The emergence of Gricean maxims in goal-oriented interaction

(or, how Pragbot learned to play well with others)

Max Bodoia, Adam Vogel, and Christopher Potts

SUBTLE MURI review, Penn, October 12, 2012

Plan

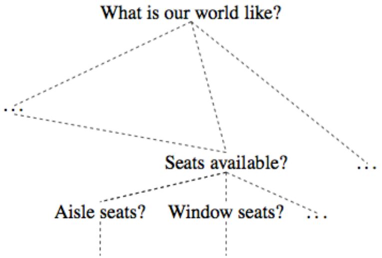
1. Release 2 of the Cards corpus (1,266 transcripts; ≈260,000 words)
2. Grounded interpretation in the Cards world
3. Decision-theoretic agents:
   - Listener bot
   - Dialogue bot
   - And presenting, at last, a true Pragbot

Pragmatics and the question under discussion

(10) Traveler Are there aisle seats available on the 7:30 flight?

Agent There are *seats* available. (That's all that matters.)

(11)



Task oriented dialogue corpora

Corpus	Task type	Domain	Task-orient.	Dialogues	Format
Switchboard	discussion	open	very loose	2,400	aud/txt
SCARE	search	3d world	tight	15	aud/vid/txt
TRAINS	routes	map	tight	120	aud/txt
Map Task	routes	map	tight	128	aud/vid/txt

(See also Blaylock & Allen, 'Generating artificial corpora for plan recognition')

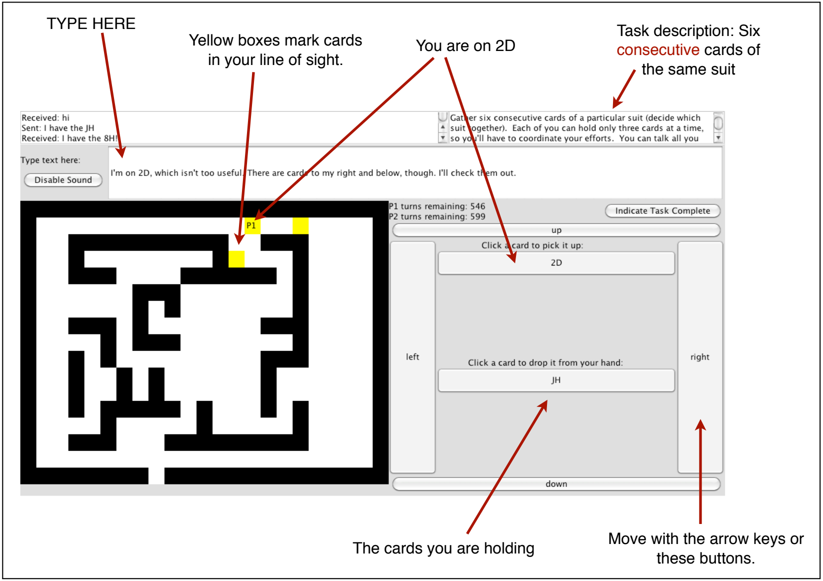
The Cards Corpus

`http://CardsCorpus.christopherpotts.net/`

Included:

- The transcripts in CSV format
- Python classes for working with the transcripts
- Examples of the Python classes in action
- R code for reading in the corpus as a data frame
- **All the annotations used in the work described here**

Gameboard


TYPE HERE
Yellow boxes mark cards in your line of sight.
You are on 2D
Task description: Six consecutive cards of the same suit
Received: hi
Sent: I have the JH
Received: I have the 8H!
Gather six consecutive cards of a particular suit (decide which suit together). Each of you can hold only three cards at a time, so you'll have to coordinate your efforts. You can talk all you
Type text here:
Disable Sound
I'm on 2D, which isn't too useful. There are cards to my right and below, though. I'll check them out.
P1
P1 turns remaining: 546
P2 turns remaining: 599
Indicate Task Complete
up
Click a card to pick it up:
2D
left
right
Click a card to drop it from your hand:
JH
down
The cards you are holding
Move with the arrow keys or these buttons.

Scenario

Gather six consecutive cards of a particular suit (decide which suit together), or determine that this is impossible. Each of you can hold only three cards at a time, so you'll have to coordinate your efforts. You can talk all you want, but you can make only a limited number of moves.

Version 1 numbers

- 744 transcripts
- Game length mean: 414.44 actions (median 325.50, sd 261.88)
- Actions:
  - Card pickup: 11,027
  - Card drop: 7,202
  - Move: 255,734
  - Utterance: 23,532
    - Utterance length mean: 5.84 words (median 5, sd 5.08)
    - Total word count: 137,323
    - Total vocabulary: 4,004 (3,453 if card references are normalized)

Version 2 numbers

- 1,266 transcripts
- Game length mean: 373.21 actions (median 305, sd 215.20)
- Actions:
  - Card pickup: 19,157
  - Card drop: 12,325
  - Move: 371,811
  - Utterance: 45,805
    - Utterance length mean: 5.69 words (median 5, sd 4.74)
    - Total word count: 260,788
    - Total vocabulary: 3,398 (assumes regularized card references)

Constants and points of variation

Constants

- Task description ('six consecutive cards of the same suit')
- Max cards in hand: 3

Randomness

- Players' initial positions
- All card positions

Constants and points of variation

Systematic variation (highlights)

- Some game are infeasible because areas of the board are walled off.
- Most games are symmetric: the players each have the same line of sight and number of moves.
- Around 500 games are asymmetric: one player has a very limited number of moves but infinite line-of-sight; the other has a large number of moves but very limited line of sight.
- There are a few different games boards, in a few different sizes.
- The number of moves each player has varies from 100 to 600; these values result in very different play.

