

# Joint Process Games: From Ratings to Wikis \*

Michael Munie  
Computer Science Department  
Stanford University  
munie@stanford.edu

Yoav Shoham  
Computer Science Department  
Stanford University

Microsoft Israel R&D Center  
Herzliya Pituach, Israel  
shoham@stanford.edu

## ABSTRACT

We introduce a game setting called a joint process, where the history of actions determine the state, and the state and agent properties determine the payoff. This setting is a special case of stochastic games and is a natural model for situations with alternating control. Joint process games have applications as diverse as aggregate rating sites and wiki page updates. These games are related to Black's median voter theorem and also strongly connected to Moulin's strategy-proof voting schemes. When each agent has a personal goal, we look at how the play converges under a simple myopic action rule, and prove that not only do these simple dynamics converge, but the actions selected also form a Nash equilibrium. The convergence point is not the mean or the median of the set of agent goals; instead we prove the convergence point is the median of the set of agent goals and a set of focal points. This work provides the first theoretical model of wiki-type behavior and opens the door to more questions about the properties of these games.

## Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Multiagent Systems

## General Terms

Theory, Economics

## Keywords

Game Theory, Wiki, Joint Process

## 1. INTRODUCTION

We propose a new class of games called joint process games. These games are general enough to describe many different situations, from rating systems to voting and wikis. There are four important features of these games. First, agents take turns performing actions in a round robin format. Second, each agent has a personal goal that is not influenced by the goals of other agents. Third, the state of the world is described by a single value, influenced by all

\*This work was supported by NSF grant IIS-0205633-001 and in part by a BSF grant.

**Cite as:** Joint Process Games: From Ratings to Wikis, Michael Munie and Yoav Shoham, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. XXX-XXX.

Copyright © 2010, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

agent's actions, so each agent's payoff is the distance from the outcome in the round to the agent's goal. Finally, this is an infinite game, so utility is determined by the limit average of rewards in each round.

To get a feel for joint process games, let's look at two different examples. A very simple example that joint process games model well is ratings on a Trip Advisor-like site that accepts many reviews from users but only displays one aggregate review. Imagine a group of travelers, each with their own opinion of how a particular hotel should be rated. The travelers submit ratings to the site in a round robin format. The ratings can only be positive or negative. If these travelers behave by attempting to make the aggregate rating as close to their own rating at each time step, what will the properties of the joint behavior be? What will the players' payoffs be? Will this process reach a steady state? We answer these questions later, but first we look at a second example of a joint process.

From Wikipedia's founding in 2001, it has grown to almost thirteen million articles. However, Wikipedia, and the class of wikis in general, still lack a crisp theoretical model of article evolution. While a wiki can be modeled in many ways, we look at arguably the most simple model of all: an article consisting of a single bit. A wiki is the history of updates to this bit, and the outcome at time  $t$  is the historical average of the article's value. We have a set of agents that update the wiki in a round robin fashion. Each agent has a personal goal between zero and one, not drawn from any common distribution, but exogenously given. While we admit our model is basic and lacks some of the higher order interactions involved in a wiki's development, we feel it captures the intuition behind the evolution of articles. The same questions we had about the rating site apply here: will the value of the wiki converge, and to what point?

We will shortly discuss equilibrium, but we begin by considering a natural myopic play rule similar to fictitious play or best response dynamics.

Myopic play in wiki games has the important property that the outcome and the empirical distribution of each player's actions converge. The outcome converges to the median of the set of user beliefs unioned with a set of partition points – more on this in the next section. We show a formal connection between this convergence point and the mean of agent goals. Later, we also characterize all payoffs obtainable in a Nash equilibrium of a joint process game and show that the myopic strategies in fact form a Nash equilibrium themselves. These results give us insight into what behavior suffices to produce near optimal results in a joint process, why joint processes work well, and when they might not work as well.

The rest of the paper is organized as follows. We cover the general model for joint processes and a specialized model for wikis in

Section 2. In Section 3 we prove convergence results, some characterizations of the convergence point, and a folk theorem. In Section 4 we discuss related work. Finally in Section 5 we conclude and discuss future work.

## 2. FORMAL MODEL

First, we define a joint process.

DEFINITION 1 (JOINT PROCESS). *A joint process consists of*

- a finite set  $N$  (the set of  $n$  agents)
- for each time  $t \in \mathbb{N}$ , a vector  $g^{(t)} = [g_1^{(t)}, \dots, g_n^{(t)}]$ , where  $\forall j \in N, g_j^{(t)} \in [0, 1]$  (the agents' goals)
- for each agent  $i$ , a set of possible actions  $A_i$
- a function  $v : \mathbb{N} \mapsto N$  that determines the active agent at time  $t$
- for each time  $t \in \mathbb{N}$ , the action taken by the active agent  $a^{(t)} \in A_{v(t)}$  (we will refer to the history of actions as  $A^t = [a^{(1)}, a^{(2)}, \dots, a^{(t)}]$ )
- a function  $f : A^t \mapsto [0, 1] = o^{(t)}$  (the outcome at time  $t$ )
- a distance function  $d : [0, 1] \times [0, 1] \mapsto [0, 1]$
- for each time  $t \in \mathbb{N}$ , a vector  $r^{(t)} = [r_1^{(t)}, \dots, r_n^{(t)}]$ , where  $\forall j \in N, r_j^{(t)} = 1 - d(g_j^{(t)}, o^{(t)})$  (the reward of each player)

At every time step only one agent will take an action (no simultaneous actions) and we assume that agents take their actions in a fixed order, in round robin fashion. The state of the world is always observable by every agent.

