets

FRANCE: Greece - Ionian Ionian - Tyrr

TURKEY: hm, you're right

TURKEY: good ideas

FRANCE: Then in fall you take Rome and Austria collapses.

Cicero Plays with Humans

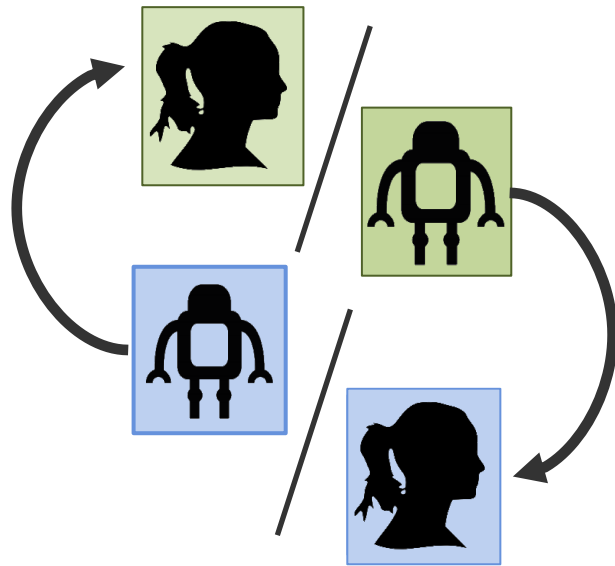
- ▶ This summer, we entered Cicero anonymously in an online Diplomacy league
- ▶ Cicero **placed in the top 10% of players**, and 2nd of 19 players who played ≥ 5 games. It achieved **more than double** the average human score.

Rank	Avg Score	# Games
1	35.0%	11
2	25.8%	40
3	24.5%	6
4	22.7%	8
5	21.0%	5
...		
19	3.0%	6
20	2.6%	7

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.