Task oriented dialogue corpora

Corpus	Task type	Domain	Task-orient.	Dialogues	Format
Switchboard	discussion	open	very loose	2,400	aud/txt
SCARE	search	3d world	tight	15	aud/vid/txt
TRAINS	routes	map	tight	120	aud/txt
Map Task	routes	map	tight	128	aud/vid/txt
Cards	search	2d grid	tight	1,266	txt **in context**

Task oriented dialogue corpora

Corpus	Task type	Domain	Task-orient.	Dialogues	Format
Switchboard	discussion	open	very loose	2,400	aud/txt
SCARE	search	3d world	tight	15	aud/vid/txt
TRAINS	routes	map	tight	120	aud/txt
Map Task	routes	map	tight	128	aud/vid/txt
Cards	search	2d grid	tight	1,266	txt **in context**

Chief selling points for Cards:

- Pretty large.
- Controlled enough that similar things happen often.
- Very highly structured — the only corpus whose release version allows the user to replay all games with perfect fidelity.

Interpreting locative expressions

Example (Players establishing a basic search strategy)

Player 2: where are you?
Player 2: I am on the left side in the middle.
Player 1: i'm on the right side in the middle
Player 2: want to split left and right?

Example (Players seeking to exchange cards)

Player 2: ok i am going to drop a card for you
Player 1: ok, where at?
Player 2: i am at the very bottom right in front of the long skinny corridor across the bottom
Player 2: right at the opening on the left side

Variation

bottom
bottom left
left middle
I'm dead center
in the middleish
i'm inside the dead-end hallway
in the small box at the left center
I am in the middle toward the bottom.
far right, 7 blocks up from the bottom
I am in the middle just under the C room
i'm in the narrow room in the upper left
im inside the sideways C at the top left
The bottom left corner above the second line
I am two squares away from the upper left corner
I am just to the left of the C room in the middle.
I am in the long rectangle towards the bottom center
I am 2 spaces off the top 3 from left wall in the center
bottom right corner inside the box just below the single black square
i am at the very bottom right in front of the long skinny corridor across the bottom

Annotations

Player 2,23804,CHAT_MESSAGE_PREFIX,[where are you?]_{sem=where(you); engagesGoal=search}

Player 2,20236,CHAT_MESSAGE_PREFIX,hello, [where are you located?]_{sem=where(you); engagesGoal=search}

Player 2,13931,CHAT_MESSAGE_PREFIX,[where are you P?]_{sem=where(you); engagesGoal=meet}

Player 1,204774,CHAT_MESSAGE_PREFIX,[where are you exactly]_{sem=exactly(where(you)); engagesGoal=card}

(a) Questions.

Player 2,31931,CHAT_MESSAGE_PREFIX,I am [on the left side in the middle]_{sem=located(Player 2; @left middle); answers=23804}

Player 1,69058,CHAT_MESSAGE_PREFIX,hi....i am [in the lower right corner of the center of the board]_{sem=located(Player 1; @middle<bottom right>)); answers=20236}

Player 1,22344,CHAT_MESSAGE_PREFIX, i am [on the left bottom corner]_{sem=located(Player 1; @bottom left corner); answers=13931}

i'm [at the very very top right]_{sem=located(Player 1; @precise top right); answers=imp(where(you)); engagesGoal=card}

(b) Answers.

Table: Some of the 599 annotated locative answers, with associated questions.

Annotations

Text	Domain	Semantics
middle box towards the right up top	middle	approx;right
3 left of the c	BOARD	4.8;5.4;6.4
to middle box	BOARD	middle;room
on the left bottom corner	BOARD	bottom;corner;left
in the bottom you see the opening on the bottom row	BOARD	bottom;entrance
in the center towards the bottom	BOARD	approx;bottom;middle
middle right	BOARD	middle;right

Table: Example answer annotations. (Showing just three of the 27 columns in the extracted annotations file.)

Word	Count	Word	Count	Word	Count	Word	Count
BOARD	547	corner	91	hall	31	U_room	2
right	227	approx	77	room	18	T_room	2
middle	195	SQUARE	71	sideways_C	11	deadend	2
top	183	precise	68	loop	7	wall	1
left	178	entrance	59	reverse_C	3	sideways_F	1
bottom	169	C_room	35				

Table: Semantic lexicon with token counts.

Numerically

$$\textsc{board}(\textbf{top} \wedge \textbf{right})$$

	6	4	2	3	5	7
			3	0	1	1
	2		3	0	4	3
	1		1	2	0	1
	0		2	1	4	1
	0		1	0	2	0
			0	0	0	0
...	0		0	0	0	0
			0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	1	0
	0		0	1	0	0
	0		0	0	0	0
			0	0	0	0
	0	0	0	0	0	0

(a) Counts.

	2.30	2.54	2.51	2.54	2.61	2.65
			2.27	2.31	2.40	2.45
	0.89		1.93	1.97	2.04	2.09
	0.67		1.57	1.60	1.65	1.69
	0.36		1.22	1.24	1.26	1.28
	0.24		0.85	0.86	0.88	0.89
			0.52	0.52	0.53	0.54
...	0.00		0.28	0.27	0.28	0.29
			0.14	0.14	0.15	0.16
	0.03	0.04	0.08	0.10	0.11	0.12
	0.02	0.05	0.08	0.10	0.11	0.11
	0.01		0.10	0.10	0.11	0.11
	0.00		0.08	0.08	0.08	0.08
			0.05	0.05	0.05	0.05
	0.00	0.01	0.03	0.03	0.03	0.03

(b) Smoothed.

