Multi agent negotiation : Adjusting speed of cooperation

Roopal Vaid(201502023)

Kritika Prakash(20161039)

Abstract

Our approach to the multi-party automated negotiation is to create an improvement of the classic Boulware strategy. We assume that the "Stacking Alternate Offers Protocol" is used. Usually, negotiation problems have constraints such as a complex and unknown opponent's utility in real time, or time discounting. We try to model our agent to use previous bids proposed to us, to become more co-operative with a very hard headed agent. Our agent estimates the alternatives the opponent will offer in the future based on the opponent's offers. Our agent tries to compromise to the estimated optimal agreement point by the end of the negotiation. It goes through 2 phases.

Phase I - In this phase, the agent sticks to a particular bid for a fixed amount of time, each time period decreasing exponentially. This fraction of time in which the agent makes the same bid is decided by the discount factor. Thus, discount factor sets the pace of Phase 1.

Phase II - In phase two agent tries to become more co-operative by taking into account the history of bids it has been offered, from the very beginning. It proposes the bids such that the utility received by itself is average of the previous bids it has been offered.

Algorithm

Phase I

The number of round spent on the first bid is ***timeFactor*** times the number of rounds spent on the second one, and so on, where, timeFactor $\in$ [2,4]. This is decided by the ***discountFactor*** as follows:

$$\boldsymbol{timeFactor} = 2^{\,2-discountFactor}$$

where $discountFactor \in [0,1]$.
Thus, the time spent on the i^{th} generated bid = $total\ no.\ of\ rounds\ /\ timeFactor^{i}$.
The agent stays in this phase until we reach a threshold amount of time spent on a bid.

$$\boldsymbol{timeThreshold} = \sqrt{total\ no.\ of\ rounds}$$

Phase II

The agent tries to become more cooperative in this phase. Our agent estimates the alternatives the opponent will offer in the future based on the opponent's offers so far.
Note that the below metrics are calculated from the very beginning of the negotiation, so as to get a better estimate of the respective opponents' behaviours.
Our behaviour is decided by following set of equations, calculated for each of the opponents:

$$emax(t) = \mu(t) + (1 - \mu(t))d(t)$$
$$target(t) = 1 - (1 - emax(t))t^{\alpha}$$

, where ***emax(t)*** is the estimated maximum utility of a bid proposed in the future by an opponent.
emax(*t*) is calculated by ***μ(t)*** (the mean of the opponent's offers in our utility space), ***d(t)*** (the width of the opponent's offers in our utility space) when the time fraction is t, $t \in [0,1]$ (0 at the beginning of the negotiation, and 1 at the end).
d(t) is the deviation. It is a measure of variance in the utilities offered to our agent by all the bids of a particular opponent. It is related to variance as follows:

$$\sigma^2(t) = \frac{1}{n}\sum_{i=0}^{n} x_i^2 - \mu^2 = \frac{d^2(t)}{12}$$

or

$$d(t) = \sqrt{12}\sigma(t)$$

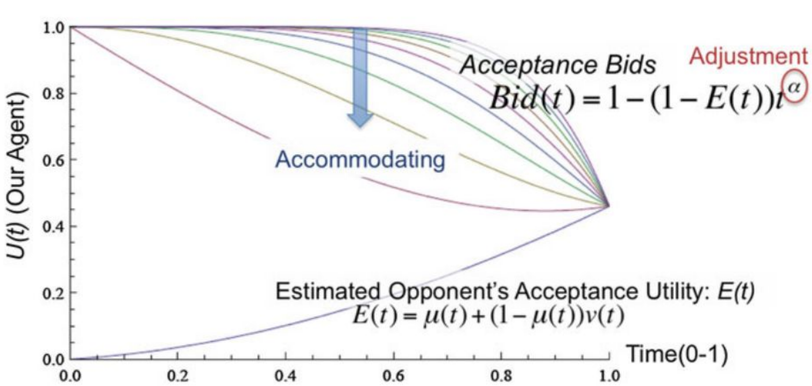
target(t) is an estimate of utility of a bid the agent should generate at time fraction t, (taking into account all the opponents) and ***α*** is a coefficient for adjusting the speed of compromise.

Generated bids approach asymptotically to *emax*(*t*) as the number of negotiation round increases.

Our agent judges whether to accept an opponent's bid based on *target*(*t*) and the mean of the opponent's offers. The following equation defines the ***probability of acceptance*** of the opponent's bid:

$$P = \frac{t^5}{5} + (Offer - emax(t)) + (Offer - target(t))$$

Offer is the utility of the opponent's bid in our utility space.



Once we calculate all these metrics for every opponent at time step t, we calculate the overall values:

overallTarget(t) = f(*target(agent 1,t), target(agent 2,t), … target(agent n, t), reservationValue* ***)***
overallEmax(t) = f(*emax(agent 1,t), emax(agent 2,t), … emax(agent n, t), reservationValue* ***)***

We set ***α*** to 10. To maintain ***continuity*** between Phase 1 and Phase 2, we ensure that the Utility of the last bid generate in Phase 1 is close to the utility generated by the first bid in Phase 2. We constantly ensure that any bid either accepted or generated has a Utility value greater than the ***reservation value***. We also introduce a little bit of ***randomness*** in the bids generated (range of +/- 0.005).

Future Extension

We would consider the dollowing exploration approaches:

1. Learn from the data of past negotiation sessions.
2. Model the opponents more extensively.
3. Explore BOA framework.
4. Test against winners of ANAC 2016.

References

1. "Automated Negotiating Agent with Strategy Adaptation for Multi-times Negotiations" - Katsuhide Fujita