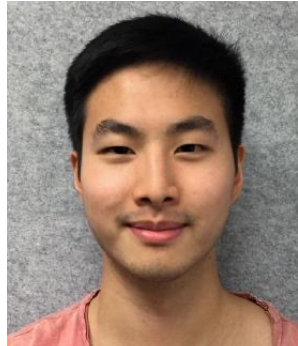
Grounding Collaborators



Jacob Andreas



Taylor Berg-
Kirkpatrick



Justin Chiu



Volkan Cirik



Trevor Darrell



Ronghang Hu



Dan Klein



Louis-Philippe
Morency



Anna Rohrbach

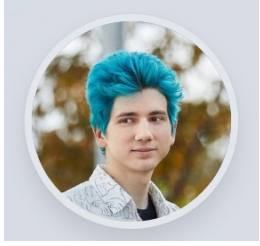


Kate Saenko

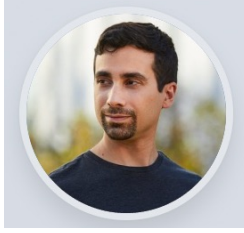


Sheng Shen

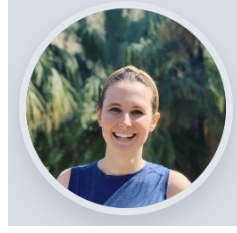
FAIR Diplomacy Team



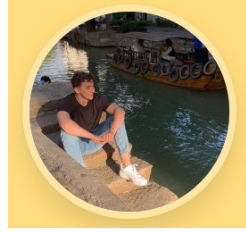
Anton Bakhtin



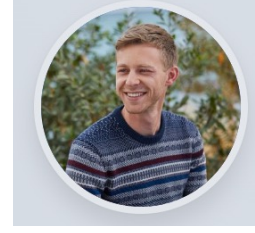
Noam Brown



Emily Dinan



Colin Flaherty



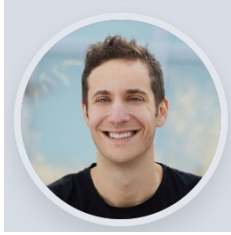
Jonathan Gray



Hengyuan Hu



Athul Paul Jacob



Adam Lerer



Mike Lewis



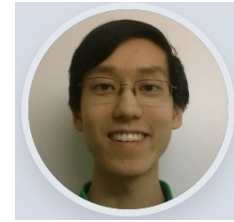
Alexander Miller



Adithya
Renduchintala



Weiyan Shi



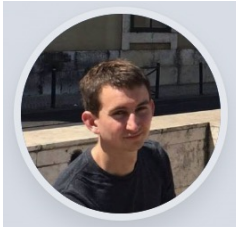
David Wu



Hugh Zhang



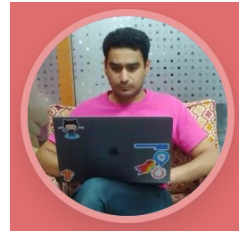
Gabriele Farina



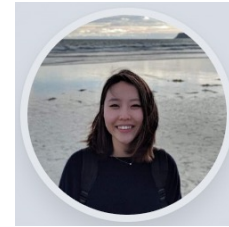
Daniel Fried



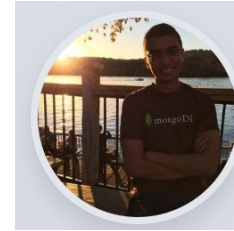
Andrew Goff



Mojtaba Komeili



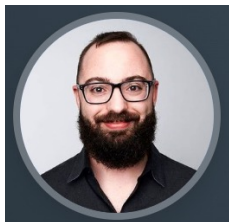
Minae Kwon



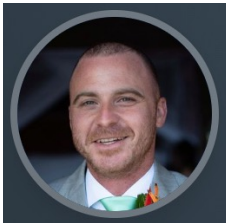
Karthik Konath



Sasha Mitts



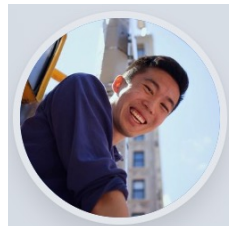
Stephen Roller



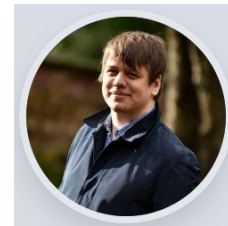
Dirk Rowe



Joe Spisak



Alex Wei



Markus Zijlstra

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*

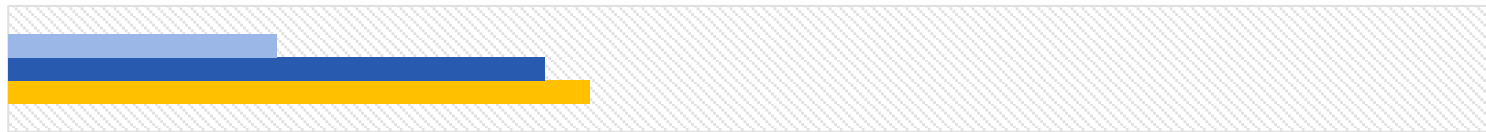
✓

Instruction Quality: Alchemy

Base speaker Pragmatic speaker Human instructions

Amount of Information

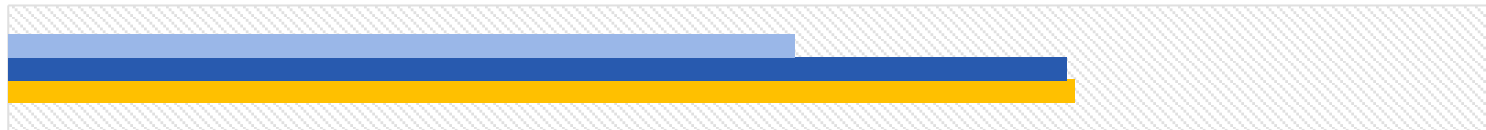
Too Little



Too Much

Difficulty of the Task

Very Hard



Very Easy

Confidence in Reaching End State

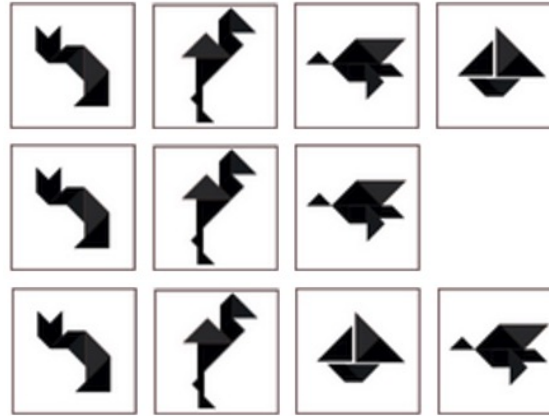
Not Confident



Confident

Averaged from 3 or 5 point Likert scales [Daniele et al. 2017]. Differences between base and pragmatic all statistically significant by χ^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 \rightarrow “Language”

AX \rightarrow x

AX \rightarrow x

BX \rightarrow x

BX \rightarrow x

True, but
under-informative

AX \rightarrow ax

BX \rightarrow bx

True and informative

Base Speaker

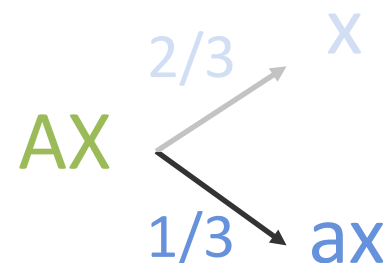
$$P_S(x | AX) = 2/3$$

$$P_S(ax | AX) = 1/3$$

Base Listener

$$P_L(AX | x) = 1/2$$

$$P_L(AX | ax) = 1$$



Pragmatics as best response [Franke 2009; Jäger 2014]

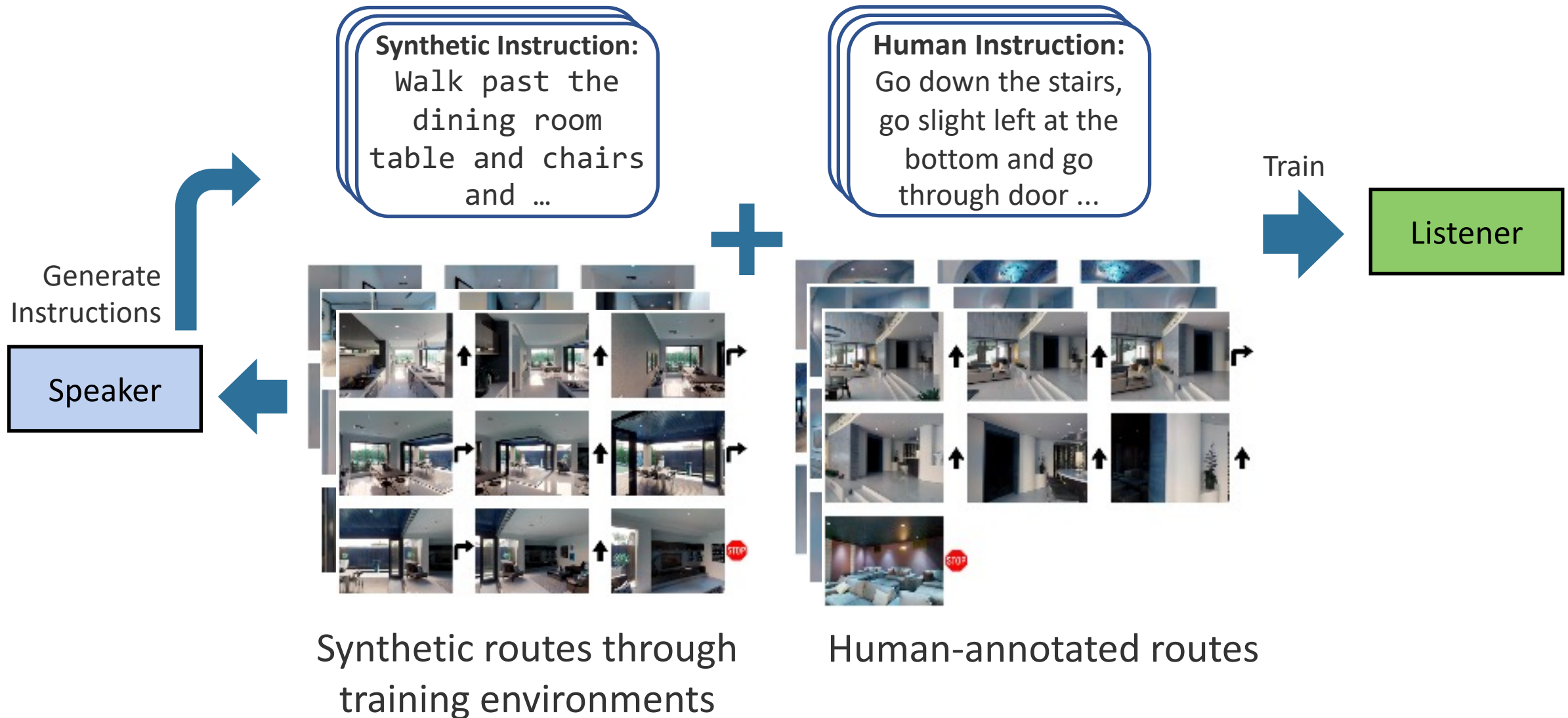
Other formalisms:

Recursive Bayesian agents [Frank and Goodman 2012; Jeon et al. 2020]

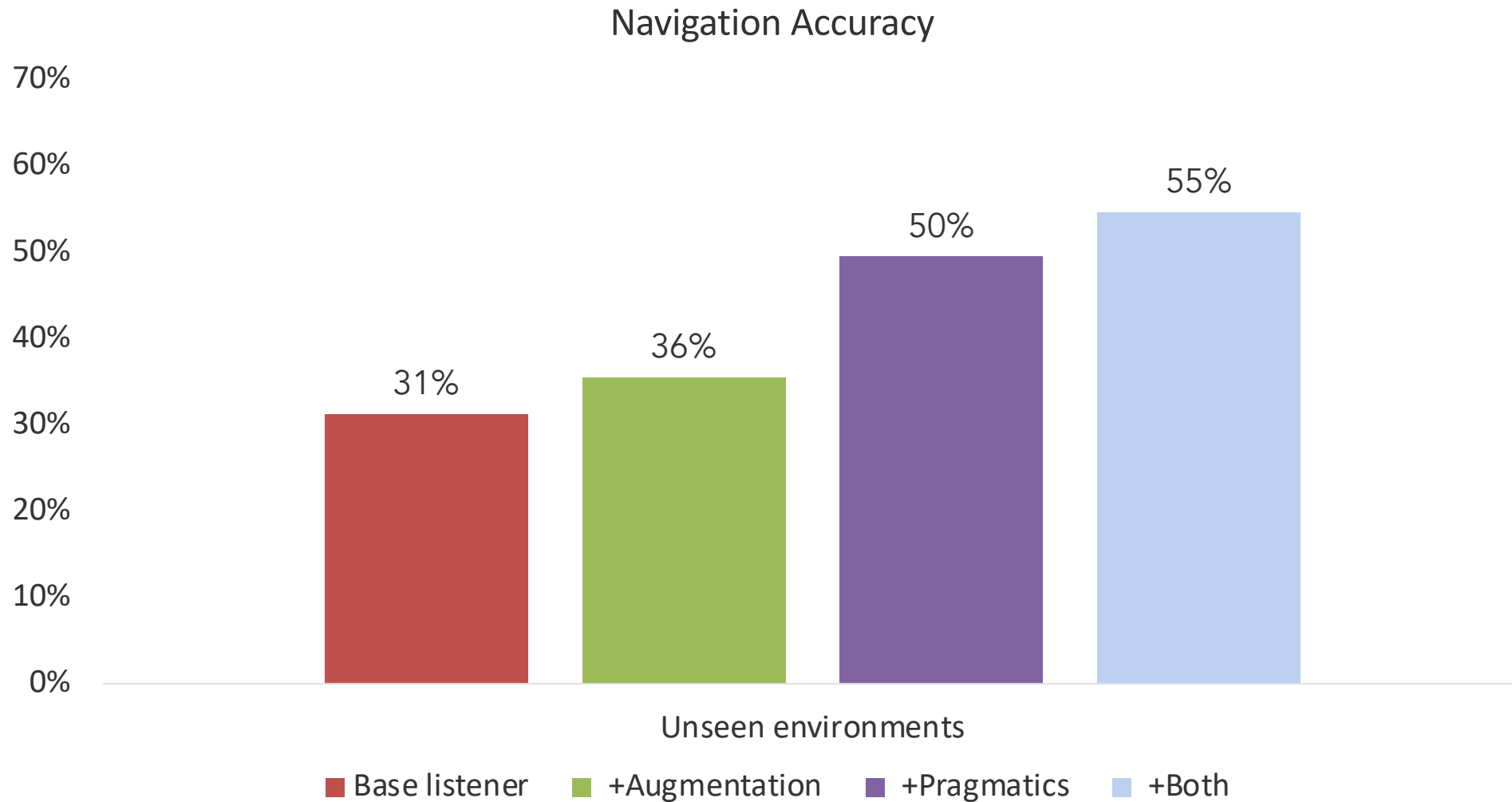
Optimal transport of beliefs [Wang et al. 2020]

Rate-distortion communication [Zaslavsky et al. 2020]