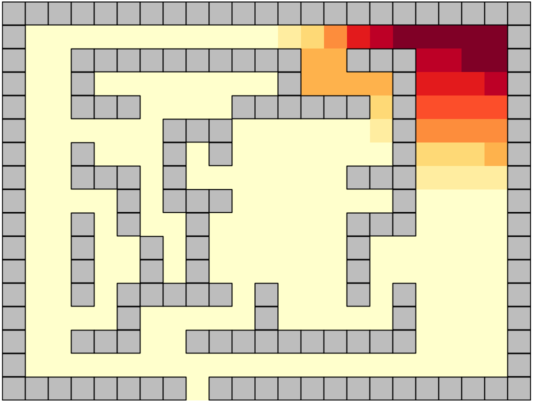
	0.035	0.038	0.038	0.038	0.039	0.040
			0.034	0.035	0.036	0.037
	0.013		0.029	0.030	0.031	0.031
	0.010		0.024	0.024	0.025	0.025
	0.005		0.018	0.019	0.019	0.019
	0.004		0.013	0.013	0.013	0.013
			0.008	0.008	0.008	0.008
...	0.000		0.004	0.004	0.004	0.004
			0.002	0.002	0.002	0.002
	0.000	0.001	0.001	0.002	0.002	0.002
	0.000	0.001	0.001	0.002	0.002	0.002
	0.000		0.001	0.002	0.002	0.002
	0.000		0.001	0.001	0.001	0.001
			0.001	0.001	0.001	0.001
	0.000	0.000	0.000	0.000	0.000	0.000

(c) Probabilities.

Table: The rightmost seven columns of the gameboard. The blank spaces correspond to walls (undefined values). The counts table shows an outliner four rows from the bottom.

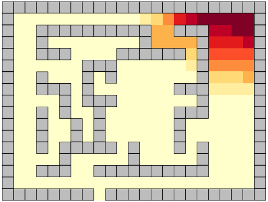
Heatmaps

BOARD(top & right); H: 5.68

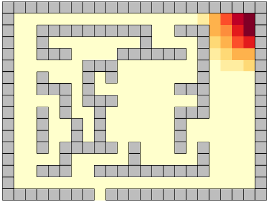


Heatmaps

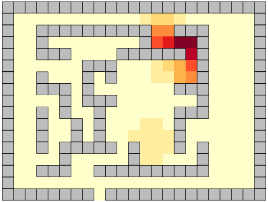
BOARD(top & right); *H*: 5.68



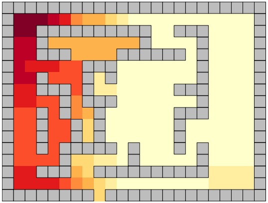
BOARD(precise & top & right); *H*: 4.4



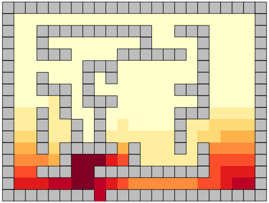
middle(top & right); *H*: 5.27



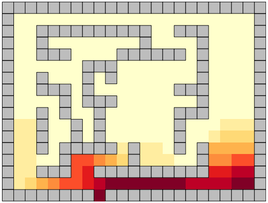
BOARD(left); *H*: 6.82



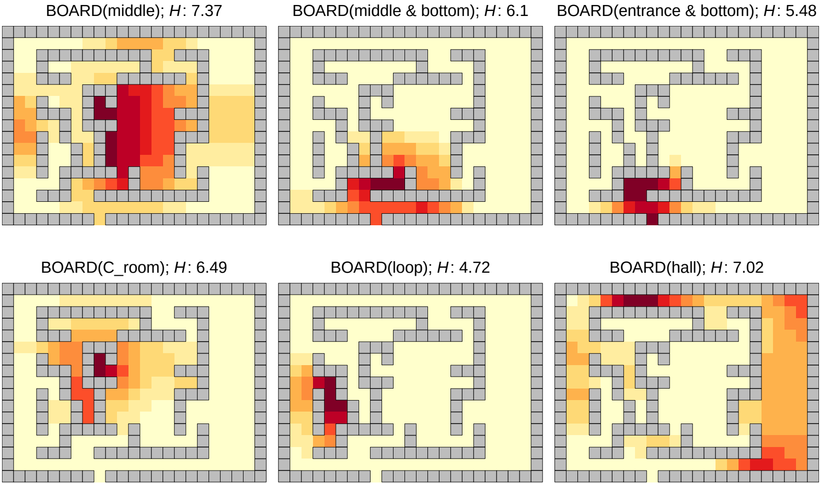
BOARD(bottom); *H*: 6.54



BOARD(precise & bottom); *H*: 6.05



Heatmaps


BOARD(middle); H: 7.37
BOARD(middle & bottom); H: 6.1
BOARD(entrance & bottom); H: 5.48
BOARD(C_room); H: 6.49
BOARD(loop); H: 4.72
BOARD(hall); H: 7.02

Experimental result (Potts 2012, WCCFL)

Specificity hypotheses (informal versions)

1. When the players need to meet up or direct each other to specific cards, their answers will tend to be more specific.
2. When the players are developing a general search strategy, their answers will tend to be less specific.

