History-based Decision Making in Networks

Chris Bernt and Zhiyong Zhao

September 15, 2016

Abstract

In this report, we consider decison making in networks based on historical strategies and corresponding rewards. We analyze the decision making by separating the information sharing network and interaction network. If there is a distingushed center in the network, the center can use fully bayesian update or following the discounted leader(FTDL). If the information or interaction graphs are general graphs, we use a modified decision rule to simulate the decision making process for all agents. We observe that in two strategy congestion games, mixed-strategy Nash equilibrium is achieved. In a two-strategy coordination game, we consider convergence time to the dominant equilibrium. In particular, convergence time decreases with a higher weight on past actions, but only after a certain threshold. Convergence time also decreases with more past actions considered, but only up to a certain point.

Keywords: information network, interation network, decision-making, following the discounted leader, Bayesian update, Coordination game, Congestion game

1 Introduction

In decision making literature, a common assumption is that people interact with each other in a network and simultaneously share information through the same network. The authors of paper [1] consider these two activities as distinct. In real life, existence of interactions do not guarantee existence of information sharing, and existence of information sharing may not lead to interactions. For example, in viruses spreading processes, people get contaminated even when they do not know the person they were interacting with [2]. In addition, consider someone watching TV who hears about a particular way to resist getting infected. The person likely will not interact with whomever is saying that, but obtains strategy information. In these cases, to understand the behaviours of agents, we need to separate the information and interaction networks.

Each agent in the network wants to maximize his/her expected utility. In [2], they considered the decision making based on two rules: best response and imitate-the-best. The best response rule cares only information sharing, while imitate-the-best depends both on interaction and information. Through random simulations on 100 agents and 500 repetitions, they obtained that the strategies for agents will converge for their chosen coordination game.

In this report, we consider the effect of history data on the decision making for agents who interact in interaction networks and share information in information networks. A special case of our model is that there exists one distinguished center who knows all the information of the other agents. For example, in a class the teacher knows all the history quiz grades of the whole class, while students only know their own grades. The teacher and the students play a teaching-learning game. And in this game the teacher is the distinguished center. If there is a center who knows all the history information of all the agents, we use Bayesian update of decision in the future steps for the center and all the other agents use following the discounted leader(FTDL). Here the discount factor will influence the convergence of decision making and the convergence time. In this special case, we will provide theoretical formulation and simulation results. If there is no such distinguished center, we will use FTDL rule to update decisions for all agents.

The report is organized as follows. Section 2 describes the model and the decision making rule. Section 3 provides the results from simulation for different discount factors etc. Section 4 describes the conclusion and further research.

2 The Model

In this report, we will consider decision-making processes in a model that separates information and interaction networks of agents. The agents only know information of the other agents through information network and they only interact through interaction network. At each step, agents have a set strategy, play a game with all of the agents in their interaction network, and get average of utilities through all of the interactions. For example, the following figure[1] is an example of interaction and information network.

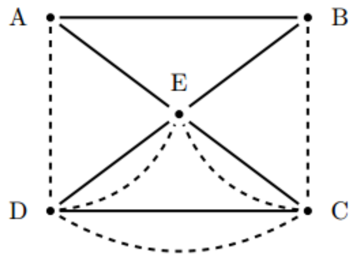


Figure 1: An example network: Interaction (Solid) and Information (Dashed)

In this figure, the interactions are: A-B, A-E, B-E, D-E, C-E, C-D. The information exchanges are: A-D, D-E, D-C, C-E, B-C. In particular, there is no information exchange through A and B, or interaction between A and D.

2.1 A Special Model with A Distinguished Center

Suppose there is a center who knows all the history information of all the agents, but all the other agents only know their own history strategies and rewards. For simplicity, assume there are only two strategies available for each agent. Denote the pure strategy by 0 and 1. We will use the following notations for this report.