Speaker Also Helps In Training

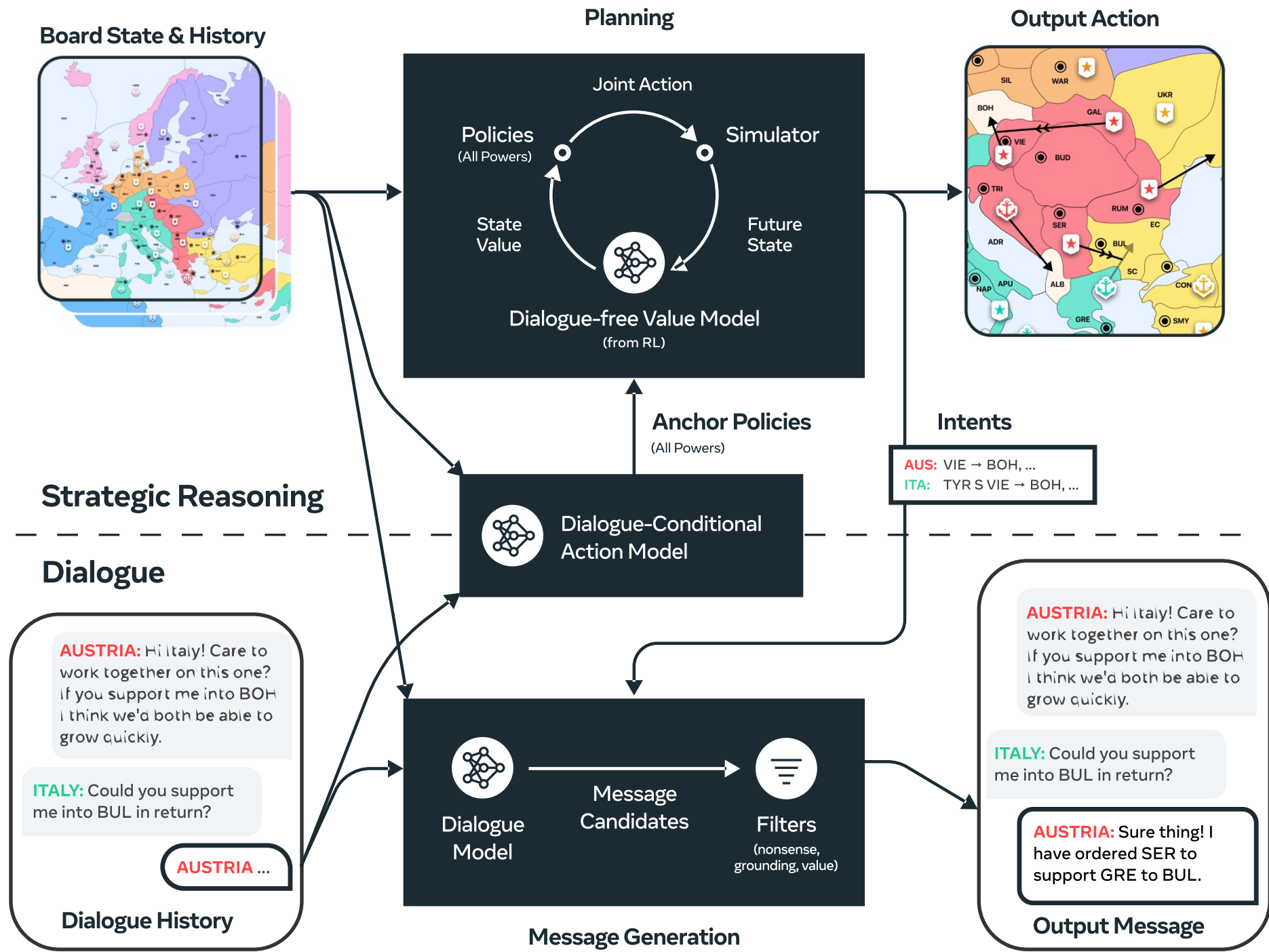


Speaker Also Helps In Training




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

 **Eliia** Yesterday at 2:11 AM

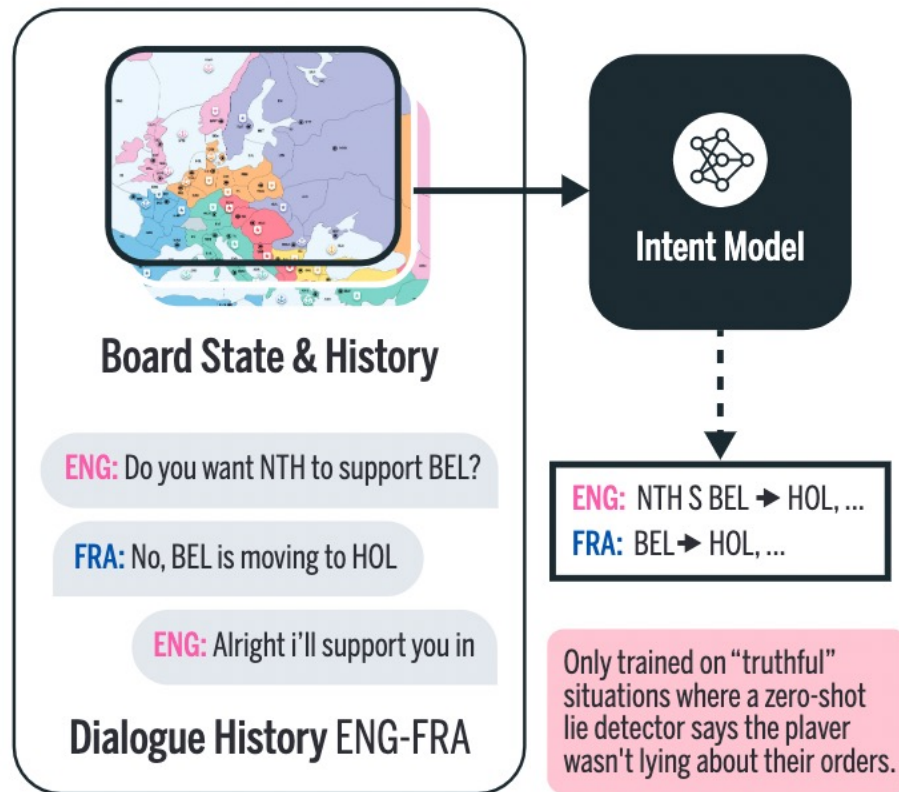
I got the email and like

Holy s**t what, I played with an AI? I don't ever remember playing with someone that didn't feel human like