Example (search)

Player 1: on the left

Example (card)

Player 1: at the very top right corner

Experimental result (Potts 2012, WCCFL)

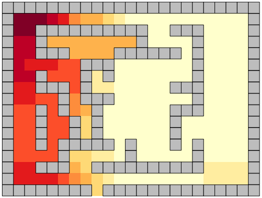
Specificity hypothesis (experimental version)

Answers to questions annotated as `engagesGoal=search` will tend to have higher entropy (*H*) than answers to questions annotated as `engagesGoal=card` or `engagesGoal=meet`.

Example (search)

Player 1: on the left

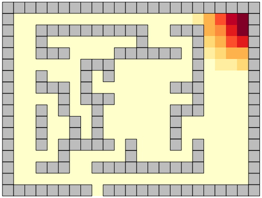
BOARD(left); *H*: 6.82



Example (card)

Player 1: at the very top right corner

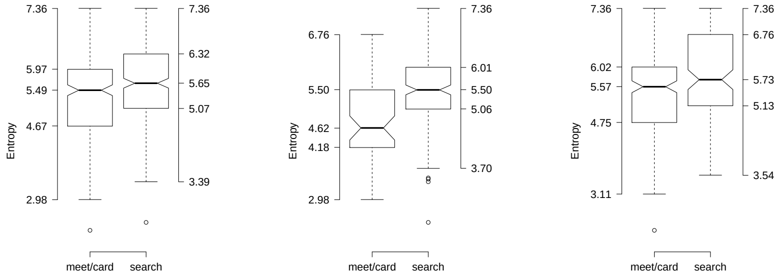
BOARD(precise & top & right); *H*: 4.4



Experimental result (Potts 2012, WCCFL)

Specificity hypothesis (experimental version)

Answers to questions annotated as `engagesGoal=search` will tend to have higher entropy (*H*) than answers to questions annotated as `engagesGoal=card` or `engagesGoal=meet`.



(a) The full annotated set. (b) Symmetric games only. (c) Asymmetric only.

Figure: The distribution of entropy values relative to task types.

Decision-theoretic agents

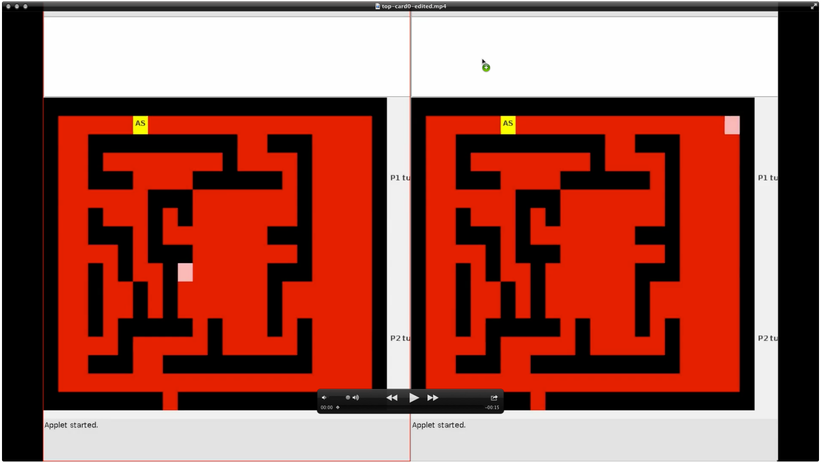
- The previous result supports the idea that people's answers are goal-oriented, a version of the general question-driven model of pragmatics that we began with.
- But what does it mean to be goal-oriented?
- And *how* does cooperative linguistic behavior emerge from the goals?
- More generally: to what extent do general Gricean pressures on rational communication emerge from basic decision-theoretic considerations?

Outline

What we want	How we get it
Semantic representations	Linear classifiers
Grounded language interpretation	As in the previous section
Decision-theoretic agents	Partially Observable Markov Decision Processes (POMDPs)
Language as a representation for planning	Generalizing from human behavior in the corpus
Modeling others beliefs	Add others' beliefs to the state space

Simplified cards scenario

Both players must find the ace of spades.



Semantic representations

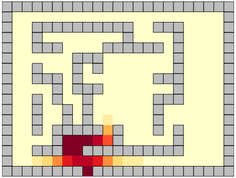
- Utterances are bags of words (for now). No preprocessing (yet) for spelling correction, lemmatization, etc.
- Assign semantic tags using log-linear classifiers trained on the corpus data.
- Binary classifiers for semantic tags and a multi-class classifier for domain tags.
- Micro-averaged F1 (10-fold cross-validation) is 81.9% — but the real test is how well it works in the extrinsic (bot vs. bot) evaluation I describe later.

Grounded language interpretation

“in the bottom you see the opening on the bottom row”

⇓

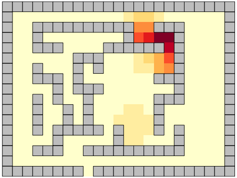
BOARD(entrance & bottom); H: 5.48



“in the top right of the middle part of the board”

⇓

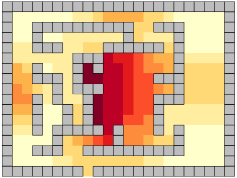
middle(top & right); H: 5.27



“i’m in the center”

⇓

BOARD(middle); H: 7.37



Agent framework

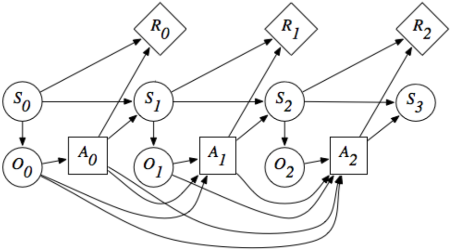
We want our agent to:

- Make moves that are likely to lead it to the card.
- Change its behavior based on observations it receives.
- Respond to locative advice from the other player.
- Give locative advice to the other player.