- n = the number of agents
- $s_{t,i}$ = the pure strategy of agent i of t steps before the current step
- $r_{t,i}$ = the reward of agent i of t steps for choosing strategy $s_{t,i}$

In the beginning, all agents do not know how to choose their strategies, and we assume they will randomly choose their strategies for a few steps (the exploratory period). After this exploration, we assume agents except the center, follow the discounted leader(FTDL) to decide their next strategies. Denote

$$\delta(s_{t,i}) = \begin{cases} 0 & \text{if agent } i \text{ play strategy 0 in step } t \\ 1 & \text{if agent } i \text{ play strategy 1 in step } t \end{cases}$$

Let the discount factor be γ. Assume we consider $T+1$ steps of history information. The reward of strategy 1 from history is

$$\sum_{t=0}^{T} r_{t,i}\delta(s_{t,i})\gamma^t$$

the reward of strategy 0 from history is

$$\sum_{t=0}^{T} r_{t,i}(1-\delta(s_{t,i}))\gamma^t$$

The agents except the center will choose strategies that maximize their history reward (FTDL).

The center can also use FTDL rule to choose his/her next strategy. If so, then the center will follow the same formula. In this case, all agents will update their strategies according to which strategy has larger results.

Since the center knows all the history data including the history strategies and rewards of all the other agents, the center can use all the history data to update his/her strategy in the future steps. For simplicity, the center thinks that agent i will choose strategy $s_{t,i} \sim$ i.i.d. Bernoulli(θ_i), where $\theta_i = \frac{\sum_{t=0}^{T} s_{t,i}}{T+1}$. Since the center dose not have any other information, the center will have uniform belief of the parameter θ_i.

$$p(\theta_i) = 1 \qquad \text{for } \theta_i \in [0,1]$$

2.2 Posterior Distribution

Now we calculate the posterior distribution of each parameter. We will use Bayesian rule to update the posterior predictive distribution.

The posterior distribution for agent i is

$$\begin{aligned} p(\theta_i|s_{1,i}, s_{2,i}, ..., s_{T,i}) &\propto \prod_{t=0}^{T} p(s_{t,i}|\theta_i)p(\theta_i) \\ &\propto \prod_{t=0}^{T} \theta_i^{\delta(s_{t,i})}(1-\theta_i)^{1-\delta(s_{t,i})} \\ &\propto \theta_i^{\sum_{t=0}^{T}\delta(s_{t,i})}(1-\theta_i)^{T+1-\sum_{t=0}^{T}\delta(s_{t,i})} \\ &= Beta(A_i, B_i) \end{aligned}$$

where $A_i = \sum_{t=0}^{T} \delta(s_{t,i}) + 1$, and $B_i = T - \sum_{t=0}^{T} \delta(s_{t,i}) + 2$.

The posterior predictive distribution for agent i is

$$
\begin{aligned}
& p(s_{new,j}|s_{0,j}, s_{1,j}, ..., s_{T,j}) \\
= & \int_0^1 p(s_{new,j}|\theta_j)p(\theta_j|s_{0,j}, ..., s_{T,j})d\theta_j \\
= & \int_0^1 \theta_j^{\delta(s_{new,j})}(1-\theta_j)^{1-\delta(s_{new,j})}\frac{\Gamma(A_j+B_j)}{\Gamma(A_j)\Gamma(B_j)}\theta_j^{A_j-1}(1-\theta_j)^{B_j-1}d\theta_j \\
= & \frac{\Gamma(A_j+B_j)}{\Gamma(A_j)\Gamma(B_j)}\frac{\Gamma(A_j+\delta(s_{new,j}))\Gamma(B_j+1-\delta(s_{new,j}))}{\Gamma(A_j+B_j+1)}
\end{aligned}
$$

From the Posterior Predictive Distribution, the corresponding probabilities for a new strategy is

- If $s_{new,j} = 1$, $p_j = \frac{A_j}{A_j+B_j}$
- If $s_{new,j} = 0$, $1 - p_j = \frac{B_j}{A_j+B_j}$

From the posterior predictive distribution, the center can predict which strategy he should choose in the next step. If the center chooses 1, the expected average reward is

$$R_1 = \frac{\sum_{j \in N(\text{center})} u(1,1) \times p(s_{new,j} = 1|s_{0,j}, ..., s_{T,j})}{|N(\text{center})|}$$

If the center chooses 0, the expected average reward is

$$R_0 = \frac{\sum_{j \in N(\text{center})} u(0,0) \times p(s_{new,j} = 0|s_{0,j}, ..., s_{T,j})}{|N(\text{center})|}$$

Thus according to this Bayesian update, the center will choose strategy 1 if $R_1 > R_0$. The center will choose strategy 0 if $R_1 < R_0$. If $R_1 = R_0$, the center will be indifferent to 0 and 1.