How f**king far is AI going holy s**t

(1) Label Training Messages with Intents

(a) Intent Model Training



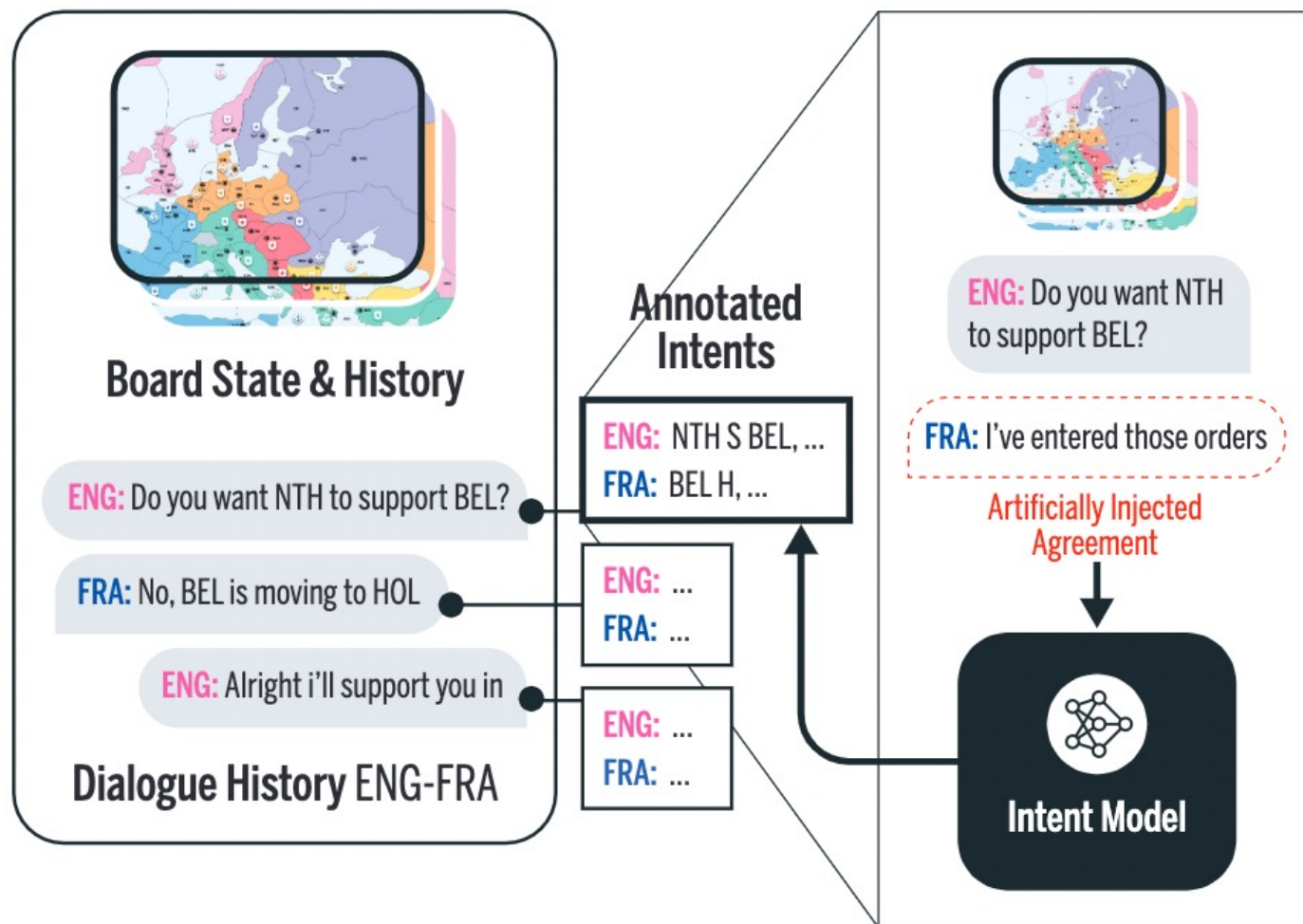
1. Zero-shot classify turns where people were truthful

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

2. Train an **intent model** that predicts what actions people will take at the end of **truthful** turns.

(2) Use Intents as Control Tokens

(b) Intent Annotation

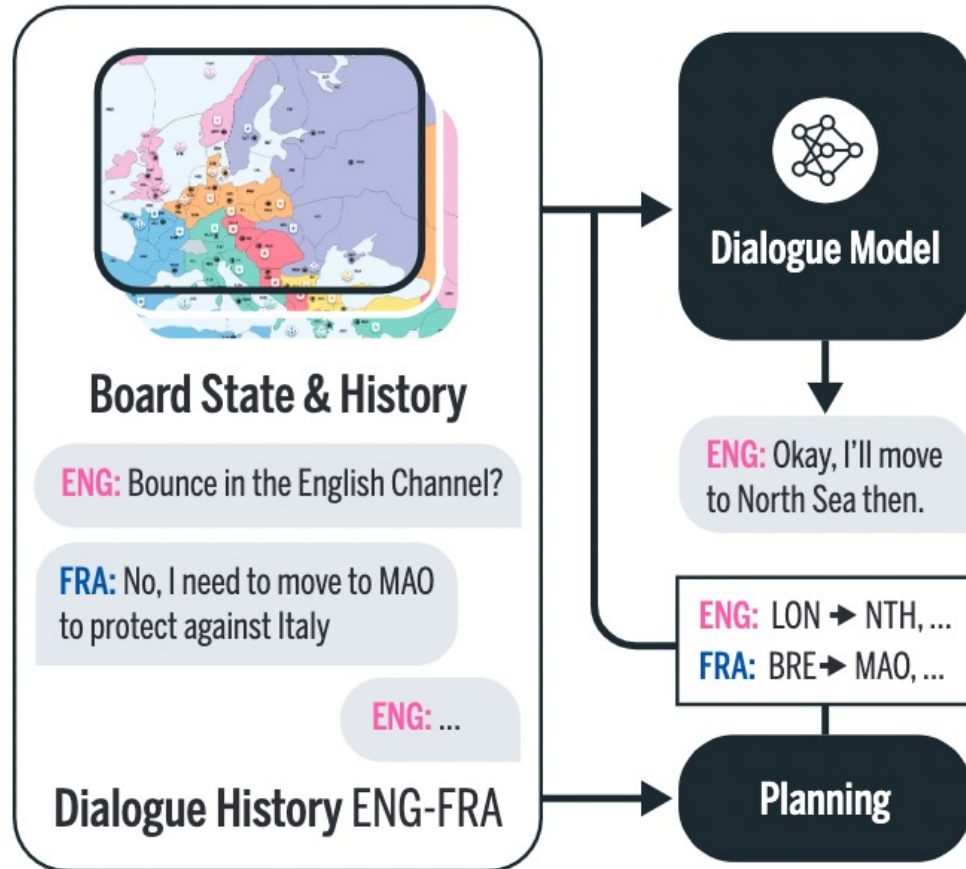


Use intents as **control tokens** in dialogue model

Every message conditions on its intent

(3) Control Dialogue with Planner

(d) Dialogue Model Inference



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

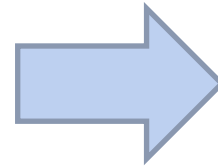
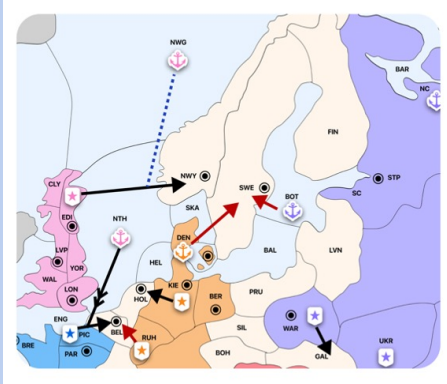


Tension
between
these goals!