The reward at each time step is a function of the agent's goal and outcome in that time step. From now on we will assume the use of the Euclidean distance function  $d(\cdot)$ . At time  $t$ , agent  $j$ 's reward is denoted as  $r_j^{(t)} \in [0, 1]$ . The distance function  $d$  implies that an agent's goal at time  $t$  is the outcome for which the agent receives his greatest payoff.

This formal model can be seen a special case of stochastic games [21], although with an infinite number of states and a deterministic state transition function. Other work has also considered turn-taking games [2], but without our outcome and reward structure. We will return to discuss related work in more depth after covering our main results. As in stochastic games, we use a method common in the literature of defining a player's payoff in an infinitely game.

DEFINITION 2 (AVERAGE REWARD). *Given an infinite sequence of payoffs  $r_i^{(1)}, r_i^{(2)}, \dots$  for player  $i$ , the average reward of player  $i$  is*

$$\limsup_{k \rightarrow \infty} \frac{\sum_{j=1}^k r_i^{(j)}}{k}.$$

We have given the formal model, but left unspecified the action space. As we define our specific model in the next sections, this will be filled in. However, the features of the formal model capture the intuition behind joint process games. Agents are involved in a repeated interaction where they only care about the distance from the current global outcome to their own goal. The average reward is influenced by changes in strategy and agent's past actions continue

to influence future rewards. In the next section, we will discuss a refinement of this model.

Returning to our wiki example from the introduction we ask: how can we use a joint process to model control over a wiki? First, we specify the set of possible actions to be  $\{0, 1\}$ , corresponding to a single bit update. Second, the value of the article is the historical average of the past bits. Finally, the goals of agents are constant over time.

Formally, we specialize the class of joint process games as follows. Let the set of possible actions for each player at every time step be equal to  $\{0, 1\}$ . Let the outcome be the historical average of past actions,  $f(A^{(t)}) = \frac{\sum_{i=1}^t a^{(i)}}{t} = o^{(t)}$ .

In this setting agents have goals that do not change over time, which we enforce to be unique. Thus for all players  $j$  and all times  $t$  and  $t'$ ,  $g_j^{(t)} = g_j^{(t')}$ . We will refer to  $g_j^{(t)}$  simply as  $g_j$ .

Although this special case of the joint process model we have just defined can be used to describe many different situations, for the rest of the paper we will refer to it as a wiki game or the wiki model.

## 3. ANALYZING WIKI GAMES

With the model specified we still ask: what dynamics occur when this game is actually played? Looking for Nash equilibria is one possible option, but instead we start by looking at natural strategies. We propose modeling the behavior under myopic play, where agents choose the action giving the highest payoff at the current time, and ignore its effect on future rewards. Later in the paper we reduce the space of possible strategies by looking at Nash equilibrium, but we prove even later that there are many such equilibrium points. This model can just as easily model the process of submitting ratings to a review site.

We find that the outcome converges to an interesting value and almost all of the player's actions become stationary after some finite time. In the next subsection we define myopic play, then we prove convergence, and finally we discuss some fascinating connections to voting and their implications.

### 3.1 Myopic Play

Myopic play for wiki games is defined as follows.

DEFINITION 3 (MYOPIC PLAY). *At time  $t$ , the agent that is active, agent  $i$ , chooses an action  $a \in \{0, 1\}$  such that*

$$a = \operatorname{argmin} \left( d \left( g_i, \frac{a + \sum_{i=1}^{t-1} a^{(i)}}{t} \right) \right)$$

The motivation behind this myopic play rule is that agents are able to easily optimize their current period rewards, and it is natural for them to select actions for this immediate payoff.

Just as we stated in Definition 3, at each time step, an agent will have to decide between the action 0 or 1. It is useful both conceptually and in our later proofs to look the pair of points resulting from the agent's possible actions. We call this range the active window.

DEFINITION 4 (ACTIVE WINDOW). *The active window is the interval between the outcome when the active agent chooses action 0 and the outcome with action 1. This is written as,*

$$\left[ \frac{\sum_{i=1}^{t-1} a^{(i)}}{t}, \frac{1 + \sum_{i=1}^{t-1} a^{(i)}}{t} \right] \quad (1)$$

The active window is a useful tool in our following proofs, and allows an easy way to visualize the myopic decision process.

### 3.2 Convergence Under Myopic Play

Under the simple action choice rule of myopic play, the outcome converges to a value that we will refer to as the *central value*. The central value has connections to the median and mean of the set of agent goals, a connection that we will make more exact later in this section.

**DEFINITION 5 (CENTRAL VALUE).** *Given a set of agents  $N$ , let  $G$  be the set of agent goals,  $G = \{g_i | i \in \{1, 2, \dots, n\}\}$ . Let  $P$  be the set of partition values, where  $P = \{\frac{x}{n} | x \in \{1, 2, \dots, n-1\}\}$ . The central value,  $C(N)$ , is the median of the multiset  $G \cup P$ .*

The following lemma clarifies how the central value divides the set of agents' goals. This lemma is important to later proofs in the paper and connects the partition points with the agent goals.