Modeling the problem as a POMDP allows us to train agents that have these properties.

POMDPs

- S: states
- b_0: initial belief state (distribution over S)
- A: actions
- O: observations
- T: distributions $P(s'|s, a)$
- Ω: distributions $P(o|s, a)$
- R: rewards $(S \times A) \mapsto \mathbb{R}$



Approximate solutions take us (only) part of the way

- An exact solution specifies the value of every action at any reachable belief state.
- In practice, only approximate solutions are tractable. We used the PERSEUS solution algorithm.
- Even approximate solutions are generally only possible for problems with $<$ 10K states.

Card location		Agent location		Partner location		Partner's card beliefs
231	$\times$	231	$\times$	231	$\times$	231
		$\approx$ 50K		$\approx$12M		$\approx$3B

Table: Size of the state-space for the one-card game.

Language as a representation for planning

- Divide the board up into *n* regions, for some tractable *n*
- Generate this partition using our locative phrase distributions.
- *k*-means clustering in locative phrase space.

231

	(0,0)	(0,1)	...	(9,6)	(9,7)	(9,8)	(9,9)	...	(17,23)
BOARD(bottom;entrance):	0	0	.01	.03	.04	.05	.04	.12	0
BOARD(middle;top):	0	0	.02	.02	.03	.02	.02	.07	0
middle(middle):	0	0	.02	.07	.08	.09	.10	.02	0
BOARD(bottom;corner;right):	0	0	.03	.06	.05	.06	.07	.03	0
...	0	0	.11	.02	.01	.01	.01	.06	0
	0	0	.03	.13	.13	.14	.12	.01	0
	0	0	.07	.04	.05	.06	.07	.01	0
	0	0	.01	.01	.02	.01	.01	.04	0
	...	...	...	...	...	...	...	...	...

130

Clusters induced

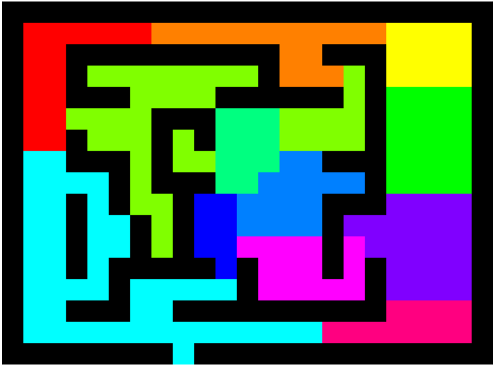


Figure: 12-cell clustering.

Clusters induced

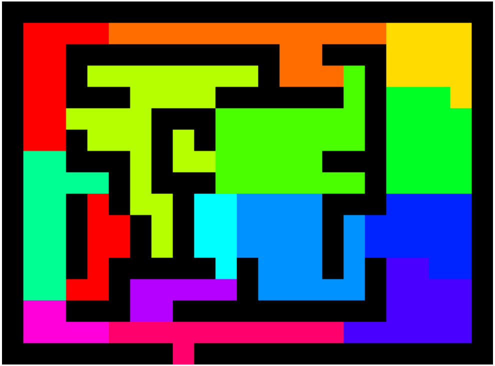


Figure: 14-cell clustering.

Clusters induced

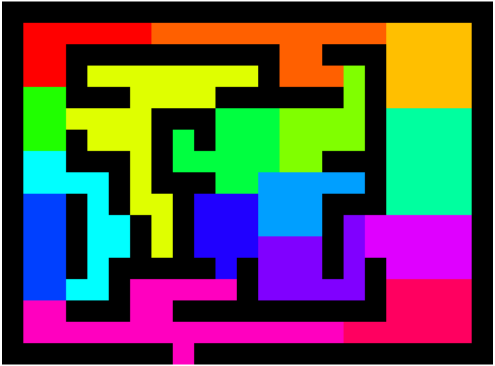


Figure: 16-cell clustering.

Clusters induced

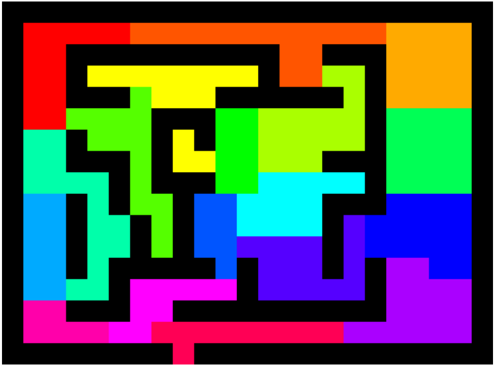
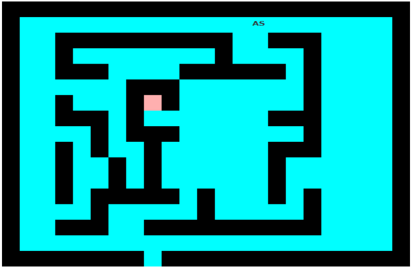


Figure: 18-cell clustering.

