Word Games

J Chiu

March 15, 2022

Dialogue and information gathering

- Resolve ambiguity and coordinate through dialogue
- OneCommon: Interactive, symmetric reference game
  - Isolates info gathering (and coordination)
  - Environment (dots) are completely static
  - Dynamism comes from dialogue only
- 20 questions with symmetric information constraints

Previous SotA

- Purely supervised
  -
- Uncalibrated beliefs: overconfidence
  - Pushes for to select a dot that will not work
- Research goal: Improve purely supervised models via model-based planning

Fixing strategy with planning

- Prior: Fully supervised neural encoder-decoder
  - Encode past interactions with a neural net
  - Generate what to say with a neural net
  - Brittle strategy, less brittle language
- Next: Model-based planning
  - Choose what to say by imagining how partner would respond
  - Say utterance with best expected outcome
  - Potentially stronger player than expert demonstrations

Challenges in model-based planning

- Partner modeling is hard
  - Variable amount of information
  - Random strategies
- Multi-turn planning
  - Accuracy of planning depends greatly on the partner model
  - Errors from the partner model will compound over time
- Single-turn planning
  - Removes compounding errors
  - Optimize a dialogue progress heuristic: uncertainty reduction
  - Requires belief
- Use what dots partner also sees as belief

Planning

- Plan by imaging partner response

$$\max_u \mathbb{E}_{p(r|h,u)} [\text{Utility}(h, u, r)]$$

- Utterance u, response r, history h
- Utility should approximate dialogue progress
  - Goal of dialogue is information gathering and coordination
  - Focus on information gathering
- Utility a function of belief

Planning with Belief

- Introduce belief state $p(s \mid h)$
  - State s is what dots partner can also see
- Incorporate belief in planning

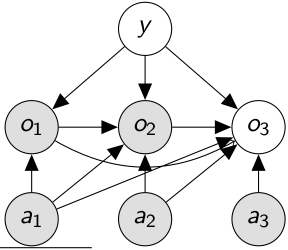
$$\max_u \mathbb{E}_{p(r \mid h,u,s)p(s \mid h)} \left[\text{Certainty}(p(s \mid h, u, r))\right]$$

- Obtain

$$p(s \mid h, u, r) = \frac{p(r \mid h, u, s)p(s \mid h)}{\sum_s p(r \mid h, u, s)p(s \mid h)}$$

Belief model

- Static latent quantity y: which dots do they also see
  - Alternative: actual field of view
- Actions a_t, observations o_t
  - Prior work: yes/no questions (20 questions)
  - OneCommon: unrestricted language for both
- Uniform prior $p(y)$ over 7 choose 4 dots partner also sees
- Observation model $p(o_t \mid a_{1:t-1}, o_{1:t-1}, y)$
  - Pick y that is observed during training
  - Ideally fully supervised
  - Prior work[1] makes naive Bayes assumption $p(o_t \mid a_t, y)$



[1]Yu et al. (2019); Padmakumar and Mooney (2020)

Belief update

- Interaction history $h_t = (a_0, o_0, \ldots, a_t, o_t)$ contains all previous actions and observations
- Given an initial belief $p(y \mid h_t)$ + next action/observation, obtain next belief via

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

- Belief calibration depends on accuracy of observation model

Observation model

- Example exchange
  - Action: Do you see a red dot?
  - Observation: No, but I see a blue one.
- Utterances are multifaceted
  - Responses contain more information than asked
  - New information injected by partner due symmetric roles
  - Very difficult to model new information
- Simplifying assumption: only model response, not new information
  - Update belief state afterwards by pretending we asked corresponding question
  - Allows reduction to 20 questions / assymetric role
- Supervised training needs observed o, h, a, y
  - Main question: How to extract responses o?

Response extraction

- Heuristic: Use repeated mentions from response
  - Do you see a red dot? Yes, the one next to the blue one?
- Generalization: TBD
- Recap: we have
  - Belief over shared dots $p(y \mid h)$
  - Observation model $p(o \mid h, a, y)$
  - Update $p(y \mid h, a, o)$
  - Reduced to assymmetric case by extracting response only
- Next: Single-turn planning

Planning: Use prior work in assymmetric setting

- Given history h, we need to chose an action a by optimizing heuristic utility

$$\max_a U(h, a)$$

- Utility $U =$ information gain - utterance - pragmatic cost
  - IG: Reduce uncertainty
  - Utterance cost: Can't send a full paragraph
  - Pragmatic cost: Want utterance to be accurate
- Ideally would estimate and optimize future reward directly
  - Heuristic approximation of future reward U
  - Limited-horizon planning to minimize impact of model error

Expected information gain

- Maximizing expected information gain equivalent to minimizing uncertainty

$$\min_a \sum_o \sum_y \underbrace{p(o \mid h, a, y)}_{\text{observation model}} \underbrace{p(y \mid h)}_{\text{belief}} \text{Uncertainty}(\underbrace{p(y \mid h, a, o)}_{\text{new belief}})$$

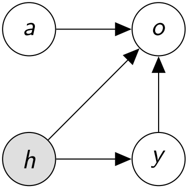
Summary

- Goal: Extend methods from 20 questions to symmetric, language setting
- Extract relevant information from partner utterances
- Use explicit belief state + single-turn planning heuristic

End

Information Gain

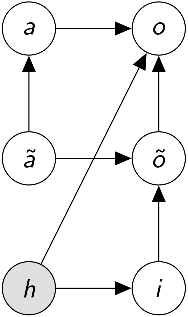
- A good action should decrease uncertainty
- Requires
  - Belief distribution over selection item given history $p(y \mid h)$
  - Partner response model $p(o \mid h, a, y)$
- Represent a turn as



- 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



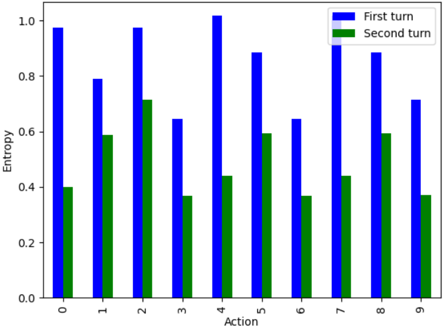
- Abstract observation $\tilde{o} \perp\!\!\!\perp h \mid \tilde{a}, i$

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

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.