**LEMMA 1.** *Let  $x$  be the least integer that satisfies  $C(N) \leq \frac{x}{n}$ . Then  $|\{i \in N | g_i < C(N)\}| = n - x$  and  $|\{i \in N | g_i \geq C(N)\}| = x$ .*

**PROOF.** There are  $x - 1$  partition points less than  $\frac{x}{n}$  and consequently  $n - x$  partition points greater than or equal to  $\frac{x}{n}$ . Let  $l$  represent the number of goals less than  $C(N)$ . If  $C(N)$  is a goal, by the definition of the median and definition 5 we have,

$$\begin{aligned} x - 1 + l &= n - x + (n - l - 1) \\ 2l &= 2n - 2x \\ l &= n - x, \end{aligned}$$

and consequently the number of goals greater than or equal to  $C(N)$  is  $x$ .

If  $C(N)$  is a partition point,

$$\begin{aligned} x - 1 + l &= n - x - 1 + (n - l) \\ 2l &= 2n - 2x \\ l &= n - x, \end{aligned}$$

reaching the same conclusion.  $\square$

The next results we cover are necessary to prove convergence and concern the behavior of the active window over time. Not only does the size of the active window decrease over time, but its position stabilizes.

**LEMMA 2.** *There exists a  $T_0$  such that for all  $t > T_0$ , the active window contains at most one  $g_i$ .*

**PROOF.** First, we note that the length of the active window is  $\frac{1}{t}$ . Given the set of agent goals  $G$ , where  $d$  denotes the Euclidean distance metric, define the distance between the closest pair of goals as,

$$\rho = \min_{i,j \in G, i \neq j} d(i,j) > 0. \quad (2)$$

Let  $T_0 = \lceil \frac{1}{\rho} \rceil$  and we can see that  $\forall t > T_0, \frac{1}{t} < \rho$ , so there can be at most one value in the active window because the length of the interval is small enough to prevent any other points from simultaneous being included.  $\square$

The definition of myopic play gives no explicit guarantee on the long term behavior of the agents. But, by the nature of myopic play, agents will respond in a predictable way according to the relative position of their goal and the active window. Because the length of the active window will eventually cover at most one agent's goal, it follows in the next lemma that the actions of all agents but at most one are determined by what side of the active window their goal is on.

**LEMMA 3.** *For all times  $t$  and all agents  $i$  such that  $g_i \leq \frac{\sum_{i=1}^{t-1} a^{(i)}}{t}$ ,  $a^{(t)} = 0$ . For all times  $t$  and all agents  $i$  such that  $g_i \geq \frac{1 + \sum_{i=1}^{t-1} a^{(i)}}{t}$ ,  $a^{(t)} = 1$ .*

**PROOF.** In order to minimize the distance from  $g_i$  to either endpoint of the active window, if  $g_i$  is less than the lower bound of the active window, the distance will be minimized by agent  $i$  setting  $a^{(t)} = 0$ . Likewise, if agent  $i$  has  $g_i$  greater than the upper bound of the active window, the distance will be minimized by setting  $a^{(t)} = 1$ .  $\square$

Under the assumption that agents use myopic play, the active window will cover  $C(N)$  infinitely often. The bounds of the active window will also approach  $C(N)$  as time goes to infinity. Because the size of the active window goes to zero as time goes to infinity, the outcome of the wiki model will converge to  $C(N)$  and the limit average payoff of all agents will be equal to the distance from their goal to  $C(N)$ . However, as seen in the proof of the following theorem, even though all agents with goals other than the central value will have their play converge, if an agent's goal is the central value, then the play of that agent will not converge even though his average reward will converge to one.

**THEOREM 1.** *For all  $\epsilon > 0$  there exists a time  $T_1$  such that for all  $t > T_1$ ,  $\left| C(N) - \frac{\sum_{i=1}^t a^{(i)}}{t} \right| < \epsilon$*

Although the proof of this theorem is complicated, it is roughly composed of three subclaims. First, the active window will eventually cover  $C(N)$ . Second, if the active window covers  $C(N)$  at time  $t$ , it will again cover  $C(N)$  at some future time. Third, the distance from the active window's endpoints to  $C(N)$  will decrease.

**PROOF.** If the active window is less than  $C(N)$  at time  $t > T_0$ , the first time in which the agent (if there is one) whose goal is  $C(N)$  is acting, then by the definition of  $x$  in lemma 1, the active window is less than  $x/n$ . By lemma 1 there are  $x$  agents with  $g_i \geq C(N)$ , which is greater than the active window. Since there are at least  $x$  agents whose action will be 1, Then at time  $t + kn$ , The right bound of the active window will be  $\geq \frac{kx + \sum_{i=1}^t a^{(i)}}{t + kn}$ . The limit as  $k$  goes to infinity is  $x/n$  so there exists some time in which this quantity will be closer to  $x/n$  than  $C(N)$  is, and the right bound of the active window will pass  $C(N)$ . Let this time be denoted as  $T_2$ . A similar argument holds if the active window is greater than or equal to  $C(N)$ .

If  $C(N)$  is a partition point, then we're done because once the active window is between two agent's goals, and no agent will ever change their action from this point on under the myopic play rule causing the active window to converge to  $x/n$ .

