Word Games

J Chiu

March 4, 2022

Dialogue

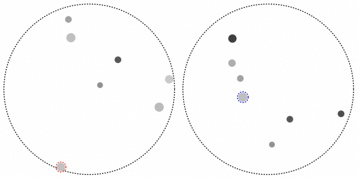
- Communication is rarely unambiguous
  - Ambiguity resolution through dialogue
  - Clarification questions
- Interactive, symmetric reference games
  - Isolates ambiguity resolution
  - Both give and request information

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

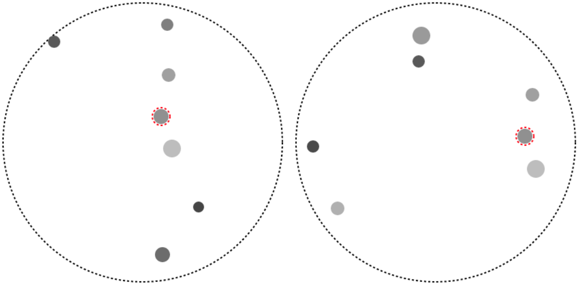
Mutual Friends and OneCommon

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.

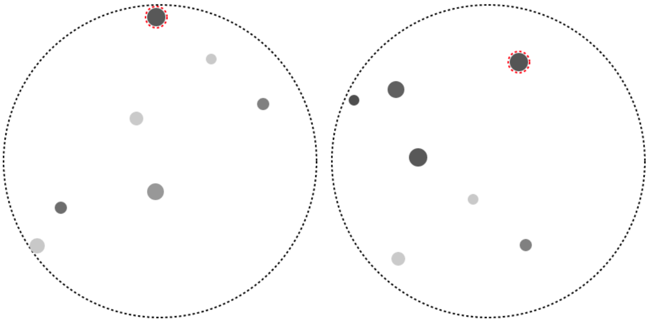
Issue: Poor neural reasoning



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

System	C	C_T	C_S
Human	.89	.07	.36
Rule	**.88**	**.06**	**.29**
StanoNet	.76	.04	.23
DynoNet	.87	.05	.27

- Rule-based text generation and understanding is viable for Mutual Friends
- Continuous and spatial nature of OneCommon makes writing rules difficult

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 belief state and planning

A dialogue turn

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

Problem setup

- Given history representation h, we need to chose an action a that maximizes the expected future reward
- Must solve

$$V(h) = \max_a \underbrace{\mathbb{E}_{p(s|h)}\left[R(s,a)\right]}_{\text{immediate reward}} + \underbrace{\mathbb{E}_{p(o|h,a)}\left[\tilde{V}(hao)\right]}_{\text{future reward}}$$

- Two questions we will answer
  - How do we represent h, a, o?
  - How do we compute future reward $\tilde{V}$?
- But first look at past work under this lens

Past approaches: Rule-based

$$V(h) = \max_a \underbrace{\mathbb{E}_{p(s|h)}\left[R(s,a)\right]}_{\text{immediate reward}} + \underbrace{\mathbb{E}_{p(o|h,a)}\left[\tilde{V}(hao)\right]}_{\text{future reward}}$$

- History encoding h: Table with mention counts of attributes
- Action a: Attribute, utterance from a template
- Reward R: Defined by rule
- Observation o: Not used
- Future reward $\tilde{V}$: Not used

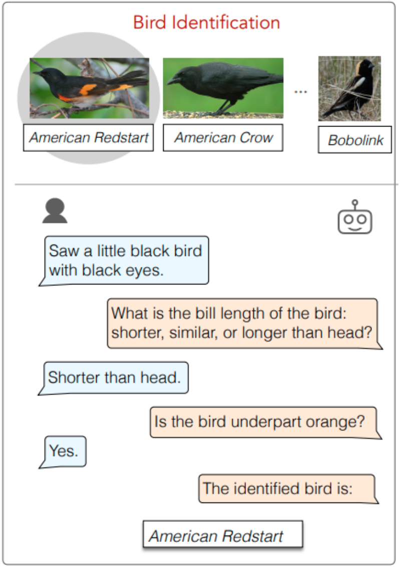
Past approaches: Neural

$$V(h) = \max_a \underbrace{\mathbb{E}_{p(s|h)}\left[R(s,a)\right]}_{\text{immediate reward}} + \underbrace{\mathbb{E}_{p(o|h,a)}\left[\tilde{V}(hao)\right]}_{\text{future reward}}$$

