Pragmatic Dialogue Agents for Negotiation

Nicholas Tomlin (Advisor: Chris Potts)
Stanford CSLI: *Language, Cognition, and Computation* (REU Site #1659585)

Abstract

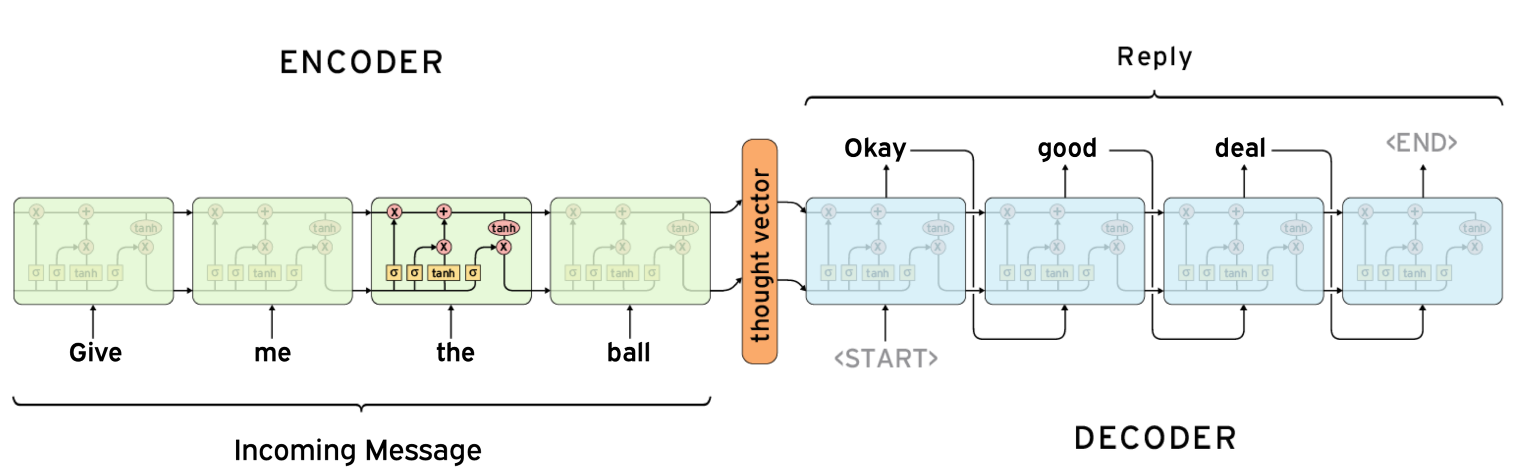
Artificial conversational agents are playing an increasingly large role in our daily lives, so it is important that they are attuned to our goals, rather than simply producing very common responses. Modern deep learning architectures for these agents rely on sequence-to-sequence models, which are trained to generate mappings between dialogue turns, outputting the most likely response given a conversational context. This poses an issue for goal-oriented dialogue systems, where the most likely response may not achieve the desired goal. We propose a solution via the Rational Speech Acts (RSA) model of pragmatic reasoning, in which agents reason recursively about utterances and intents in order to communicate more effectively with each other.

Rule-Based and Neural Chatbots

Conversational dialogue agents (chatbots) can be divided into two main categories: rule-based and neural agents. Rule-based chatbots (ELIZA, Cleverbot, Siri) rely on complex hand-coded systems based on slot-filling. The earliest of these, ELIZA, was created in the 1960s:

YOU: I am unhappy.
ELIZA: Are you unhappy often?
YOU: Yes.
ELIZA: Can you elaborate on that?
YOU: Too much work.
ELIZA: Why do you say too much work?

