Bounded Rationality of Restricted Turing Machines*

(Extended Abstract)

Lijie Chen
IIIS, Tsinghua University, Beijing, China
wjmzbmr@gmail.com

Pingzhong Tang
IIIS, Tsinghua University, Beijing, China
kenshinping@gmail.com

ABSTRACT

Bounded rationality aims to understand the effects of how limited rationality affects decision-making. The traditional models in game theory and multiagent system research, such as finite automata or unrestricted Turing machine, fall short of capturing how intelligent agents make decision in realistic applications. To address this problem, we model bounded rational agents as restricted Turing machines: restrictions on running time and on storage space. We then study our model in two-person repeated games. In the case where the running time of Turing machines is restricted, we show that computing the best response of a given strategy is much harder than the strategy itself. In the case where the storage space of the Turing machines is restricted, we show the best response of a space restricted strategy can not be implemented by machines within the same size (up to a constant factor). Finally, we study how these restrictions affect the set of Nash equilibria in infinitely repeated games. We show restricting the agent's computational resources will give rise to new Nash equilibria.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Economics, Theory, Algorithms

Keywords

bounded rationality, repeated games, restricted Turing machine

1. INTRODUCTION

In this paper, we studied a realistic model of bounded rationality, where agents are confined to use time-restricted or space-restricted Turing machines to implement their strategies. We first use computational complexity models to rigorously define time and space restrictions. We then study the important game theoretical question of how to compute and implement the best response of a restricted Turing machine, and how such restrictions affect the set of Nash equilibria.

2. PRELIMINARIES

For ease of exposition, we consider the simplest non-trivial case: repeated game [2, 3] where the stage game is the well-known Prisoner's Dilemma(described below). Our results and approach apply to general 2×2 games. In the remainder of this paper, G denotes only the Prisoner's Dilemma.

1,1	0,5
5,0	3,3

We call the two actions of a player *cooperate* and *defect*. Map cooperate to 1 and defect to 0, a strategy is then equivalent to a function: $\{0,1\}^* \to \{0,1\}$.

And a deterministically strategy s can be viewed as a function : $\{0,1\}^* \to \{0,1\}$. Define the language of s by $\{x \mid s(x) = 1\}$. That is, the set of histories that plays cooperate. Then we can define a strategy's complexity by its language's complexity class.

A strategy s is a C-strategy if its language is within complexity class C.

3. TIME-RESTRICTED STRATEGIES

In this section, we study how many time resources are needed to find or implement a best response of a time-restricted strategy.

Here we restrict the running time of a strategy explicitly. Let f be a function and M be a TM, define M_f as a TM such that it runs M on input x of size n for $f(n)$ steps, if it halts, it returns M's output; otherwise it rejects. Then for polynomial f, M_f will always be a P-time strategy.

Then we define the decision problem $\mathrm{BR_f} = \{\langle M, 1^n, k\rangle\}$ such that there exists a strategy that can gain at least utility k against the strategy M_f in the game G^n.

We can prove that with oracle accesses to the decision problem $\mathrm{BR_f}$ and M, we can construct the strategy itself in polynomial time. So it suffices to study the complexity of $\mathrm{BR_f}$.

With a reduction from 3SAT, we can prove the following result:

THEOREM 1. *There exists a polynomial f such that BR_f is NP-complete. And for every polynomial f, BR_f is in NP.*

Theorem 1 suggests that, in order to compute the best response for a general P-strategy, one must be within the class of $\mathrm{P^{NP}}$!

Then it is natural to ask the same question for a $\mathrm{P^{NP}}$-strategy, define $M_{f,O}$ and $\mathrm{BR}_{f,O}$ similarly as above for TM M with oracle

*This work was supported by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the Natural Science Foundation of China Grant 61033001, 61361136003, 61303077, a Tsinghua Initiative Scientific Research Grant and a China Youth 1000-talent program.

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

access to an language O. Since $3SAT$ is NP-*complete*, it suffices to study the complexity of $\mathrm{BR}_{f,3SAT}$.

With a similar reduction from $\sum_2 SAT$, we can prove an analogous result for P^{NP}-strategy:

THEOREM 2. *There exists a polynomial f such that $BR_{f,SAT}$ is $\sum_2^{\mathrm{P}}$-complete. And for every polynomial f, $BR_{f,SAT}$ is in $\sum_2^{\mathrm{P}}$.*

Generalize the above theorem, we have the following corollary.

COROLLARY 1. *There is a polynomial f such that $BR_{f,\sum_i SAT}$ is $\sum_{\mathrm{i+1}}^{\mathrm{P}}$-complete. For all polynomial f, $BR_{f,\sum_i SAT}$ is in $\sum_{\mathrm{i+1}}^{\mathrm{P}}$.*

Recall that $\mathrm{PH} = \bigcup_i \sum_i^{\mathrm{P}}$ for $i \in \mathbb{N}$ and $i \geq 1$.

Since $\mathrm{BR}_{f,\sum_i SAT}$ is the decisional problem of computing best response against strategy in $\mathrm{P}^{\Sigma_{\mathrm{i}}^{\mathrm{P}}}$, which implies finding the best response for strategy in PH is also in PH. In other words, PH is closed up to best response.

