Word Games

J Chiu

March 9, 2022

Dialogue

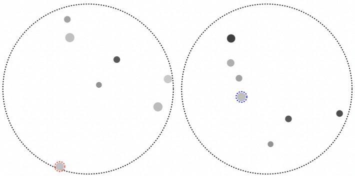
- Communication is rarely unambiguous
  - Ambiguity resolution through dialogue
  - Clarification questions
- Interactive, symmetric reference games
  - Isolates ambiguity resolution
  - Both give and request information

Games

Friends of agent A:

Name	School	Major	Company
Jessica	Columbia	Computer Science	Google
Josh	Columbia	Linguistics	Google
...	...	...	...

A: Hi! Most of my friends work for Google
B: do you have anyone who went to columbia?
A: *Hello?*
A: I have Jessica a friend of mine
A: and Josh, both went to columbia
B: *or anyone working at apple?*
B: SELECT (Jessica, Columbia, Computer Science, Google)
A: SELECT (Jessica, Columbia, Computer Science, Google)



Human A's view Human B's view

Human B:	three light grey dots in a diagonal line
Human A:	i dont have that but i have a black dot neer the top to the right, the only black dot in the circle
Human B:	i have two black dots. find something else
Human A:	ok i have a light grey dot by itself at the bottom to the left. right on the line
Human B:	how big is it
Human A:	its one of the bigger ones
Human B:	okay just pick it then
Human A:	ok
Human B:	SELECT blue
Human A:	SELECT red

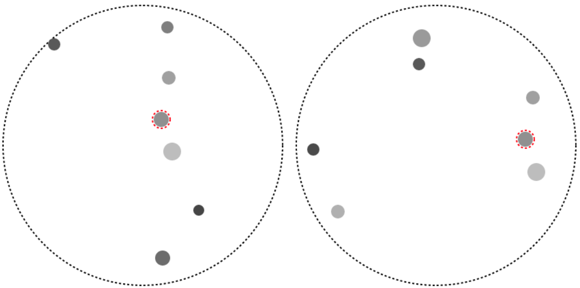
Mutual Friends and OneCommon

Issue: Poor neural reasoning

From Mutual Friends: Neural + Human

- A: Know anyone who likes chess?
- B: None of my friends like chess.
- (conversation continues)
- A: Crocheting?
- B: None like crocheting.
- A: Chess?
- B: None like chess either, haha.

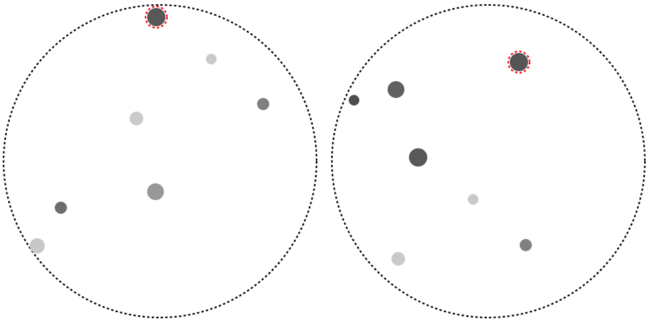
Issue: Poor neural reasoning



Agent 0: human || 1: human

	0	1
0	1	I see a large grey dot with a smaller black dot right below it
1	0	is the smaller black dot to the right and below
2	1	No, it is slightly to the left and below
3	0	Might not be the same one. Do you have a lot of dots in a fairly vertical conformation
4	1	I have three dots in a kinda vertical line with different shade of gray but close in size
5	0	Is the middle dot curved to the left or right a little bit?
6	1	the middle dot is slighty to the left of the other two and is also a little bit darker than the rest
7	0	Yes pick that middle dot

Issue: Poor neural reasoning



Agent 0: pragmatic_confidence || 1: human

	0	1
0	0	i have one large black dot by itself . do you have it ?
1	1	Yes, I do have that.
2	0	let 's pick that one
3	1	ok
4	0	ok