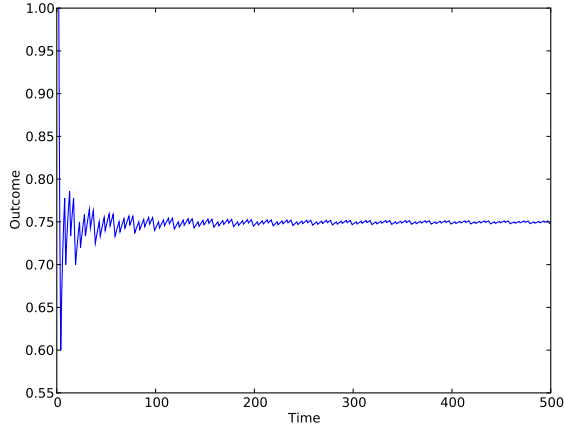
Denote the active window at time  $t$  as  $[\frac{z}{t}, \frac{z+1}{t}]$ . If at time  $t$  the active window is covering  $C(N)$  and without loss of generality the active agent's action is 0, then at time  $t + yn$  where  $y = \lceil \frac{t}{2xt - 2nz - n} \rceil$  the active window will again cover  $C(N)$ . Starting from our assignment of  $y$ ,

$$y(xt - nz - .5n) \geq \frac{t}{2}, \quad (3)$$

$$yxt - ynz - .5yn \geq \frac{t}{2}, \quad (4)$$

$$\frac{z + yx}{t + yn} \geq \frac{z + .5}{t}, \quad (5)$$

thus the right bound of the active window is greater than  $\frac{z + .5}{t}$ . Given that the agent's action is 0, under the myopic play rule we



**Figure 1: Convergence for the set of agents with goals  $G = \{.95, .90, .85, .75, .65\}$  using myopic play.**

know that  $C(N) \leq \frac{z+0.5}{t}$ . This inequality and the inequality deduced from the value of  $y$ , give us that  $C(N) \leq \frac{z+yx}{t+yn}$  and  $C(N)$  is covered at time  $t+yn$ . An analogous argument holds if the agent's action is 1.

We also note that for every  $1 \leq y' < y$ , the chosen endpoint of the active window will be closer to  $C(N)$  than at time  $t$ . Assume the agent's action is zero and thus has chosen  $z/t$  under the myopic rule as the closest point, then  $C(N) < \frac{z+.5}{t}$ . If at time  $y'$  the active window is not covering  $C(N)$  we know that  $\frac{z+y'x}{t+y'n} < C(N)$ , and because  $z/t < C(N)$  and  $C(N) \leq x/n$ , then

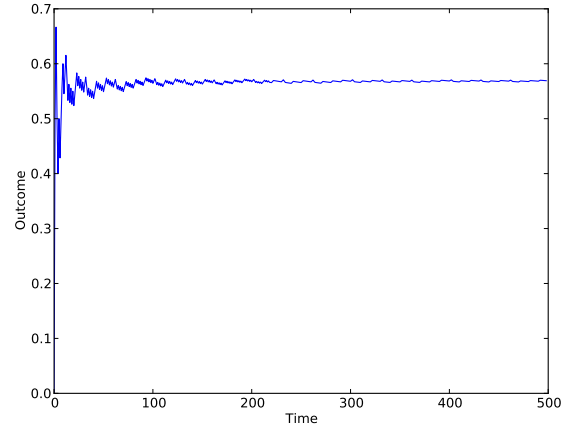
$$\begin{aligned} \frac{z}{t} &\leq \frac{x}{n}, \\ zt + y'nz &\leq zt + y'tx, \\ \frac{z}{t} &\leq \frac{z+y'x}{t+y'n}, \end{aligned}$$

showing that at time  $t+y'n$  the active window will be closer than at time  $t$ .

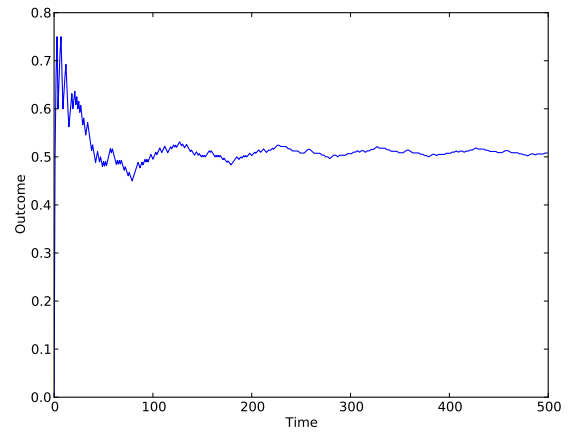
Thus, if we let  $T_1$  be the maximum of  $T_2$  and the first time greater than  $\lceil \frac{2n-1}{2\epsilon} \rceil$  such that the active window is again covering  $C(n)$ , (which we have already proved will happen) then at time  $T_1$  the active window is covering  $C(N)$ . The length of the active window is  $\frac{1}{T_1} < \epsilon$ , and the maximum distance  $C(N)$  can be from the active window is  $\frac{1}{2T_1}$ . As we showed before, until the active window again covers  $C(N)$  the distance to  $C(N)$  will only be less than at time  $T_1$ . Finally, once the active window again covers  $C(N)$  the maximum distance will be  $\frac{1}{2T_1'} < \frac{1}{2T_1}$ .  $\square$

We show the value of the outcome over time for an example set of five agents in Figure 1. Although the convergence is far from monotonic, it still converges with some regularity. For this set of agents, the reader can check that  $C(N) = .75$ . In Figures 2,3, and 4 we show the behavior for larger sets of agents with goals drawn from a uniform distribution over  $[0, 1]$ .