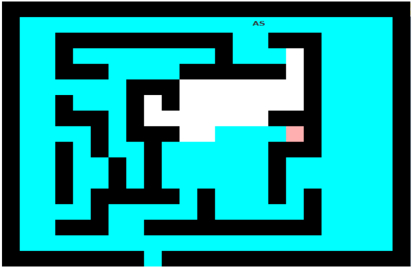
Listener Bot

- S: all combinations of the player's region and the card's region
- b_0: initial belief state (distribution over S)
- A: *travel* actions for each region, and a single *search* action
- O: {AS seen, AS not seen}
- T: distributions $P(s'|s, a)$, except *travel* actions fail between nonadjacent regions
- Ω: distributions $P(o|s, a)$; *travel* actions never return positive observations; *search* actions return positive observations only if the player's current region contains the AS
- R: small negative for not being on the card, large positive for being on it. No sensitivity to the other player.

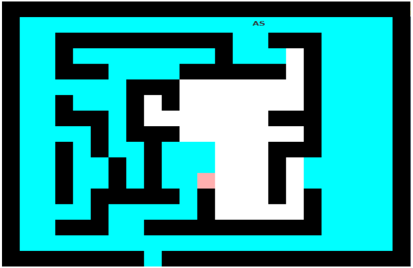
Listener Bot example



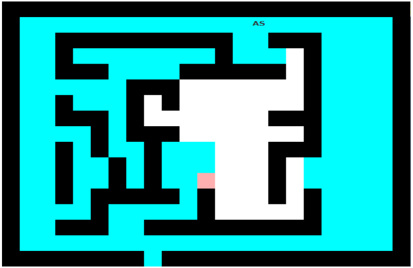
Listener Bot example



Listener Bot example



Listener Bot example

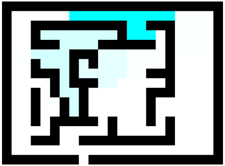


"by the middle of the top hallway"

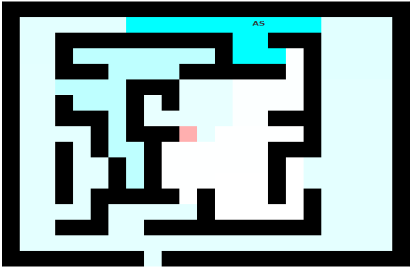
$\Downarrow$

$\textsc{board}(\textbf{middle} \wedge \textbf{top})$

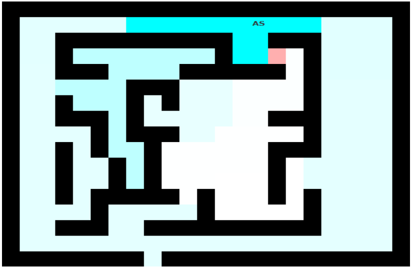
$\Downarrow$



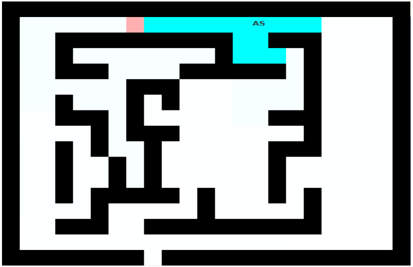
Listener Bot example



Listener Bot example



Listener Bot example



ListenerBot home movies

- “in the room on the left of the board”
- “in the top west corner”
- “in the top entrance to the middle room”
- “in the middle of the top hallway”
- “on the right ... the right of the middle room that is”
- “in the middle of the right hallway”
- “in the corner”
- “at the entrance to the main room”
- “halfway up”

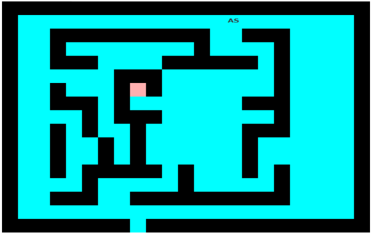
DialogBot (Baby Pragbot?)

DialogBot is a strict extension of Listener Bot:

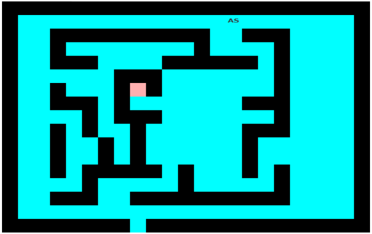
- The set of states is now all combinations of
  - both players' positions
  - the card's region
  - the region the other player believes the card to be in
- The set of actions now includes dialogue actions.
- (The player assumes that) a dialogue action U alters the other player's beliefs in the same way that U would impact his own beliefs.
- Same basic reward structure as for Listenerbot, except now also sensitive to whether the other player has found the card.

