Iterated Communication Through Negotation

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

The abstract paragraph should be indented ½ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

1 Introduction

One of the first philosophers of language, Ludvig Wittgenstein, posited that "language is use" [2]. This idea, that the use of language is what gives it its meaning, is a profound statement that also has consequences for how we think of language. Wittgenstein saw language as wholly tied to its use, there could be no language separate from reality or possible use. To this end, he defined language games as games with simpler forms of language "consisting of language and the actions into which it is woven".

Recently, the AI community has taken this philosophy of language and sought to use it as the basis for the communication of autonomous agents [1]. The field of "emergent communication" seeks to understand language starting from the most basic of language games; the goal is to teach agents to communicate amongst themselves grounded in a simpler world described by some "game." This game can be one of

2 Related Work

3 Reproduction and Criticism

3.1 Reproduction

The reproduction code is an extension of Hugh Perkins' attempt at reproducing the paper `https://github.com/ASAPPinc/emergent_comms_negotiation` with many notable fixes and extensions. The paper itself is relatively reproducible with strong prosocial agent results especially along the linguistic channel and weaker non-linguistic

3.2 Criticism

Though the paper is quite good, there are still questions and issues that can be raised.

3.3 Prosocial Agents Don't Negotiate

A negotiation game where the interests of two parties do not clash is not as much about negotiation as it is just sharing preferences and calculating the optimum. This is exactly what is seen in case of prosocial agents with only a linguistic channel: one agent communicates their preferences and the

Submitted to 31st Conference on Neural Information Processing Systems (NIPS 2017). Do not distribute.

other agent finds the optimal division and proposes it. In this way, the goal and even the whole task is completely different for prosocial agents compared to selfish agents whose selfish opponents do not have their best interests in mind. In this way, selfish agents have no incentive to communicate their preferences as it could potentially weaken their negotiating ability. In this way, we see that it is not selfish agents failing to learn to communicate but learning to not communicate not because they are selfish but because the situation does not incentivise them to in a one-shot environment.

3.4 Unmotivated Agents

The paper and unfinished repository both allow for sampling situations that give no possible reward to an agent, either through sampling a pool with 0 objects of any kind or sampling a utility for an agent such that there are no object for which they have non-zero utility (e.g. [1,0,1]). In this case, any agent will theoretically accept any opponent's offer but their reward will be $\frac{0}{0}$ and so pose an issue for calculating the fraction of possible reward which the paper uses as a metric.

3.5 Fractional Reward is Biased

The metric used by the paper "fraction of joint reward" compares the reward achieved by each agent (utility over the accepted proposal $u_i \cdot P$) with the total reward possible. In the case of prosocial agents the metric measures the optimality as both agents' rewards are the same (an equally weighted sum of their individual rewards) and they seek to optimize the total possible reward.

$$FR_{prosocial} = \frac{u_A \cdot P + u_B \cdot P}{\max_P u_A \cdot P + u_B \cdot P} \tag{1}$$

But in the case of selfish agents, the total possible reward is the not the objective the agent is optimizing. The agents optimize their own reward and using the total possible reward metric is therefore a bad way of judging performance. Instead, you could measure how well each agent does at maximizing their own reward, average across agents, and divide by the total possible reward

$$FR_{selfish} = 0.5 * \frac{u_A \cdot P}{\max_P u_A \cdot P} + 0.5 * \frac{u_B \cdot P}{\max_P u_B \cdot P} \tag{2}$$

But this is also wrong because the maximum possible reward for one agent will usually decrease the reward for the opponent. In this way, the only situation where $FR_{selfish} = 1$ is possible is if there do not exist any items for which both agents have a non-zero utility. Since this situation is quite rare (the whole point of negotiation is to resolve conflicting preferences) this metric is unfair to selfish agents by usually being lower than for prosocial agents and artificially makes them appear worse. Total possible reward is therefore a bad metric for selfish agents and it is not conclusive that their performance is worse.

4 Exploratory Extensions

4.1 Utility Channel

One way to test whether selfish agents are learning to communicate or whether they are specifically learning not to communicate is to give their opponents full information of their preferences and see whether they achieve better results. To this end, we add the opponent's utility as a separate communication channel.

4.2 Equitable Reward Metric

Since fractional reward is biased as explained in 3.5, a better metric for selfish agents is proposed to measure the performance of their negotiation: equitable reward. Since each agent's preferences may be at odds with the other, it is better to measure the agents' performances taking into account a comparison of their respective utilities. We consider optimality if each agent recieves a proportion of the items in the pool corresponding to the fraction of their utility over their utility and their opponents. Essentially, if each agent recieves items according to how much they care about them.

$$ER_{selfish} = \frac{u_A \cdot P}{\frac{u_A}{u_A+u_B} \cdot P} \tag{3}$$

Since for any given P for selfish agents, $ER(A) = 1 - ER(B)$, we want to just measure the optimality

4.3 Utility Sampling

One consequence of randomly sampling utilities is that there is no guarantee on the clash of utilities in negotation as the players could have non-zero utility only for the items that their opponent has zero utility and negotiation is simplified. To combat this issue, it is proposed to guaranteee non-zero utility to every item.

Another issue is that in the social case, one player's utilities could dominate the other's $u_j^1 > u_j^2 \forall j$. In such a case, the optimal strategy for both players is to give all items to the player with the dominating utility, and again the player. The final split is therefore pareto optimal, but doesn't feel "fair" for the side of the dominated player. A similar situation for a selfish agent would generally lead to a more even split with a smaller total reward. For this reason, we can experiment with avoiding domination situations by normalizing the utilities so that each agents utilities all sum to 15.

5 Iterative Negotation

5.1 Pareto Optimality

6 Conclusion

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments go at the end of the paper. Do not include acknowledgments in the anonymized submission, only in the final paper.

References

References

[1] Kyle Wagner, James A Reggia, Juan Uriagereka, and Gerald S Wilkinson. Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11(1):37–69, 2003.

[2] Ludwig Wittgenstein. *Philosophical investigations*. John Wiley & Sons, 2009.