- History encoding h: Neural representation (hierarchical RNN, graph NN)
- Action $a = (i, u)$: First pick an item i to ask about, then produce an utterance $u \mid i$ from NN (or directly produce u from NN)
- Reward R: Specificity of utterance ($p(i \mid u)$) or just probability of utterance
- Observation o: Language (but not used)
- Future reward $\tilde{V}$: Not used

Segway: Prior work in 20 questions and belief-states

- Target items i
- Belief distribution $p(i)$



Past approaches: Belief-state (20 questions)

$$V(h) = \max_a \underbrace{\mathbb{E}_{p(s|h)}\left[R(s,a)\right]}_{\text{immediate reward}} + \underbrace{\mathbb{E}_{p(o|h,a)}\left[\tilde{V}(hao)\right]}_{\text{future reward}}$$

- History encoding h: Categorical over target items $p(i \mid h)$
- Action a: Finite set of questions
- Reward R: None
- Observation o: Yes/no or categorical response
- Future reward $\tilde{V}$: Information gain for $p(i \mid hao)$

Belief representation

- Just considering target items is not enough
- If ask about the same item, should ask in a way that improves information
  - Pushes all work onto observation model $p(o \mid h, a) = \sum_i p(o \mid i, a) p(i \mid h)$
  - Additionally, representation of *hao* may not be adequate
- There should be no information gain from taking the same action multiple times
- Need a belief state based on attributes to prevent repeats

Proposal

End

Reference Dialogue Games

- Simplest form of collaborative task-oriented dialogue
- Convey a referrent to your dialogue partner
  - Ex: Jointly selecting a star
- Successful joint attention = task success
- Simplest form: Information gathering via discriminative questions
  - Can also include coordination (OC, MutualFriends) focus on this
  - Or even navigation (Cards)

Solving Reference Games

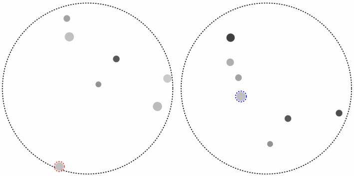
- Multiple agents with equal capabilities
  - Pro: More efficient task completion
  - Con: May be more difficult to reason about other agents' belief states ('agent uncertainty')
- Natural language dialogue
  - Pro: Great for interfacing with non-experts
  - Con: Extremely large search space when planning
- Planning in a structured action / observation space
  - Pro: Computationally efficient planning
  - Con: Must convert to and from language, leading to 'sensor uncertainty'

Research question

- How can we achieve joint attention, a necessary first step towards collaboration?
- Hypothesis: We can solve joint attention with efficient and accurate planning
  - Past work planning directly in language is inefficient and inaccurate
  - Reference games admit a natural structured action / observation space
- Challenge: Planning is computationally difficult in multi-agent systems
  - Large number of belief states
  - Difficult to calibrate sensor uncertainty with language

OneCommon

Collaborative reference game where the goal is to find one dot both players have in common, given different but overlapping views



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

Method

- Separate language (sensors) from planning
- Formalize problem as a decentralized POMDP
  - Players are performing hypothesis testing together
- Sources of uncertainty
  - Language input (sensors)
  - Language output (generation errors)
  - Other agents
- Focus on uncertainty induced by other agents
  - Is modeling other agents' beliefs necessary?

Research Questions Summary

- Goal: Solve joint attention
- Question: Can we solve joint attention by utilizing structure to improve the efficiency and accuracy of planning?
- Additionally: When is modeling other agents' beliefs necessary?

Current Status

- POMDP formulation as best-arm identification (Simplification, ignores dec)
  - Symmetric actions and observations for both partners (Good!)
  - Actions and observations are structured
  - Include extra information as a ‘free turn’ in POMDP
  - Ignores partner belief
- Planning ignores structure (inefficient, but ok for now)
- Solves small state space (each can see 3 dots with 1 shared) and all dots unique / easily identified
- Next step: See if partner belief modeling is necessary when dots are not as easily identified

Deprecated below

See belief writeup for more up to date stuff

Related Work

- Lots of work decoupling actions from language (to avoid catastrophic forgetting)
- Aligning text to (continuous) actions (Zhao and Eskénazi, 2018)
- Modern implementation of Cards (Vogel et al., 2013)?
  - Is the Dec necessary? (How different is I'm sure from they're sure)
  - Is the PO part important, i.e. can we use certainty equivalence? (low risk, how often do edge cases occur?)
  - Amortized inference, better approximation algorithms / search?