2.3 General Model

For a general graph, complete Bayesian updating is unfeasible. Each agent will have different beliefs, and different beliefs of what others believe about their opponents' beliefs. Due to this complexity, we will use the discounted rule to update the strategies for each agent.

In each step of iteration, we only allow one agent to update his/her strategy in the next step. Let Agent a_i be the agent who can update his/her strategy in the next step. a_i looks at all her information neighbors $N(a_i)$, and calculates their rewards. She then takes the mean of the rewards for S_0 and S_1, and chooses the higher one as his/her own strategy in the next step.

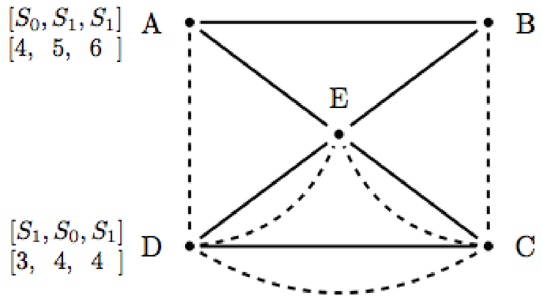


Figure 2: An example network with strategies and outcomes: Interaction (Solid) and Information (Dashed)

We will give a simple example for one update in the situation shown in figure[2]. Suppose Agent A is choose to update his/her strategy in the next step, with $\gamma = 0.9$. The expected utility from strategy S_0 is

$$V(S_0) = \frac{4(\gamma^0) + 4(\gamma^1)}{2} = 3.8$$

The expected utility from strategy S_1 is

$$V(S_1) = \frac{5(\gamma^1) + 6(\gamma^2) + 3(\gamma^0) + 4(\gamma^2)}{4} = 3.9$$

Since $V(S_1) > V(S_0)$, according to the FTDL, A chooses S_1 as his/her strategy in the next step.

For a larger graph, calculating the updates and analyzing the convergence are complicated. In the next section, we will show the results of simulations, from which we can draw interesting conclusions about our model.

3 Numerical Results

In all experiments, except where stated, the maximum step number was 1000, and the number of simulations run was 1000. The final data point for each simulation was the mean of the 1000 individual results. The information and interaction networks for the simulations were independently randomly generated for each simulation according to the simulation parameters. In addition, all agents were in their own information networks, and agents were never in their interaction network. Both graphs were undirected: if an agent i had agent j in their network, then j must also have i in their network.
Frequently referenced simulation parameters include:

- Number of agents n
- The discount factor at which the agents value past actions γ
- The probability of an agent playing 1 for each step in the exploratory phase q_1

- The number of historical actions considered T
- Connection probability of the information network κ_{inf}. When this network is randomly generated, κ_{inf} is the probability that agent j is in the information network of agent i - the same applies for interaction networks.
- Connection probability of the interaction network κ_{int}

In all simulations, agents play in an exploratory phase for T steps, in which they choose a strategy randomly according to q_1.

3.1 Convergence to Equilibrium in Congestion Games

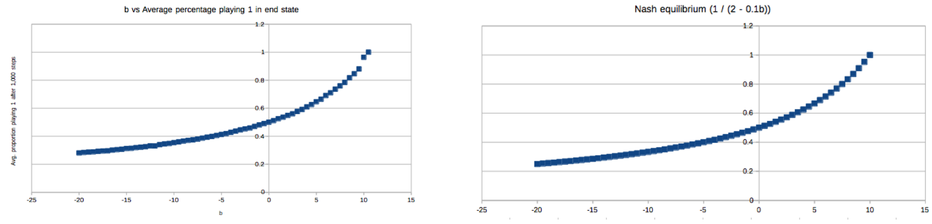
The type of game under consideration in these simulations are congestion games, in which agents are mutually benefited to play different strategies. The specific congestion games we consider are of the form:

	S_0	S_1
S_0	0	a
S_1	a	b

where $a > b$. Note that both agents receive the same values in all considered games, so only one value is shown for simplicity. The pure Nash equilibria in this game are clearly the corners: $\{\ (S_0,\ S_1)\ ,\ (S_1,\ S_0)\ \}$. However, it is very unlikely that all agents in the simulation find themselves in a position to play a pure Nash equilibrium strategy (playing the same strategy for the whole simulation) since their neighbors' strategies may be changing. A mixed strategy Nash equilibrium can be determined for the case when $a = 10$ by calculating the probability p_1 with which a player should play 1 so that the opponent is indifferent between their two options:

$$p_1 a = a - p_1 a + p_1 b$$
$$p_1 = \frac{a}{2a - b}$$
$$p_1 = \frac{1}{2 - 0.1b}$$

This simulation was run with parameters $n = 100$, $\gamma = 0.8$, $q_1 = 0.2$, $\kappa_{inf} = 0.2$, $\kappa_{int} = 0.2$, with $a = 10$ in the game matrix. It was run for 2000 steps: the first 1000 proceeded as normal, to allow the simulation to converge to a stable state. The mean proportion of agents playing 1 was then taken for the latter 1000 steps, to account for random variation. This process was repeated for varying values of b in the interval $[-20, 10]$ with an increment of 0.5. In the figures below, the simulation results are on the left, with the expected number of agents playing 1 based on the Nash equilibrium on the right:



(a) Simulation Results

(b) Nash equilibrium $p_1 = \frac{1}{2-0.1b}$

The two graphs look nearly identical. It is safe to say that in this example with congestion games, our proposed decision rule converges to a mixed strategy Nash equilibrium.

3.2 Convergence to Equilibrium in Coordination Games

We now consider coordination games, in which agents are mutually benefited to play the same strategy. The specific game we use for the following examples is:

	S_0	S_1
S_0	1	0
S_1	0	5

In this game, there are two Nash equilibria, { (S_0, S_0) , (S_1, S_1) }, with (S_1, S_1) strictly better than (S_0, S_0). We therefore expect the simulation to eventually converge to a state in which all agents play 1. The amount of steps it takes to achieve convergence is the main piece of information we are interested in. As a result, parameters were chosen to ensure a high probability of convergence at some point. If too few agents play 1 in the exploratory phase, the simulation will converge to all agents playing 0. $q_1 = 0.2$ was determined in initial trials to be around the minimum needed.

3.2.1 Convergence Time and γ

Here we intend to investigate the relationship between γ and average convergence time. All of the following simulations were run with $T = 10$. Within each simulation set, parameters were kept constant and γ was varied in the interval $[0, 1]$ with an increment of 0.01. The parameters were changed in a way that kept the expected connectivity of the networks constant at 20. The results of one simulation with parameters $n = 100, \kappa_{inf} = \kappa_{int} = 0.2$ are below (See Appendix A for full results).

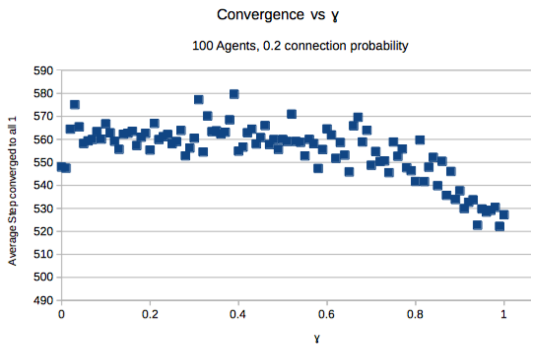


Figure 3: Simulation results for convergence time versus γ

From these results we draw the conclusion that γ has no effect on simulation convergence time, up to a certain γ threshold. Beyond this threshold, increasing γ decreases convergence time. Intuitively, the decline makes sense as the more information is available to agents, the more likelihood there is that an agent observes 1 in their information network enough times to switch to the better strategy.
In addition, the threshold at which γ becomes important changed (See appendix A). In all simulations, the expected connectivity of agents in both their information and interaction networks was 20, achieved by varying n, κ_{inf}, and κ_{int}. The threshold is therefore related to one of these three factors. In further simulations, n was determined to be the most important factor, followed by κ_{inf}, then κ_{int} (See Appendix B).