Issue: Scaling rule-based

- Rule-based text generation and understanding is somewhat viable for Mutual Friends
  - Very optimistic selection, but can be tuned
- Continuous and spatial nature of OneCommon makes writing rules difficult
  - Size, color, and positions all continuous
  - Descriptions are relative

Current approaches: Two extremes

- Neural encoder-decoder
  - Encode past interactions with a neural net
  - Generate what to say with a neural net
  - Brittle strategy, less brittle language
- Rule-based
  - Encode past interactions in a table
  - Use rules for what to say next
  - Nonparametric lookup of utterances
  - Brittle language, less brittle strategy
- Meet in middle with interpretable planning + neural language

A dialogue turn

- Engaging in dialogue requires
  - Inference: What do I know? How do I represent it?
  - Planning: What should I do and say?
- Formulate as model-based planning
  - Plan what to say through a simple model of our partner
  - Model of partner conditions on past information

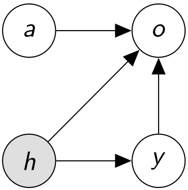
Model-based planning

- Goal: Use a supervised static model in an interactive setting
- Belief state (What do I know)
  - Interpretable summary of past information that is relevant to task
  - Model latent partner information
  - Able to enforce constraints
- Model-based planning (What should I say)
  - Pick best action by imagining how partner would respond
  - Can train partner response models on static data
  - Allows policies to perform better than the data they trained on
  - Need a measure of utility to determine which action to take (heuristic: reduce uncertainty)

Belief model

- Latent quantity y
- Actions a_t,observations o_t (e.g. yes/no questions)
- Interaction history $h_t = (a_0, o_0, \dots, a_t, o_t)$ contains all previous actions and observations
- Given an initial belief $p(y \mid h_t)$ + next action/observation, obtain next belief via

$$p(y \mid h_{t+1}) = p(y \mid h_t, a_{t+1}, o_{t+1}) \propto \underbrace{p(o_{t+1} \mid h_t, a_{t+1}, y)}_{\text{observation model}} p(y \mid h_t)$$



Belief state

- Use belief state $p(y \mid h)$ to capture information relevant to planning
- Most of the burden is on the observation model $p(o \mid h, a, y)$
  - Must be able to predict partner
- Prior work assumes conditional independence $p(o \mid h, a, y) = p(o \mid a, y)$
- Why is this bad?

Conditional independence in partner modeling

- Prior work assumes conditional independence $p(o \mid h, a, y) = p(o \mid a, y)$
- If you ask the same question twice, your belief changes both times!
  - $p(\text{yes} \mid h = \emptyset, y)$ can vary depending on the latent y
  - $p(\text{yes} \mid h = (\text{red dot?}, \text{yes}), a = \text{red dot?}, y) = 1$, since we just asked!
- 'Questions with correlated answers' and deficient observation model lead to uncalibrated beliefs, and therefore poor strategy
- Contribution: relax independence assumption
  - Use copy attention from past answers to current answer
  - Probably solved by using Transformer (is there enough data)

Belief calibration: Other issues

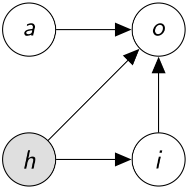
-

Planning: Use prior work

- Goal: Mutually select the same item y as partner
  - Row in knowledge base, dot
  - Coordinate through dialogue
- Given history h, we need to chose an action a by optimizing utility
$$\max_a U(h, a)$$
- Utility $U =$ information gain $+$ utterance $+$ pragmatic cost
  - IG: Entropy reduction of item selection probability
  - Utterance cost: Can't send a full paragraph
  - Pragmatic cost: Want utterance to be accurate
- Ideally estimate and optimize future reward directly
  - Heuristic approximation of future reward U
  - Limited-horizon planning to minimize impact of model error

Information Gain

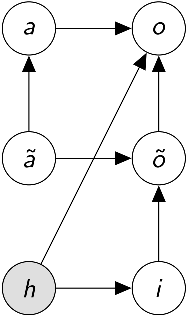
- A good action should decrease uncertainty
- Requires
  - Belief distribution over selection item given history $p(i \mid h)$
  - Partner response model $p(o \mid h, a, i)$
- Represent a turn as