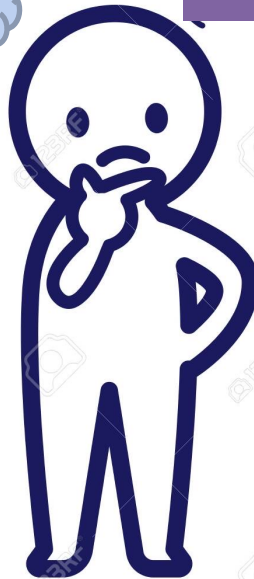
Training Language Models to Negotiate

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

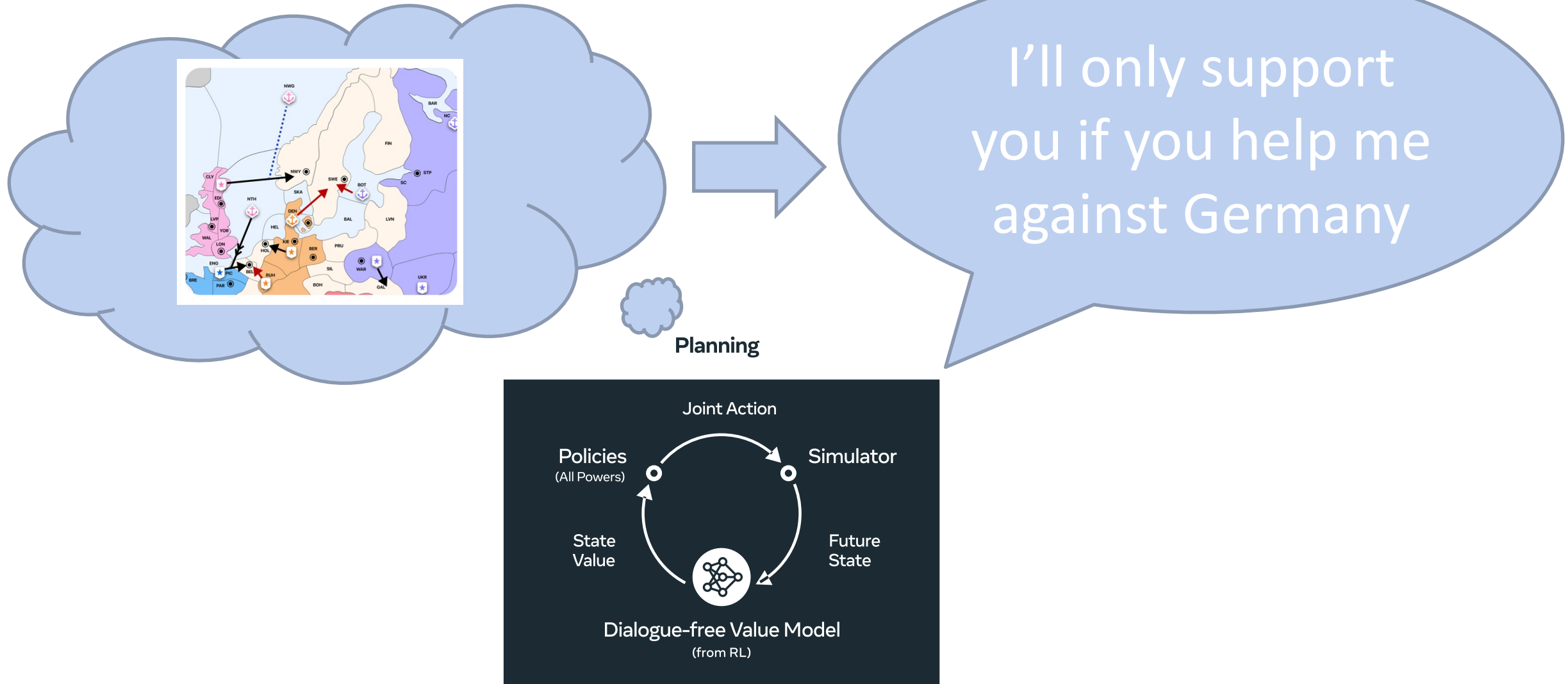
Intent-Controlled Dialogue



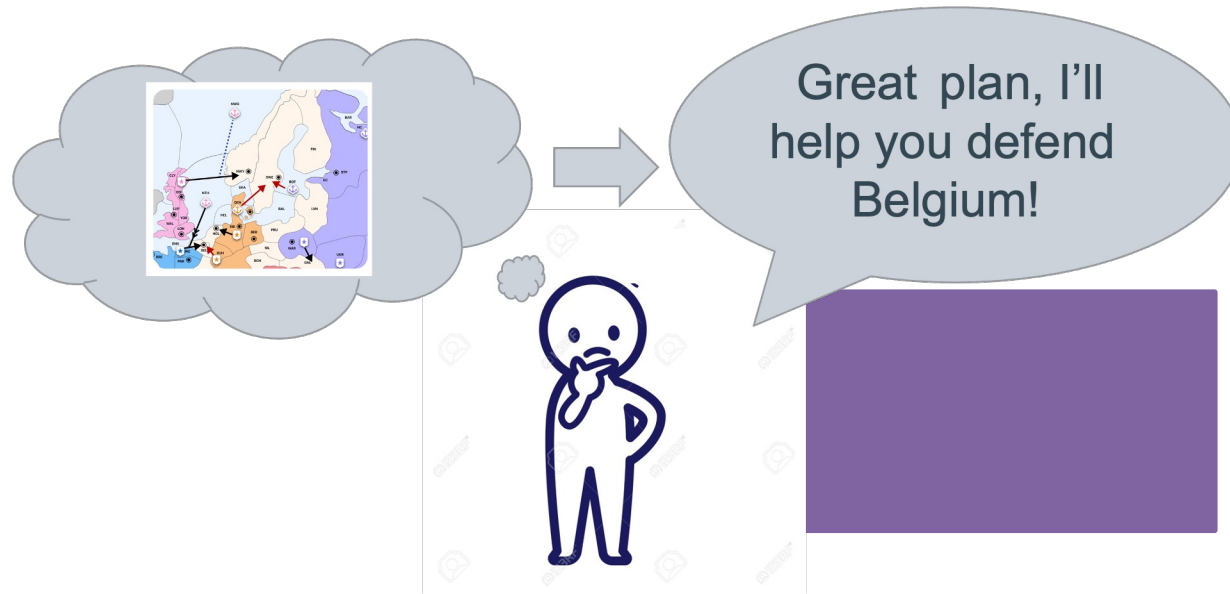
Great plan, I'll help
you defend
Belgium!