3.2.2 Convergence Time and T

Here we intend to investigate the relationship between T, the number of historical actions agents consider, and simulation average convergence time. T was varied within the interval $[0, 30]$ with an increment of 1 throughout each simulation (note that $T = 0$ is exactly the decision rule Imitate-the-Best). For all simulations, $n = 100, \kappa_{inf} = \kappa_{int} = 0.2$. The results of one simulation with $\gamma = 0.9$ is shown below (see Appendix C for full results).

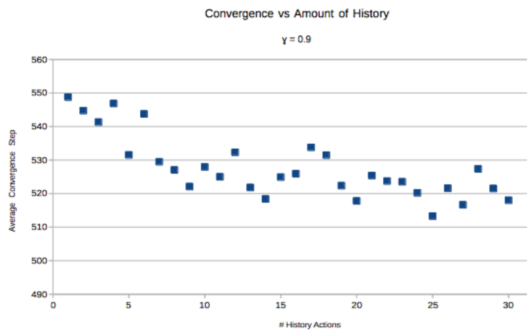


Figure 4: Simulation results for convergence time versus T

These results suggest that a larger T corresponds to lower average convergence time, up to a certain point, at which increasing T does not make a difference. The full results from Appendix C further suggest that the point at which T stops having an influence varies with γ. In particular, that point increases with larger γ, and decreases with smaller γ. This effect has the particular ramification that for this particular parameter set, if $\gamma \leq 0.8$, T has virtually no bearing on convergence time. In addition, higher γ yields less variation in convergence time.

4 Conclusion and Further Directions

In this report we have outlined a model for agents in information and interaction networks to consider their neighbors' past actions and the results of those actions in their own decision making. While full Bayesian updating for all agents is ideal, it is computationally unfeasible. We presented a modified decision rule, based on Imitate-the-Best, and explored how certain simulation and decision rule parameters affect the end state of the simulation, and speed of convergence in coordination games.

Further questions we considered are numerous, and provide rich avenues for further research. The most straightforward variation to our decision rule is to change the way the denominator is calculated. Rather than the total number of times a strategy appears, it can instead be the sum of all γ values used in the numerator. As a result, a strategy will not lose value if it appears far back in the history. In our current decision rule, a strategy will be negatively influenced for appearing far back, since the discount factor will be far too low to offset a 1 added to the denominator.

Other extensions include: agents that are resistant to changing their own behavior; networks comprising of probabilities that two agents will connect in a step; agents broadcasting potentially false information; different classes of network graphs, or directed graphs; and networks that change as the simulation progresses.

References

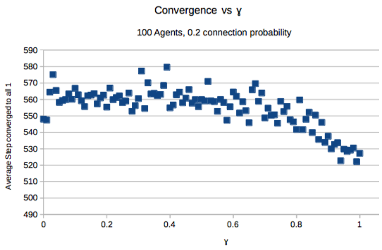
[1] C. Alos-Ferrer and S. Weidenholzer *Contagion and efficiency* Journal of Economics Theory 143 (2008),251-274

[2] S. Angus and V. Masson *The effects of information and interactions on contagion process* Research Paper No. 2010-12 June 2010.

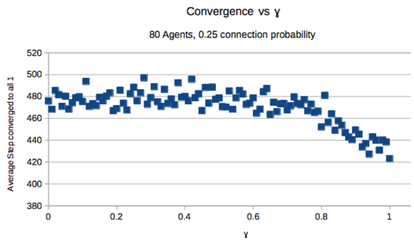
Appendix A Full Results of convergence time and γ

For all of the following simulation sets, all parameters are identical other than those explicitly stated. In addition, the parameters were set so that each agent was expected to have 20 agents in both their information and interaction networks.