OneCommon: Properties

- Action space: Can be reduced to predicates over dots and boolean response to partner predicates
- Observations: Observations are symmetric with action space
- Agents: Only one other agent, which we can assume is similar
- Strategy: Mostly common to other reference games

Simplified OneCommon

We can solve the following POMDP almost exactly with value iteration

- Small horizon $T = 5$
- State: Beta distribution for each dot. Gives prior over our partner having a dot. Combine independent Beta parameters into Dirichlet for selection.
- Actions: We communicate directly in a single predicate over our dots (2^D different predicates)
- Observations: Partner can only respond to predicate noiselessly

We can almost solve the Bellman equation

- Convert to and MDP over belief state
- Reduce belief state to sufficient statistics of positive response counts for each dot (of size T^D)
- Transitions limited because we can only add 1 or 0 to counts
- Major issue is max over actions

Next Steps

Shorter term

- Align language to actions (predicates over our dots)
- Add noise to feedback (did they understand our action/predicate?)

Longer term

- Extend response from boolean (they can offer more information, ask questions)
- Approximate methods for larger state space (MCTS used in prior work with actions = language), i.e. MutualFriends

Longerer term

- Dec-POMDP

VERY OLD SLIDES. DO NOT CONTINUE

Motivation: Why OneCommon

- OneCommon is all about referring expressions (resolving / generating)
- Focus on improving reference resolver, text has more noise
- Improve generation using reference resolver
- Goal: Reduce as much learning as possible to reference resolver
- Impact: Need to carry over to referring expressions in other settings

Current OneCommon Performance

- How can we improve the reference resolver?
  - Best model gets 93% accuracy 78% exact match
  - Human is 96% accuracy 87% exact match
  - Does fine-tuning pretrained models improve accuracy of reference resolver?
  - How can we close the gap?
- Would a better reference resolver result in a better selection model?
  - Yes, training on true refs gives 99% accuracy selection model
  - Best model currently at 83%, humans at 91%
- Does fine-tuning pretrained models improve round-trip of reference generator?
  - Generation second priority
  - Likely over-fitting

Current Models for RE

- Fried 2021
  - Conditions on utterances and marginal dot features
  - Neural referent memory (updates dots on reference)
- Udagawa 2021
  - Conditions on utterances and marginal dot features (input structure via relational networks)

Improving RE

- Resolve multiple dots
  - Use number prediction to select explicit arity potentials
  - Not enough data? Gather more data with more ambiguity, i.e. dots all have the same size and color? Would this transfer (motivation)?
- Architectures that are more biased towards composition
  - DIORA
  - Transformers? (possibly with attention constraints)
- Reason explicitly about past references
  - Belief state / checklist
  - Pragmatic inference via mutual exclusivity (check if anything is referenced multiple times)
- Coreference
  - Unfamiliar, need further research
  - How often does this happen?
- Dialogue acts
  - Pragmatics of resolution vary based on dialogue act
  - Confirmation repeats vs new references
  - How to collect data for this?

Generalization / Transfer to Other Tasks

- Low-resource datasets that need more structure / bias
- Context is useful for understanding
  -
- Understanding is useful for generation

3 Challenges

- Modular models
  - Dialogue model = local meaning representation - belief state - dialog act - utterance generation
  - Various works supervise particular parts then leave others to be implicitly learned through neural nets
  - Results in very task-specific architectures
  - Can we break down tasks to allow for more component sharing across different tasks, as well as semi-supervised learning?
- Less complicated meaning representations
  - Meaning representations vary in granularity
  - Can we learn a minimal task-specific subset of a meaning representation formalism?
- Adapting to partners
  - Partner model allows forward modeling, learned over a multiple round game or repeated games
  - Adapt opaque neural model or hierarchical bayesian model?
- Better reference resolution / referring expression models?
  - Dialogue games are often data scarce and out of domain
  - Can we use models that are more sensitive / with stronger biases?

Natural Language Interaction

- Interaction (through language) is important
  - Cannot fully automate every task, i.e. task-oriented or information seeking dialogues require human input
  - Must handle diverse non-expert human input, although input may map to a low-dimensional manifold
  - High levels of ambiguity must be resolved via interaction
- Interaction (through language) is hard
  - Human input is expensive, so supervision is limited
  - In order to make certain problem aspects tractable, must make sacrifices in other areas (toy domain = out of distribution for pretrained models)
- What are the main challenges in interaction, and what are the tradeoffs of different approaches?