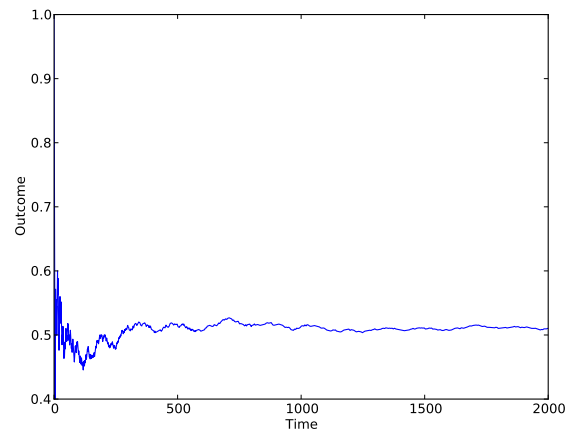
The convergence under the simple and easy to compute myopic play rule is remarkable, but it is still unclear what the meaning of the central value is, given that it is not the mean or median of the set of agent goals. In the following lemma we clarify the connection of the central value with the mean of agent goals. Generally, the



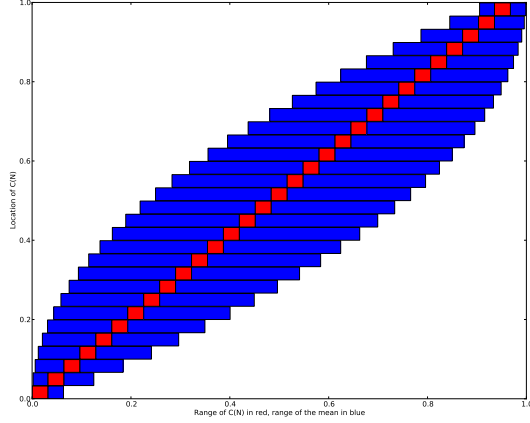
**Figure 2: Convergence for a set of 10 random agents**



**Figure 3: Convergence for a set of 100 random agents**



**Figure 4: Convergence for a set of 1000 random agents**



**Figure 5: Possible mean values of the set of agent goals given  $C(N)$ , with  $|N| = 30$**

central value is closest to the mean of agent goals when the goals are close to 0 or 1, and farthest when the central agent's goal is in the middle between 0 and 1. Specifically, the following lemma allows for the greatest distance between the central value and the mean when  $x = \lceil \frac{n}{2} \rceil$  and the greatest range for possible values of the mean when  $x = \lfloor \frac{2n+1}{4} \rfloor$ , as derived by basic calculus.

**LEMMA 4.** *Let the central value for a set of agents be  $C(N)$  such that  $C(N) \in [\frac{x-1}{n}, \frac{x}{n}]$ . Then the mean of agent goals lies in  $[\frac{x^2-x}{n^2}, \frac{2xn-x^2}{n^2}]$ .*

**PROOF.** First we prove the lower bound. By Lemma 1 there are  $x$  agent goals greater than  $C(N)$ . The least value these goals can take is  $\frac{x-1}{n}$  and the  $n-x$  goals less than  $C(N)$  can all be zero giving us a mean of  $\frac{0(n-x) + x \frac{x-1}{n}}{n} = \frac{x^2-x}{n^2}$ .

Likewise, for the upper bound, the greatest value the  $n-x$  agents less than  $C(N)$  can take on is  $\frac{x}{n}$  and the other agents can take on the value of 1. The mean in this case is  $\frac{1(x) + (n-x) \frac{x}{n}}{n} = \frac{2xn-x^2}{n^2}$   $\square$

The above lemma makes clear that although there is a connection between the central value and the mean of agent goals, these two values can possibly be quite far apart (up to  $1/4$  of the interval distance) from each other. We graph an example of the bounds where we have thirty agents in Figure 5. While joint process games have many desirable properties, close convergence to the mean is not guaranteed. But in situations where convergence to the mean is not the goal, the central value provides an alternative that balances the goals of extreme agents with the set of social partition points.

### 3.2.1 Continuous Action Space

We also note that all the previous lemmas up to this point directly apply to the case where we have an action space of  $[0, 1]$ . The active window can now be interpreted as the range of next round outcomes the active agent can achieve. Theorem 1 can also be extended to cover the continuous action space though a slight modification of the proof of the subclaim that once the active window covers  $C(N)$ , it will again cover it.

## 3.3 Myopic Play and Nash Equilibrium

Not only under myopic play does the outcome converge but the induced actions also form a stable point.

**THEOREM 2.** *The set of strategies defined by the myopic play rule form a Nash equilibrium.*

**PROOF.** Let  $A = (a_1, a_2, \dots, a_n)$  be the set of strategies defined by the myopic play rule for all possible action histories, including histories that are not reached when the set of agents follow strategy profile  $A$ . Consider any  $a'_i$  that agent  $i$  is considering. If agent  $i$  has  $g_i = C(N)$  then it follows that under myopic play this agent receives an average reward of 1, the maximum possible, so there is no profitable deviation.