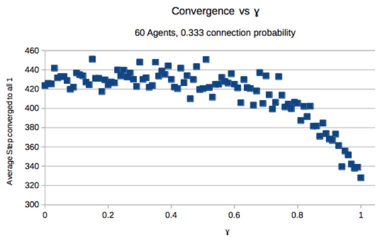
Observe how all graphs follow the same trend: increasing γ has no effect up to a certain threshold, after which increasing γ decreases average convergence time. As can be seen in these graphs of varying parameters, this threshold decreases as n decreases and κ_{inf} and κ_{int} get larger.



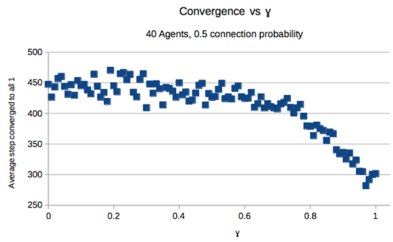
(a) $n = 100, \kappa_{inf} = \kappa_{int} = 0.2$



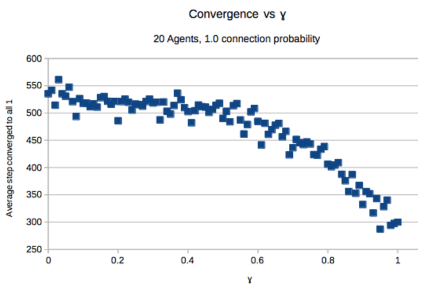
(b) $n = 80, \kappa_{inf} = \kappa_{int} = 0.25$



(c) $n = 60, \kappa_{inf} = \kappa_{int} = 0.333$



(d) $n = 40, \kappa_{inf} = \kappa_{int} = 0.5$

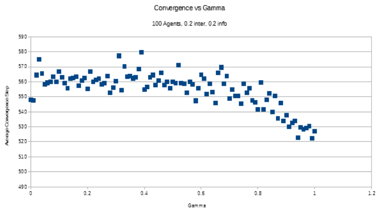


(e) $n = 20, \kappa_{inf} = \kappa_{int} = 1$

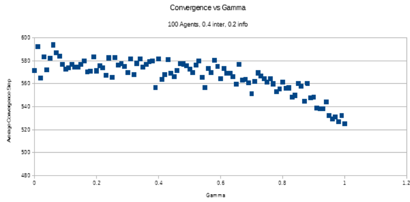
Appendix B Results of convergence time γ threshold

B.1 Effects of κ_{int}

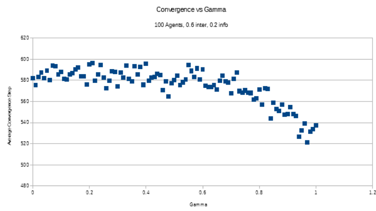
For all the following simulations, $n = 100$ and $\kappa_{inf} = 0.2$. κ_{int} is varied to observe whether it has an impact on the threshold at which γ begins to have an influence on convergence time. Observe how that threshold hardly moves throughout these trials.