Intent-Controlled Dialogue



(1) Label Training Messages with Intents

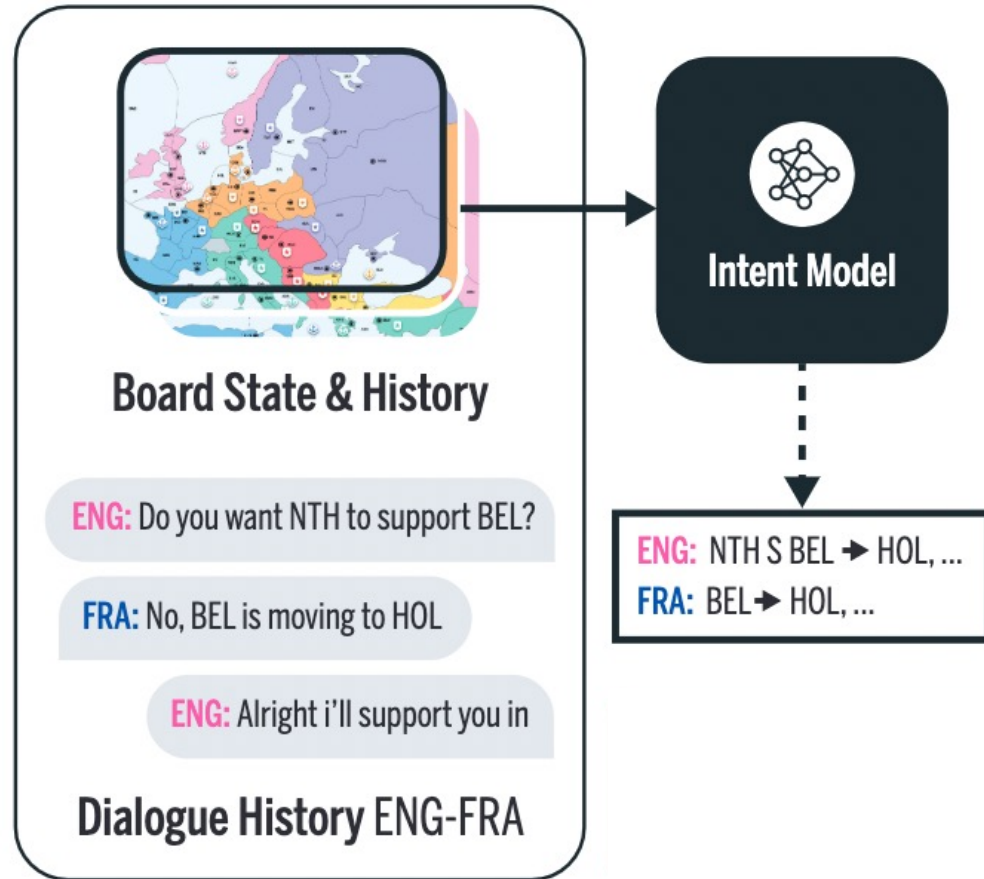


Define **intent** of a message as *the actions honest players would take if they agree to this proposal*

How can we infer these automatically?