Types of Dialogue Games

- Task-oriented: Wizard of Oz (WoZ)
  - Tseng et al. (2019): Wizard obtains task from human then executes it.
- Deliberation / reference / signal
  - Udagawa and Aizawa (2019): Visual reference game with latent translated views. Each player gets a different petri dish view of the same underlying game board, and players must select the same object on the board.
- Information seeking / inquiry
  - Yu et al. (2019): WoZ-style answer providing where asker does not know exact question. Latent true question (to all), WoZ must answer
- Persuasion / negotiation
  - Lewis et al. (2017): Negotiation over an observed set of item with latent utilities for each agent.

Types of Dialogue Games

In all cases, the game can be (indirectly) solved by resolving a latent variable

- When is this tractable, and why do no new methods do this?
- New (ie basically all) methods rely on supervision
- If they do not, it is because the game has a trivial solution

Talk with Nori

- Planning doesnt really exist in most dialogue games, the games are too simple
- Referring expressions in OneCommon can be outliers in terms of complexity. Complexity in real world image+text data comes from large noun classes and relatively simple phrases. Still seems like an interesting testbed though, and maybe there is an argument for OneCommon, despite its artificial difficulty (maybe people like talking about stars?).
- Latent variable models for adaptation still seem interesting, should read Hawkins' work.

Types of Dialogue Games: Latent goals and strategy

What are the latent variables in each type?

- Task-oriented
  - Latent task slots
- Deliberation / reference / signal
  - Varies per game
- Information seeking / inquiry
  - Infer true question, find answer
- Persuasion / negotiation
  - Infer utilities, exploit

Types of Dialogue Games: Latent goals and strategy

- Tasks must be interesting enough so that latent quantities cannot be inferred with a single utterance, reducing them to single turn games
  - High degree of ambiguity / distractors or large number of slots to fill (combinatorial)
- Break down latent quantities and use heuristics to make assumptions on structure
  - For example, choosing an ordering of WoZ slots: When choosing a restaurant, first figure out time, then cuisine, and finally price
  - Will likely remain task-specific
- What other parts can we learn?

Belief State Tracking

- The incremental inference procedure is known as belief state tracking (BST)
- Local semantics are aggregated into belief state, which informs high-level strategic decisions
- Seems difficult to learn language, high level strategy, belief state updating, and low level parsing at the same time
- Ablate how structure influences each of these

Language Games

- Games offer a testbed for the development of methods
  - Allow designers to control difficulty and simplicity
- Allowing interaction through language increases the population of players

Meaning reps

- In full generality, this problem is often encountered in hierarchical RL
  - Less bleak in the language gamesetting
  - Games are often very simple and can be constrained to small horizons, for example He He engineered a parser and policy that basically solves the negotiation task
- Many text-specific meaning representations (MR) to choose from
  - Many are too complex
  - Can we leverage existing MRs to learn a minimal task-specific representation that balances utility and expressivity?

Contributions

- Under a unified Bayesian game perspective of dialog games, present formulations for different classes of dialogs: signaling, negotation?
- Provide pipelined variational Bayesian framework for learning to play dialog games from offline data, with and without granular annotations
- Good results

Tasks

- Task-oriented dialogue: Multi-WoZ
- Negotation: Deal-or-no-deal
- Reference: OneCommon
- Information-Seeking: Birds

Generation under uncertainty

- Lemon et al. (2010)

Citations I

Lemon, O., Janarthanam, S., and Rieser, V. (2010). Generation under uncertainty. In *Proceedings of the 6th International Natural Language Generation Conference*. Association for Computational Linguistics.

Lewis, M., Yarats, D., Dauphin, Y. N., Parikh, D., and Batra, D. (2017). Deal or no deal? end-to-end learning for negotiation dialogues. *CoRR*, abs/1706.05125.

Tseng, B., Rei, M., Budzianowski, P., Turner, R. E., Byrne, B., and Korhonen, A. (2019). Semi-supervised bootstrapping of dialogue state trackers for task oriented modelling. *CoRR*, abs/1911.11672.

Udagawa, T. and Aizawa, A. (2019). A natural language corpus of common grounding under continuous and partially-observable context. *CoRR*, abs/1907.03399.

Citations II

Vogel, A., Bodoia, M., Potts, C., and Jurafsky, D. (2013). Emergence of gricean maxims from multi-agent decision theory. In *NAACL*.

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

Zhao, T. and Eskénazi, M. (2018). Zero-shot dialog generation with cross-domain latent actions. *CoRR*, abs/1805.04803.