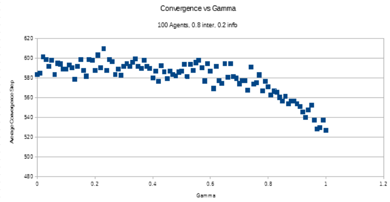
(f) $\kappa_{int} = 0.2$



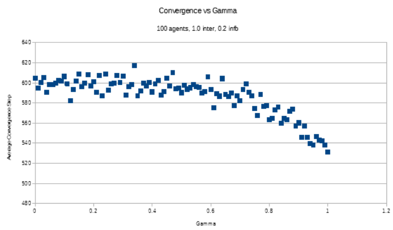
(g) $\kappa_{int} = 0.4$



(h) $\kappa_{int} = 0.6$



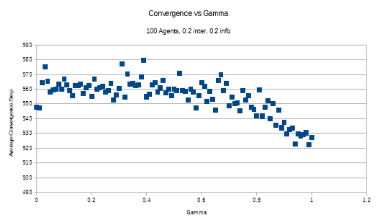
(i) $\kappa_{int} = 0.8$



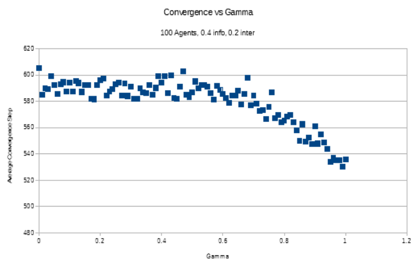
(j) $\kappa_{int} = 1$

B.2 Effects of κ_{inf}

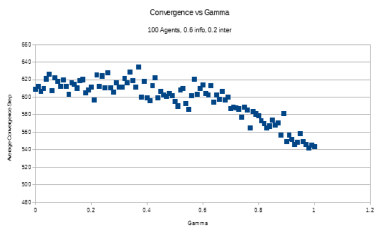
For all the following simulations, $n = 100$ and $\kappa_{int} = 0.2$. κ_{inf} is varied to observe whether it has an impact on the threshold at which γ begins to have an influence on convergence time. Observe how that threshold does decrease with higher κ_{inf} (most easily seen comparing $\kappa_{inf} = 0.2$ and $\kappa_{inf} = 0.8$). The decrease is not as pronounced as when n was changing, however, suggesting that the importance of parameters that determine γ threshold is, from greatest to least, n, κ_{inf}, κ_{int}.



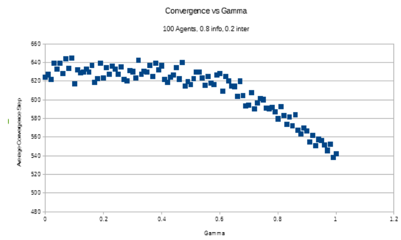
(k) $\kappa_{inf} = 0.2$



(l) $\kappa_{inf} = 0.4$



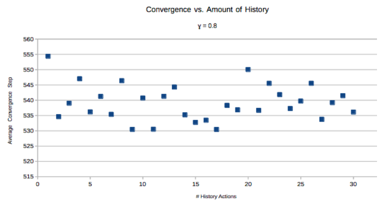
(m) $\kappa_{inf} = 0.6$



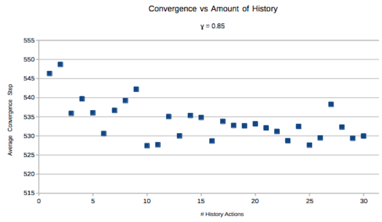
(n) $\kappa_{inf} = 0.8$

Appendix C Full Results of convergence time and T

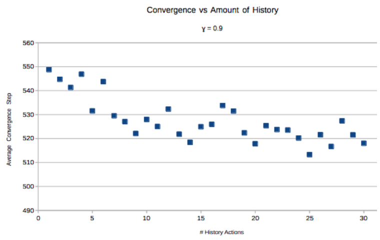
Observe how in most results, there is the trend that increasing T decreases average convergence time, up to a ceratin point after which T makes little difference. With greater γ, this point grows larger. For the parameter set used, this has the ramification that for $\gamma \leq 0.8$, considering history is no different than using Imitate-the-Best.



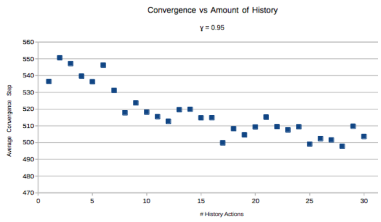
(o) $\gamma = 0.80$



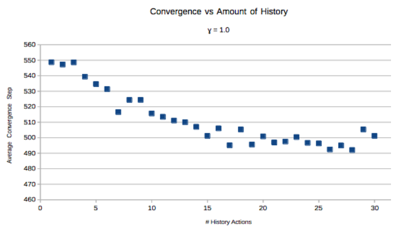
(p) $\gamma = 0.85$



(q) $\gamma = 0.90$



(r) $\gamma = 0.95$



(s) $\gamma = 1.0$