Without loss of generality, assume that agent  $i$  is such that  $g_i < C(N)$ . Thus, by Lemma 3, and Theorem 1 there exists a  $T_0$  such that for all  $t > T_0$ , agent  $i$ 's action under the myopic play rule is 0. Let the average reward for player  $i$  when the strategy profile is  $A$  be denoted  $r_i(A)$ . If for all times  $t$  except when agent  $i$  is acting, strategy profile  $A' = (a_1, a_2, \dots, a'_i, \dots, a_n)$  causes players to take the same actions as under  $A$ , it then follows that,

$$r_i(A) = \lim_{k \rightarrow \infty} \frac{\sum_{j=1}^k r_i^{(j)}}{k}, \quad (6)$$

$$= \lim_{k \rightarrow \infty} \frac{\sum_{j=T_0}^k r_i^{(j)}}{k}, \quad (7)$$

$$= \lim_{k \rightarrow \infty} 1 - \frac{\sum_{j=T_0}^k o_i - g_i}{k}, \quad (8)$$

$$r_i(A) - r_i(A') = \lim_{k \rightarrow \infty} \frac{\sum_{j=T_0}^k o'_i - o_i}{k}, \quad (9)$$

$$(10)$$

and finally, because after  $T_0$  agent  $i$ 's action is always 0, then any  $a' \neq a$  will have some number of ones. Thus,  $o_i < o'_i$ , giving us that  $r_i(A) \geq r_i(A')$ .

Now suppose that  $a'$  causes an agent other than agent  $i$  to play a different action in some round. Given that the length of the active window approaches zero and that the agents use the myopic play rule, all agents with  $g_i$  less than the active window will play zero and all agents with  $g_i$  above the active window will play one. For agent  $i$  to get a higher reward under action  $a'$  the active window would have to be less than  $C(N)$ . All agents currently less than  $C(N)$  can not cause the active window to move closer to zero since they all have converged to playing all zeros. No agent  $\geq C(N)$  will ever cause the active window to decrease below  $C(N)$ , as it would lower their reward in the round they choose zero, violating myopic play. Thus, agent  $i$  can never cause the outcome to decrease and can never increase his utility by deviating. Because there is no profitable deviation in all cases  $A$  forms a Nash equilibrium.  $\square$

## 3.4 A Folk Theorem

In the previous section we saw one Nash equilibrium, namely that defined by the myopic strategies. Can we somehow characterize *all* the Nash equilibria in these games? While we don't quite do this here, we do present a folk theorem which captures the payoffs in all such equilibria.

Roughly speaking, the payoffs obtainable by player  $i$  are any rational values between his minimax value and 1. Let

$$v_i = 1 - \max(g_i - \frac{1}{n}, \frac{n-1}{n} - g_i),$$

where  $v_i$  is the reward player  $i$  gets when all other players play minimax strategies against him, and player  $i$  plays his best response. In this joint process games, this will take the form of all players playing 0 or 1 except player  $i$  who plays the opposite.

The following two definitions are used to characterize the payoffs obtainable.

DEFINITION 6 (ENFORCEABLE). A payoff profile  $r = (r_1, r_2, \dots, r_n)$  is enforceable if  $\forall i \in N, r_i \geq v_i$ .

DEFINITION 7 (FEASIBLE). A payoff profile  $r = (r_1, r_2, \dots, r_n)$  is feasible if there exist rational, nonnegative values  $\alpha_a$  such that for all  $i$ , we can express<sup>1</sup>  $r_i$  as  $\sum_{a \in \{0,1\}^n} \alpha_a |g_i - f(a)|$  with  $\sum_{a \in \{0,1\}^n} \alpha_a = 1$ .

Thus, a payoff profile is feasible if it is a convex and rational combination of the possible one round (where a round is  $n$  time steps) outcomes in our joint process game.

THEOREM 3 (FOLK THEOREM). If the payoff profile of a joint process game is  $r = (r_1, r_2, \dots, r_n)$  then:

1. If  $r$  is the payoff profile for any Nash equilibrium  $s$ , then for player  $i$   $r_i$  is enforceable.
2. If  $r$  is feasible and enforceable, then  $r$  is the payoff profile for some Nash equilibrium

The proof follows the form of other folk theorems, similar to that in [22]. To prove item one we show that an agent can never receive less than  $1 - \max(g_i - \frac{1}{n}, \frac{n-1}{n} - g_i)$ , in any equilibrium. The proof of item two gives a construction of strategies that yields a Nash equilibrium with payoff profile  $r$ .

PROOF. *Part 1:* Suppose  $r$  is not enforceable. Then there exists some  $i$  such that  $r_i < v_i = 1 - \max(g_i - \frac{1}{n}, \frac{n-1}{n} - g_i)$ . However, consider an alternative strategy for  $i$ : let  $i$  follow the myopic play rule. The outcome can only be a distance more than  $\max(g_i - \frac{1}{n}, \frac{n-1}{n} - g_i)$  for a discrete number of time steps. The greatest value that  $\max(g_i - \frac{1}{n}, \frac{n-1}{n} - g_i)$  can take on is  $\frac{n-1}{n}$  and because agent  $i$  performs an action every  $n$  rounds, by following the myopic play rule the distance will never be greater than this. Thus if  $r_i < v_i$  then  $s$  can not be a Nash equilibrium.

*Part 2:* Because  $r$  is a feasible, enforceable payoff profile, we can write it as  $r_i = \sum_{a \in \{0,1\}^n} \frac{z_a}{y} |g_i - f(a)|$  because  $\alpha_a$  is rational and  $y$  is the least common denominator of all the  $\alpha_a$ . Because the combination is convex, we have  $y = \sum_{a \in \{0,1\}^n} z_a$ .

Let  $(a^t)$  be a sequence binary values that cycles through each  $a \in \{0,1\}^n$  (a binary sequence of player actions with length  $n$ ) with a total cycle length  $y$ , where each  $a$  is repeated exactly  $z_a$  times. Let us define a strategy  $s_i$  of player  $i$  to be the trigger version of playing  $(a^t)$ . If all other players play the action dictated by  $a^t$  in their respective time times, player  $i$  will play the action dictated to him as well. However, if at any time some player  $j$  deviates from the sequence  $(a^t)$ , then if  $g_j < .5$  all other agents will uniformly play 1 from that time onward. Likewise, if  $g_j \geq .5$  all other players will play 0.