- Language and planning coupled

Decoupling language and planning

- Compress actions a and observations o into language and abstract representations $\tilde{a}, \tilde{o}$
  - Language is high dimensional, redundant, and inefficient for planning
- Represent a turn as



- Abstract observation $\tilde{o} \perp\!\!\!\perp h \mid \tilde{a}, i$

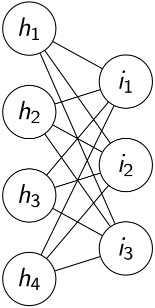
State and belief: Representation

- History: whether attributes have been confirmed $h \in \{0, 1\}^N$
- Items: $i \in [M], M << N$
- Logistic regression with attributes as features

$$p(i \mid h) = \frac{\exp(\sum_n \psi(h_n, i))}{\sum_{i'} \exp(\sum_n \psi(h_n, i'))}$$

$$\psi(h_n, i) = W_{ni} 1(h_n(i))$$

- Generate per-game $W = f(\text{context})$ from neural network
  - Many correlated features
  - How to (conditionally) sparsify?
- Dialogue = online feature selection



Attributes

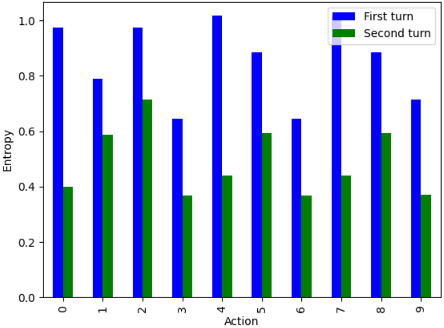
- Mutual Friends
  - Combinations of columns of knowledge base
  - Name, major, company
- OneCommon
  - Which dots are mentioned
  - Need to learn lower-level attributes
- Numerical reasoning?

Experiments

- Mutual Friends
  - Augment rule-based (prior work) to optimize info gain
  - After OneCommon: Add neural on top
- OneCommon
  - Use attributes = raw mention configurations
    - Need belief / info gain / LR weights
    - How to deal with redundancy? (i.e. correlation between features)
  - Learn latent refinement on top of mention configurations

Information gain issues

- Best info gain could be to ask the same question twice
- Usual fix: Limit to asking once only
- Would be nice to have a principled way to deal with correlated features though



- Second turn after taking action with lowest entropy

Related work: 20 questions

- Padmakumar and Mooney (2020)
  - Attribute-based classification (string heuristic to map to description) + activate learning about attributes
  - Info gain (on top of binary unweighted logistic regression) as feature for RL policy
- Yu et al. (2019)
  - Question-based classification (attributes)
  - Learn weights of features
  - Do not consider feature correlations
- More interesting language, symmetric setting
- Learn weights, account for correlation
- Symmetry, deal with unexpected features

End

Concerns

- Would a large LM solve all of this?
  - Fine tune on small onecommon dataset, are there still repeats?
  - Unlikely to solve strategy / over optimistism

End

Value: Information Gain

- drop slide
- Picture would be much better here...
- Value = expected information gain

$$IG(h, a) = H(i \mid h) - \mathbb{E}_{p(o|h,a)}\left[H(i \mid h, a, o)\right]$$

$$\mathbb{E}_{p(o|h,a)}\left[H(i \mid h, a)\right] = \sum_{o} \sum_{i'} p(o \mid h, a, i) p(i \mid h) H(i \mid h, a, o)$$

- Equivalent to minimizing expected uncertainty after receiving a response
- Cite Yu et al, White et al

Citations I

Padmakumar, A. and Mooney, R. J. (2020). Dialog policy learning for joint clarification and active learning queries. *CoRR*, abs/2006.05456.

Yu, L., Chen, H., Wang, S. I., Artzi, Y., and Lei, T. (2019). Interactive classification by asking informative questions. *CoRR*, abs/1911.03598.