4. SPACE RESTRICTED STRATEGY

In this section, we study the space restricted strategies.

However, the strategies defined in the previous section are able to sweep through the history to make a decision, which should be prohibited when the space resources are limited.

So we propose the following alternative models. Our idea is to model space-restricted strategy as a function that maps the last action of the opponent and the current information bits, to the new information bits and the action of this stage.

So a strategy on N information bits is a function $f : \{0,1\}^{N+1} \rightarrow \{0,1\}^{N+1}$. We also define the size(space needed) of a strategy to be the number of storage bits needed to evaluate the function f plus the number of bits needed to specify the function f. Note that as we need bits to store the input to f in order to evaluate it, the information bits are already counted. In the first stage, we assume the information bits and the opponent's action to be 0^N and 0.

Now we move to the implementation details of the function f. We consider two cases.

In the first case, we require f to be efficiently computable. This leads to the *circuit strategy model*. A circuit strategy of N bits is a circuit which has $N+1$ input and output gates.

In the second case, we drop the computation requirement and consider the *inplace strategy model*. An inplace strategy of N bits is a TM which runs on input of $N+1$ bits, always halts, and uses only $N+1$ space, returns the content of tape as output when it halts. For an arbitrary TM M, we define M_I as a TM such that it runs M on input x of size n restricted to n spaces. If M tries to access more than n spaces or doesn't halt after $Qn2^n$ steps, it is forced to halt. M_I returns the content on the tape when it halts. Q is the number of the states in M.

In order to study the complexity of finding the best response against them, we define the corresponding decision problem BRCT $= \{\langle C, n, k\rangle\}$ and BRIP $= \{\langle M, 1^N, n, k\rangle\}$ such that there exists a strategy can yield at least utility k against circuit strategy C(inplace strategy M_I with N information bits) in the game G^n.

We have the following result:

THEOREM 3. BRCT *and* BRIP *are* PSPACE-*complete.*

Then, we study the space complexity for implementing a best response of a particular space-restricted strategy. It is equivalent to ask what is the smallest possible size(recall the size of a space-restricted strategy is the space resources it needs.) of its best responses. From the previous result it is clear that it can be done in polynomial size, then a natural question is whether it can be done in linear space?

We have the following (surprising) theorem which shows it is impossible at least for inplace strategy under reasonable complexity conjecture.

THEOREM 4. *Unless* $\mathrm{DSPACE}(n) = \mathrm{NSPACE}(n)$*, there does not exist a constant T such that any inplace strategy of size S in super game G^n have a best response inplace strategy whose size is smaller than $T \cdot (S + \log n)$.*

5. NASH EQUILIBRIA VIA RESTRICTED TURING MACHINE

In this section, we study the case when both players are restricted TMs and this is a common knowledge. In this setting, it will definitely affect how the game plays, and thus changes the set of possible Nash equilibria.

Our focus of this section is on infinitely repeated game. For simplicity of analysis, the utility notion is the standard limit of mean.

We have the following natural definition for the potential new NEs for a complexity class C:

DEFINITION 1. *A C-NE of an infinitely super game G^* is a pair of strategy (s_1, s_2) such that s_1 and s_2 are C-strategies, and none of them can profitably deviate to another C-strategy.*

Our goal is now to investigate how such restriction affects the set of NE? It is quite obvious this restriction will disqualify some old NEs, but it is surprising that it will also produce some new interesting NEs. Indeed, we have the following two lemmas, one for TM-NE and the other for Polynomial Time-NE (P-NE). This generalize the results from Knoblauch [1].

LEMMA 1. *There exists a TM-NE that is not a NE, and a NE which is not a TM-NE.*

LEMMA 2. *There exists a P-NE that is not a TM-NE, and a TM-NE which is not a P-NE.*

Moreover, we have some stronger results summarized in the following two theorems:

We say a $f : \mathbb{N} \rightarrow \mathbb{N}$ is an increasing unbounded positive function, if f is an increasing function such that $\lim_{n\rightarrow\infty} f(n) = \infty$ and $f(0) > 0$. Let f, g be such two functions in the following two theorems.

THEOREM 5. *If $f(n)\log f(n)$ is $o(g(n))$ and $f(n)$ is $\Omega(n \log n)$. There exists a* $\mathrm{DTIME}(f(n))$*-NE which is not a* $\mathrm{DTIME}(g(n))$*-NE, and a* $\mathrm{DTIME}(g(n))$*-NE which is not a* $\mathrm{DTIME}(f(n))$*-NE.*

THEOREM 6. *If $f(n)$ is $o(g(n))$ and $f(n)$ is $\Omega(\log n)$. There exists a* $\mathrm{DSPACE}(f(n))$*-NE which is not a* $\mathrm{DSPACE}(g(n))$*-NE, and a* $\mathrm{DSPACE}(g(n))$*-NE which is not a* $\mathrm{DSPACE}(f(n))$*-NE.*

REFERENCES

[1] V. Knoblauch. Computable strategies for repeated prisoner' s dilemma. *Games and Economic Behavior*, 7(3):381–389, 1994.

[2] A. Rubinstein. Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, pages 83–96, 1986.

[3] S. Zuo and P. Tang. Optimal machine strategies to commit to in two-person repeated games. In *Proceedings of AAAI*, 2015.