DialogBot and ListenerBot play together

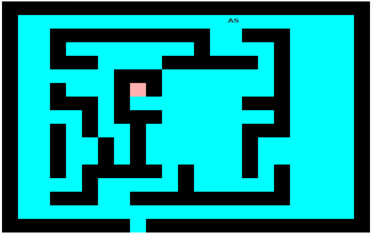
DialogBot beliefs



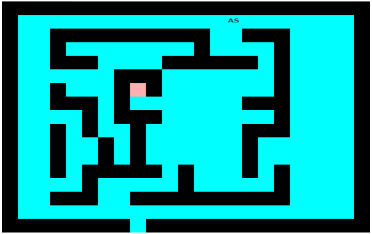
ListenerBot beliefs



DialogBot beliefs:
ListenerBot's position

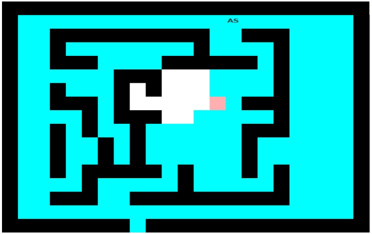


DialogBot beliefs:
ListenerBot's beliefs

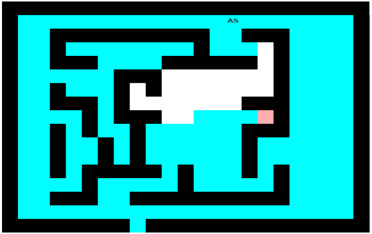


DialogBot and ListenerBot play together

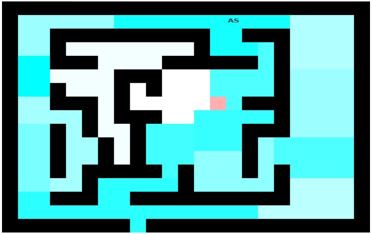
DialogBot beliefs



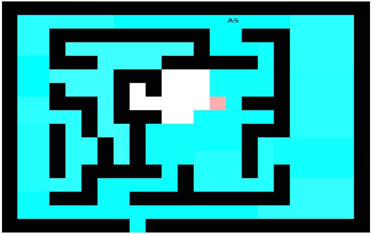
ListenerBot beliefs



DialogBot beliefs: ListenerBot's position

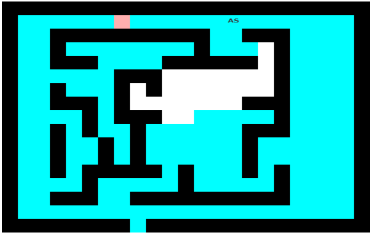


DialogBot beliefs: ListenerBot's beliefs

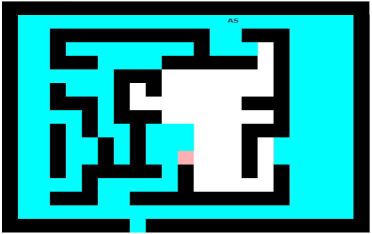


DialogBot and ListenerBot play together

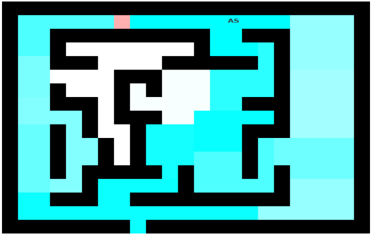
DialogBot beliefs



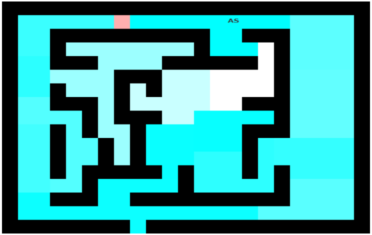
ListenerBot beliefs



DialogBot beliefs: ListenerBot's position

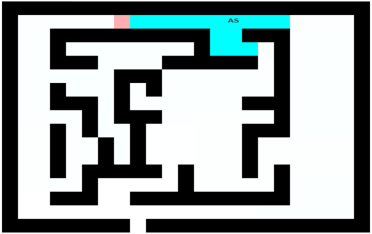


DialogBot beliefs: ListenerBot's beliefs

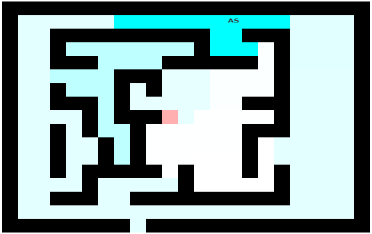


DialogBot and ListenerBot play together

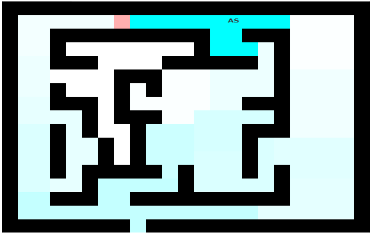
DialogBot beliefs



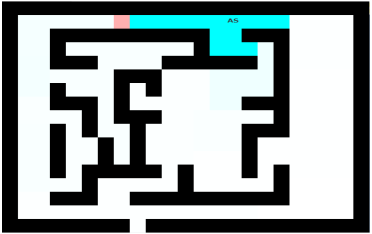
ListenerBot beliefs



DialogBot beliefs: ListenerBot's position

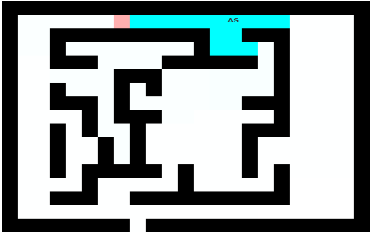


DialogBot beliefs: ListenerBot's beliefs

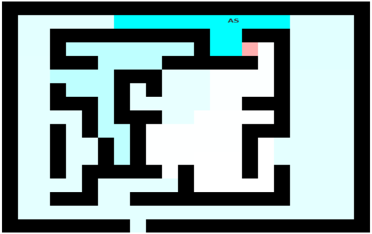


DialogBot and ListenerBot play together

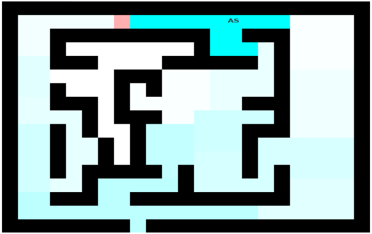
DialogBot beliefs



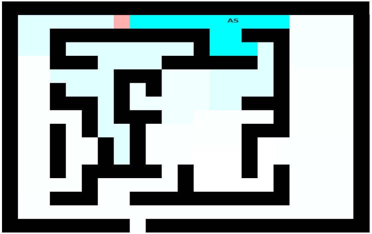
ListenerBot beliefs



DialogBot beliefs: ListenerBot's position

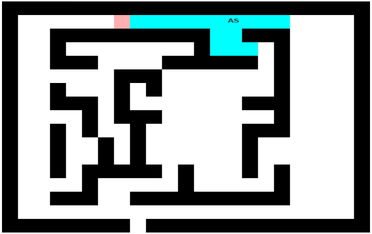


DialogBot beliefs: ListenerBot's beliefs

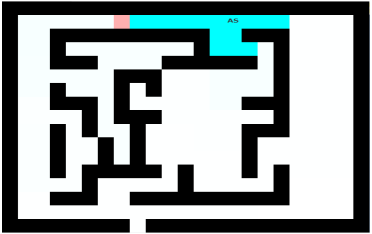


DialogBot and ListenerBot play together

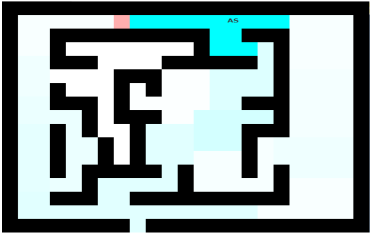
DialogBot beliefs



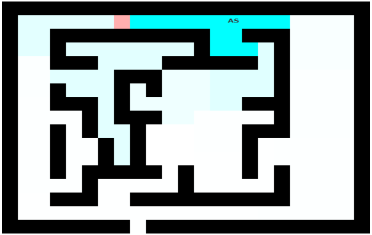
ListenerBot beliefs



DialogBot beliefs: ListenerBot's position

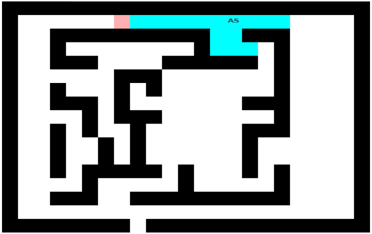


DialogBot beliefs: ListenerBot's beliefs

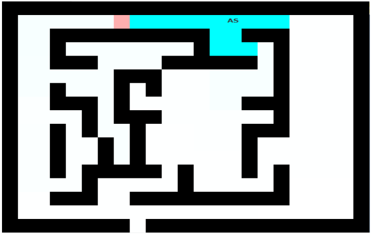


DialogBot and ListenerBot play together

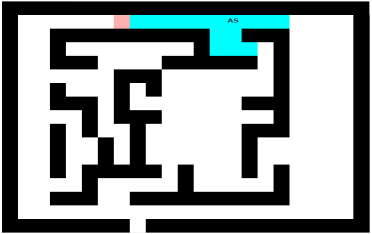
DialogBot beliefs



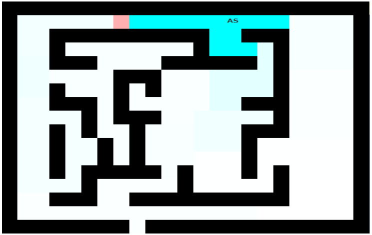
ListenerBot beliefs



DialogBot beliefs: ListenerBot's position



DialogBot beliefs: ListenerBot's beliefs



DialogBot home movies

- Right player finds the card on the top:
  - Players' card beliefs
  - Players' beliefs about other players' card beliefs
  - Players' beliefs about other players' locations
- Right player finds the card in the middle:
  - Players' card beliefs