(1) Label Training Messages with Intents

(a) Intent Model Training

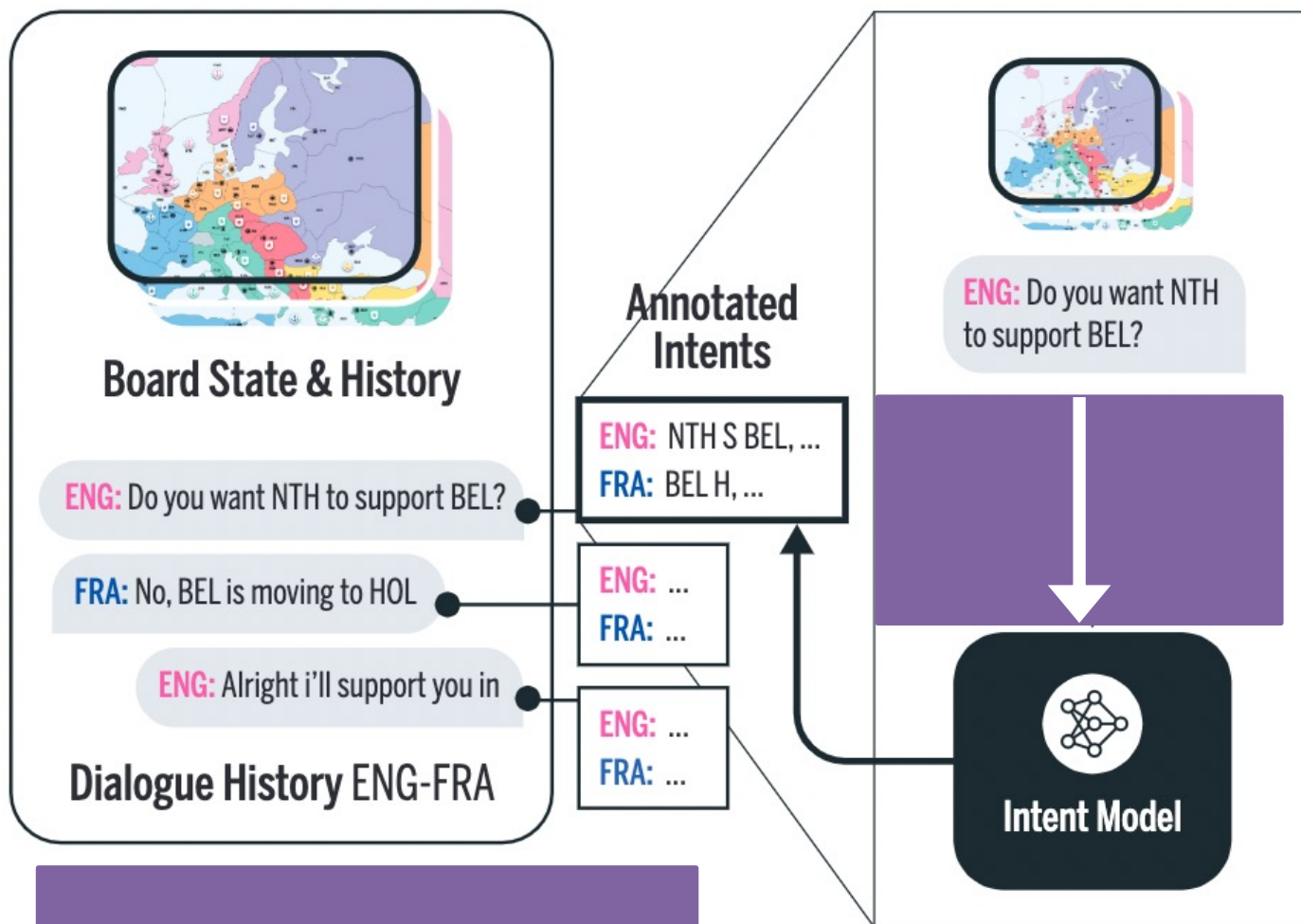


Intent Model

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

(1) Label Training Messages with Intents

(b) Intent Annotation



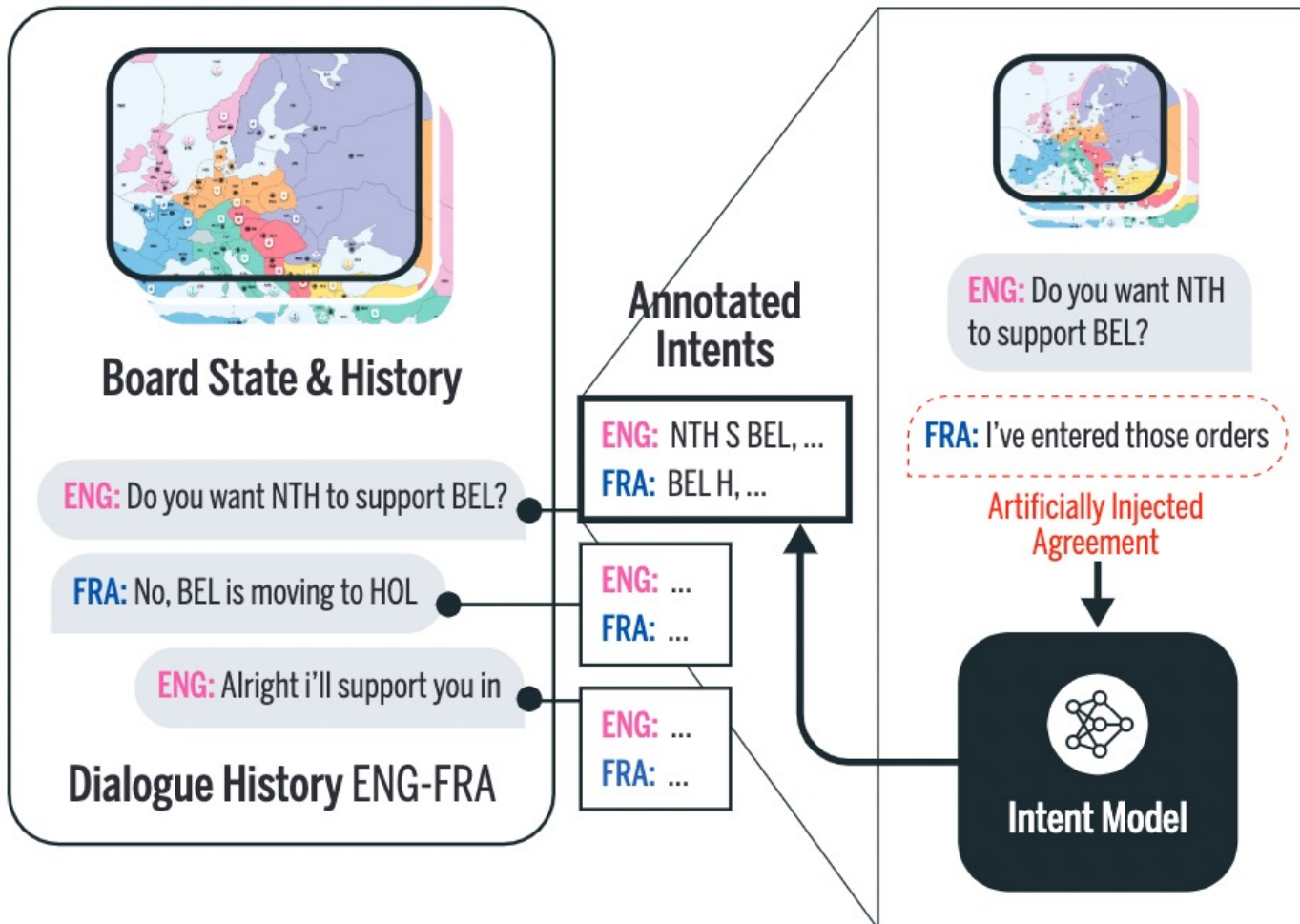
Intent Model

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



Method	% of predictions containing labeled orders
Base model	77
+ Initialized from dialogue model	87
+ Injected agreement messages	93
+ Restriction to truthful subset	97

Intent-Controlled Dialogue