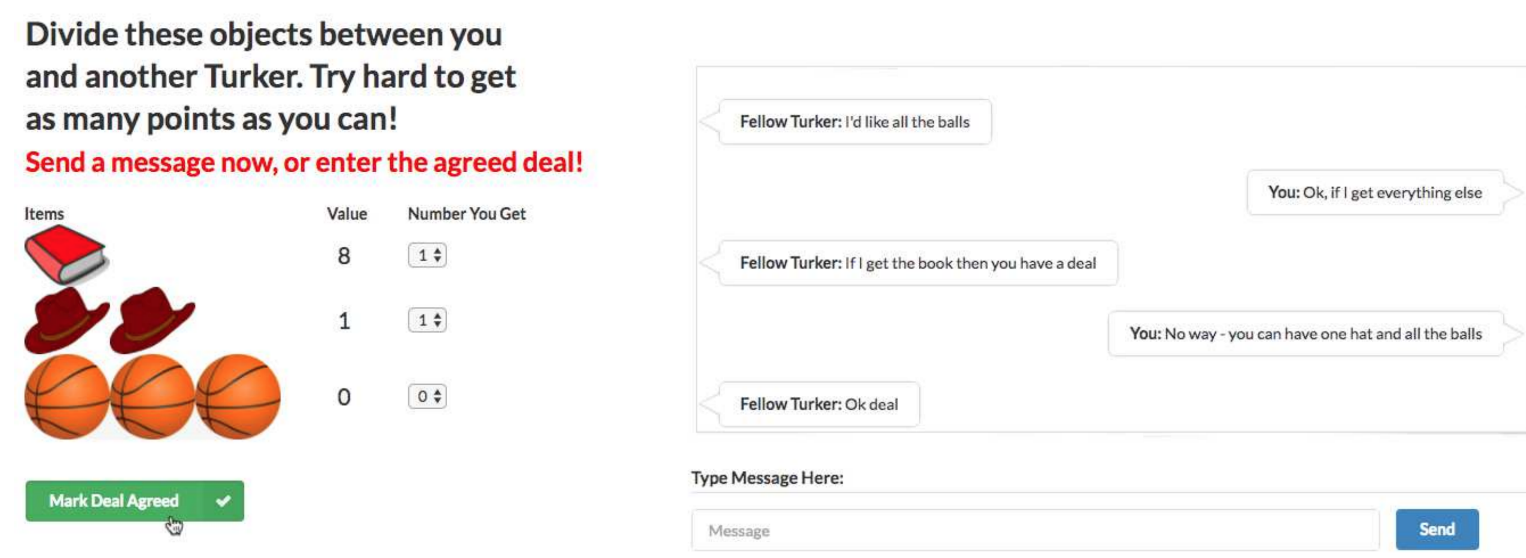
Rule-based chatbots typically search for particular common phrases or keywords and respond accordingly. Meanwhile, neural chatbots rely on general-purpose mappings between text sequences. Utterances are fed in one word (or even one character) at a time, as shown below:



The models are typically broken into **encoder** and **decoder** components. The encoder takes in text (and optionally context information) and produces a high-dimensional *sentence vector*. Based on this vector, the decoder produces a reply. The encoder and decoder rely on thousands of trainable parameters which are optimized via deep learning on a large dataset of sample dialogues. Given sufficient data, these models can learn to generalize responses to unseen inputs.

The Negotiation Task

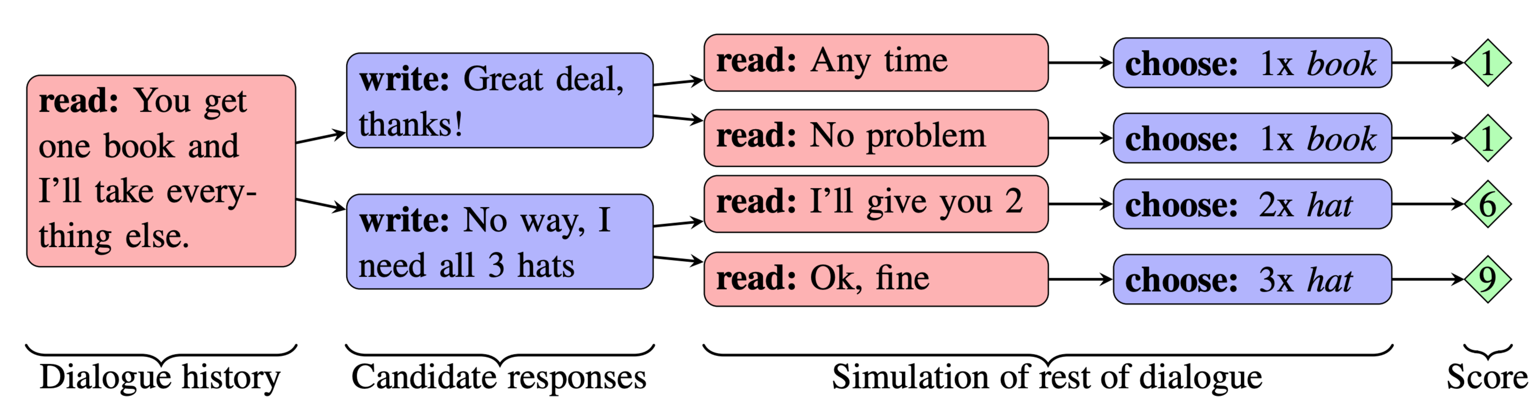
Two agents, each with their own private reward functions, are tasked with dividing a shared set of objects. The agents communicate via text, agree on a deal, and then receive their rewards accordingly. If the agents do not agree on a deal, they each receive zero reward. This task was initially described in Lewis et al. (2017).