  - Players' beliefs about other players' card beliefs
  - Players' beliefs about other players' locations

Emergent pragmatics

Quality

- The Gricean maxim of quality says roughly “Be truthful”.
- For DialogBot, this emerges from the decision problem: false information is (typically) more costly.
- DialogBot would lie if he thought it would move them toward the objective.

Quantity and Relevance

- The Gricean maxims of quantity and relevance for informative, timely contributions.
- When DialogBot finds the card, he communicates the information, not because he is hard-coded to do so, but rather because it will help the other player find it.

Experimental results

	Success rate	Moves (overall)	Moves (successful)
ListenerBot & ListenerBot	58.4%	18.90	14.20
ListenerBot & DialogBot	60.9%	19.37	15.45
DialogBot & DialogBot	90.9%	19.26	16.12

Table: Test pairs of agents from 197 random initial states.

Next steps

- Gricean manner says “Be clear and concise”. This should emerge from the right cost functions.

Next steps

- Gricean manner says “Be clear and concise”. This should emerge from the right cost functions.
- Richer utterances. At present, DialogBot keeps it simple, but we plan to bring generation into the pragmatic model.

Next steps

- Gricean manner says “Be clear and concise”. This should emerge from the right cost functions.
- Richer utterances. At present, DialogBot keeps it simple, but we plan to bring generation into the pragmatic model.
- Training with human subjects. Will Turkers be willing to help DialogBot through bad policies?

Next steps

- Gricean manner says “Be clear and concise”. This should emerge from the right cost functions.
- Richer utterances. At present, DialogBot keeps it simple, but we plan to bring generation into the pragmatic model.
- Training with human subjects. Will Turkers be willing to help DialogBot through bad policies?
- We have evidence that online pragmatic reasoning can require beliefs about beliefs about beliefs, but that it goes no deeper.

Next steps

- Gricean manner says “Be clear and concise”. This should emerge from the right cost functions.
- Richer utterances. At present, DialogBot keeps it simple, but we plan to bring generation into the pragmatic model.
- Training with human subjects. Will Turkers be willing to help DialogBot through bad policies?
- We have evidence that online pragmatic reasoning can require beliefs about beliefs about beliefs, but that it goes no deeper.
- Thus, a bot that reasoned to this level might be truly worthy of the name Pragbot.