If everyone plays according to our dictated strategies  $s_i$ , then the payoff profile will simply be  $r$  by our construction of  $(a^t)$ . However, if any player deviates, then they will receive a maximum payoff of  $1 - \max(g_i - \frac{1}{n}, \frac{n-1}{n} - g_i)$  for every subsequent cycle, giving them a limit average reward of at most  $v_i$ . Since  $r$  is enforceable,  $v_i \leq r_i$  and there is no profitable deviation.  $\square$

Although the strategies produced by myopic play form a Nash equilibrium, the folk theorem show that there are many more possible equilibria. Nevertheless, this characterization gives us more insight into joint process games.

<sup>1</sup> $\{0,1\}^n$  represents the set of possible stationary strategies for  $n$  time steps of our infinite extensive form game

## 4. RELATED WORK

The related work can be divided into two relevant threads: game theory and models created for specific applications.

In the group of work on wikis, most papers concentrate on Wikipedia, the most prominent of all wikis. One important thread of this work has been on how to understand the properties of the social community that has sprung up around Wikipedia. There has been comparisons to print media [9], estimations of the expertise of Wikipedia authors[23], and studies of the elite verses common users [10]. Others have researched how Wikipedia articles accrue edits proportional to their previous edits [26] and even comparisons across different language versions [19].

Another thread of research has been on how to improve the quality of Wikipedia through practical rule changes, community management, and user interface design. A large and diverse user base seems to be correlated to the quality of articles [1] while the costs of conflict and coordination [11] also go up with the population. More general studies have been done on how policy influences communities with mass participation, and looks at examples from Wikipedia to try to deduce design guidelines [12, 24]. Self-supervised information extraction combined with innovative interface design was studied in [25] and specifically designed to encourage communal content creation.

The AI community has recently begun to leverage Wikipedia as a means to solve problems in NLP and information retrieval, among other things. In [28], Wikipedia is used to evaluate ontologies. It has also been used to learn semantic relatedness and semantic interpretation in [27, 8].

Our work on joint process games also has deep connections to work from economics in social choice and voting. The myopic action rule we use in joint process games is similar to fictitious play [20] and the best response dynamics used in potential games [14]. For a good overview of both see [22]. Fictitious play does not apply here because we don't have a stage game. Even if we changed the model to be a repeated normal form game, the stage game payoffs would be dependent on the history and change every round, so the Nash equilibrium would also change. Further, even if fictitious play did apply, none of the known conditions [3] for convergence apply here.

There are two related areas that are worth speaking in depth about. The first is Duncan Black's work [4], now commonly referred to as median voter theory [6]. Variations of median voter theory[5, 18] have suggested models explaining why voting for extreme parties with no chance of winning can benefit the voters by allowing them to communicate and influence the subsequent positions of the winning parties in repeated election settings. Other work has shown how voting can be a process of information aggregation, like Condorcet's jury theorem[17, 16], and the work viewing voting as maximum likelihood estimation [7].

The setting for median voter theory is as follows. Let there be a set of voters with single peaked preferences, i.e. there exists some linear ordering of alternatives such that for every voter, there is a point in the linear ordering with maximum utility, and moving in either direction along this fixed linear ordering monotonically decreases the voter's utility. The strong form of the median voter theorem roughly states that the median voter always gets his most preferred candidate. Consequently, when using the majority rule, if all voters vote for the candidate(alternative) with position closest to their own peak preference, and if there are only two candidates, then the candidates maximize their votes if they commit to the policy position preferred by the median voter.

In joint process games, when we use the Euclidean distance from outcome to goal to determine each agent's reward, the agent's pref-

erences have the singled peaked property. And using the myopic play rule has a limit outcome based on the median of the agents and a set of partition points. However, the differences in turn-taking, length of games, and payoffs are too great between the joint process model and median voter theory for one to be a special case of the other.

But the key take-away of the median voter theorem also holds in our setting; if it is possible to influence the goals of the agents, then the only way to change the outcome is the influence the goal of the median agent and those agents close to the median agent.

The second area we will speak in depth about is Moulin's [15] work on strategy proof voting rules in single-peaked domains. Moulin proves that the set of strategy-proof, anonymous, and efficient voting rules in this domain is exactly equal to the set of voting rules that can be expressed as any function that chooses the median of the  $n$  voter's preferences (represented as the real valued point of the peak) and  $(a_1, a_2, \dots, a_{n-1})$  where  $a_i \in \mathbb{R} \cup \{-\infty, \infty\}$ .

If we let  $(a_1, a_2, \dots, a_{n-1}) = (\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n})$ , then the wiki joint process game has the same outcome in the limit as the strategy-proof, anonymous, and efficient voting rule in the domain of single peaked preferences. The key differences are that in Moulin's voting rule, the agents communicate by expressing a real value between 0 and 1 and reveal their true preference. In joint process games agents have a very restricted action space composed of only two different actions and do not have to fully reveal their preferences to achieve the same outcome. For example, in the instance of a joint process game shown in Figure 1, the agent with a goal of .65 only takes action 0 throughout the entire process, and other than revealing that his goal is below  $C(N)$ , this agent reveals almost nothing! We conjecture that other  $f$  functions determining the outcome of a joint process game will correspond to other assignments of  $(a_1, a_2, \dots, a_{n-1})$  in Moulin's model.