Lewis et al. (2017) collected 5808 sample dialogues via Amazon Mechanical Turk which we use as training data.

Existing Model: Dialogue Rollouts

Ideal agents for this task must produce coherent dialogue while simultaneously maximizing their score at the negotiation game. This is not achieved by simple maximum likelihood, so Lewis et al. (2017) and Yarats and Lewis (2017) use a method called dialogue rollouts:



Instead of choosing the most likely response according to the encoder-decoder model, this method probabilistically generates candidate utterances and chooses the one with the highest expected score based on simulations of future dialogue.

Rational Speech Acts (RSA) Framework

The Rational Speech Acts framework (Frank and Goodman, 2012) describes pragmatic language use through layered listeners and speakers who reason recursively about possible utterances and world states.

$$P_{S_0}(u \mid w) \propto \exp(\lambda_1(\log([[u]](w)) - C(u))) \quad (1)$$
$$P_{L_1}(w \mid u) \propto P_{S_0}(u \mid w) \cdot P(w) \quad (2)$$
$$P_{S_2}(u \mid w) \propto \exp(\lambda_2(\log P_{L_1}(w \mid u) - C(u))) \quad (3)$$

The goal of this framework is to preserve the **cooperative principle**. Speakers should be informative and concise given a dialogue context, and listeners rationally assume that speakers follow this behavior.

Pragmatic Dialogue Model

Lewis et al. (2017)'s existing model struggles from computationally expensive dialogue simulations over entire conversations. Yarats and Lewis (2017) provide an extension to this model via a *latent intention dialogue model*. This model clusters utterances into fifty discrete latent variables C_i which are trained to predict future dialogues and actions. We train latent variable prediction models and condition our language generation on these latent variables, as shown below:

$$L_0 = p_c(C_t \mid \text{msg}, \text{hist}) \quad (4)$$
$$S_0 = p_{\text{msg}}(\text{msg} \mid C_t, \text{hist}) \quad (5)$$

where msg denotes an utterance, and hist denotes the dialogue history and context. On top of these, we build pragmatic listeners and speakers:

$$L_1 = \frac{p_{\text{msg}}(\text{msg} \mid C_t, \text{hist})}{z_L} \quad (6)$$
$$S_1 = \frac{p_c(C_t \mid \text{msg}, \text{hist})}{z_S} \quad (7)$$

where z_L and z_S denote normalizations over potential utterances and latent variables. Since these latent variables are meant to signal the speaker's intent, *this pragmatic reasoning ensures that the speaker's intent is retrievable from the produced dialogue.*

Future Directions

The RSA framework requires normalization over both potential alternative utterances and potential world states. This becomes difficult at scale, but we have shown how world states can potentially be bottlenecked through latent representations.

Ongoing work involves studying how to apply RSA at the word level, rather than at the sentence level. This reduces the set of alternatives to a finite and computationally tractable number.

Acknowledgements

Thanks to Chris Potts (Stanford University) for his dedicated advising on this project. Thanks also to Will Monroe (Duolingo) for previous work on this task, and to Johnny Wei for Dynet help. Negotiation task and dialogue rollout figures are from the Lewis et al. (2017) paper.

References

Frank, M. C. and Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.

Lewis, M., Yarats, D., Dauphin, Y., Parikh, D., and Batra, D. (2017). Deal or no deal? end-to-end learning of negotiation dialogues. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2443–2453.

Yarats, D. and Lewis, M. (2017). Hierarchical text generation and planning for strategic dialogue. *arXiv preprint arXiv:1712.05846*.