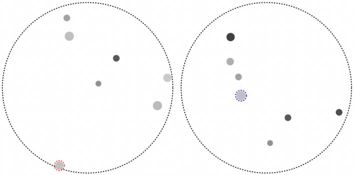
Yu, L., Chen, H., Wang, S. I., Artzi, Y., and Lei, T. (2019). Interactive classification by asking informative questions. *CoRR*, abs/1911.03598.

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.

Sample of prior work in model-based planning

- 20 questions (Yu et al., 2019; Padmakumar and Mooney, 2020)
  - Sym: Assymmetric questioner + answerer
  - Turns: Multi-turn game
  - Lang: Closed class answers (observations)
  - Heur: Expected info gain heuristic
- EVPI (**??**)
  - Sym: Assymmetric questioner + answerer
  - Turns: No interaction, single turn game
  - Lang: Open
  - Heur: Expected utility heuristic
- RSA reference game (**?**)
  - Sym: Symmetric
  - Turns: Multi-turn game
  - Lang: Symbolic language
  - Heur: Bounded depth search

Conditioning in partner modeling

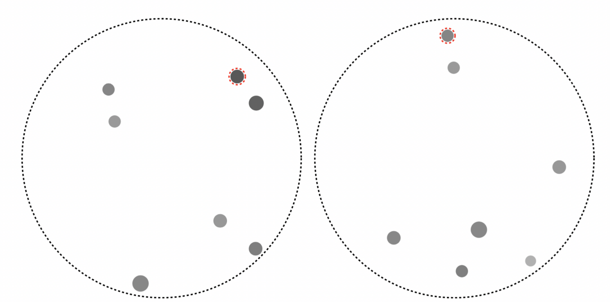
- Assuming conditional independence $p(o \mid h, a, y) = p(o \mid a, y)$ is harmful
- If you ask the same question twice, your belief changes both times!
  - $p(\text{yes} \mid h = \emptyset, a = \text{red dot?}, y)$ can vary depending on the latent y
  - $p(\text{yes} \mid h = (\text{red dot?}, \text{yes}), a = \text{red dot?}, y) = 1$, since we just asked!
- 'Questions with correlated answers' and deficient observation model lead to uncalibrated beliefs, and therefore poor strategy
- Contribution: relax independence assumption
  - Let past obs vote on current one (weighted by action similarity)
  - Probably solved by Transformers[2]

[2] Copy attention, depends on amount of data

Example dialogue 1: Overconfidence

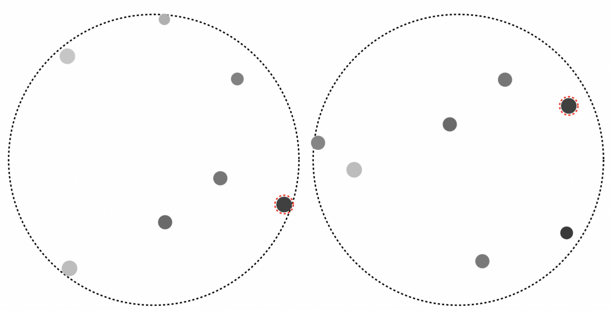
0	i have a large black dot , it 's not the same size , but it is not the darkest
1	Mine is the darkest
0	i see it . i think it is the one .
1	I have multiple dark dots so I am not sure
0	i have a large black dot , it is in the center
1	I have a large dot in the lower part with a dark smaller dot at seven o'clock
0	is the black dot the darkest ? if so click it
1	I have three the same darkness
0	pick the lightest one
1	Ok but our odds are not good.
0	pick the light one

Example dialogue 2: Overconfidence



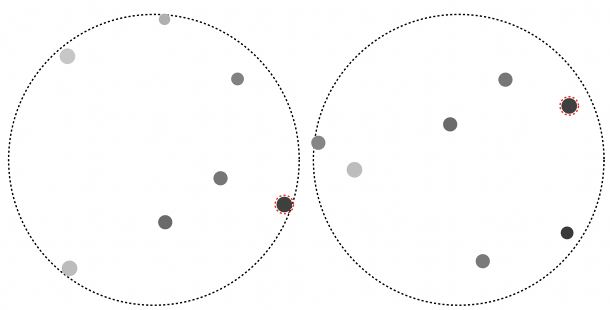
0	i have two dark dots , one on top and slightly smaller than the other
1	i see it. pick the top one?
0	ok

Example dialogue 3: Good humans



1	I have a large black dot by itself
0	I see a large, very dark dot on the edge of my screen (so I won't be able to see anything to its right). Can you see anything on the left of your large black dot?
1	Yes, my large dark dot is on the edge of the right side
0	Ok, to the left of the dark dot, and slightly above it, do you see a slightly-smaller, slightly lighter dot?
1	yes

Example dialogue 3: Good humans



0	and then far above (and a bit to the right of) that lighter one, do you see a slightly smaller, identically colored dot?
1	No, the first lighter dot is the closest dot to the top of mine
0	Okay. what do you see to the left of that lighter dot?
1	A slightly darker dot that is below it just a bit
0	ok, I think we're in the same place. Let's click that original, blackest dot
1	Okay sounds good