## 5. CONCLUSION

This paper has made the following three broad contributions. First we introduced joint process games. These games are a new model for games with alternating control and a state determined by the history of actions. They capture situations without a definite end, like aggregate ratings and wiki updates. To our knowledge this is the first proposed model to study the theoretical behavior of players updating a wiki article.

Second, we defined the central value of a set of points, and proved that under the simple myopic play rule our joint process game converges to the central value. The actions of all players (except a possible central value player) converge to play uniformly zero or one after sufficient time.

Third, we proved that the joint strategy defined by the myopic play rule is a Nash equilibrium. In addition we characterized all other payoff profiles obtainable in some Nash equilibrium by proving a folk theorem for joint process games.

For future work we hope to quantify how many bits of information agents reveal when using the myopic play rule, and investigate if there are any ways of reducing this. We also are working on characterizing all Nash and subgame perfect equilibrium of this joint process game. Finally, we are looking at other possible functions  $f(\cdot)$  mapping history to outcome to analyze their convergence properties and determine if every voting rule in Moulin's family of voting rules has a corresponding joint process game.

## 6. REFERENCES

- [1] O. Arazy, W. Morgan, and R. Patterson. Wisdom of the Crowds: Decentralized Knowledge Construction in Wikipedia. *SSRN eLibrary*, 2006.
- [2] C. Archibald and Y. Shoham. Modeling billiards games. In *AAMAS '09: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*, pages 193–199, Richland, SC, 2009. International Foundation for Autonomous Agents and Multiagent Systems.
- [3] U. Berger. Fictitious play in  $2 \times n$  games. *Journal of Economic Theory*, 120(2):139 – 154, 2005.
- [4] D. Black. On the rationale of group decision-making. *The Journal of Political Economy*, 56(1):23–34, 1948.
- [5] M. Castanheira. Why vote for losers? *Journal of the European Economic Association*, 1(5):1207–1238, 09 2003.
- [6] R. D. Congleton. The median voter model. In C. K. Rowley and F. Schneider, editors, *The Encyclopedia of Public Choice*, pages 382–386. Springer, 2003.
- [7] V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09)*, 2000.
- [8] E. Gabrilovich and S. Markovitch. Wikipedia-based semantic interpretation for natural language processing. *J. Artif. Int. Res.*, 34(1):443–498, 2009.
- [9] A. Halavais and D. Lackaff. An analysis of topical coverage of wikipedia. *Journal of Computer-Mediated Communication*, 13(2):429–440, 2008.
- [10] A. Kittur, E. H. Chi, B. A. Pendleton, B. Suh, and T. Mytkowicz. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. 2007.
- [11] A. Kittur, B. Suh, B. A. Pendleton, and E. H. Chi. He says, she says: conflict and coordination in wikipedia. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 453–462, New York, NY, USA, 2007. ACM.
- [12] T. Kriplean, I. Beschastnikh, D. W. McDonald, and S. A. Golder. Community, consensus, coercion, control: cs\*w or how policy mediates mass participation. In *GROUP '07: Proceedings of the 2007 international ACM conference on Supporting group work*, pages 167–176, New York, NY, USA, 2007. ACM.
- [13] C. F. Manski. Interpreting the predictions of prediction markets. *Economics Letters*, 91(3):425 – 429, 2006.
- [14] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124 – 143, 1996.
- [15] H. Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, January 1980.
- [16] R. B. Myerson. Extended poisson games and the condorcet jury theorem. *Games and Economic Behavior*, 25(1):111 – 131, 1998.
- [17] T. Piketty. The information-aggregation approach to political institutions. *European Economic Review*, 43(4-6):791–800, April 1999.
- [18] T. Piketty. Voting as communicating. *Review of Economic Studies*, 67(1):169–91, January 2000.
- [19] M. Rask. The reach and richness of wikipedia: Is wikinomics only for rich countries? *First Monday*, 13(6).
- [20] J. Robinson. An iterative method of solving a game. *The Annals of Mathematics*, 54(2):296–301, 1951.
- [21] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39(10):1095–1100, 1953.
- [22] Y. Shoham and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*.

Cambridge University Press, December 2008.

- [23] S. Spek. Wikipedia: organisation from a bottom-up approach, Nov 2006.
- [24] F. B. Viegas. The visual side of wikipedia. In *HICSS '07: Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, page 85, Washington, DC, USA, 2007. IEEE Computer Society.
- [25] D. S. Weld, F. Wu, E. Adar, S. Amershi, J. Fogarty, R. Hoffmann, K. Patel, and M. Skinner. Intelligence in wikipedia. In *AAAI'08: Proceedings of the 23rd national conference on Artificial intelligence*, pages 1609–1614. AAAI Press, 2008.
- [26] D. M. Wilkinson and B. A. Huberman. Cooperation and quality in wikipedia. In *WikiSym '07: Proceedings of the 2007 international symposium on Wikis*, pages 157–164, New York, NY, USA, 2007. ACM.
- [27] E. Yeh, D. Ramage, C. Manning, E. Agirre, and A. Soroa. Wikiwalk: Random walks on wikipedia for semantic relatedness. In *ACL-IJCNLP TextGraphs-4 Workshop 2009*, 2009.
- [28] J. Yu, J. A. Thom, and A. Tam. Ontology evaluation using wikipedia categories for browsing. In *CIKM'07 Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, pages 223–232, November 2007.