Characterizing Involution with Game Theory Models

Weihang Xu
Institute for Interdisciplinary Information Sciences
xuwh19@mails.tsinghua.edu.cn

Abstract

The studying object of this term paper is the "involution" phenomenon, or the scenario where redundant competition reduces every player's utility. To characterize the ineffectiveness brought by internal competition, we use the price of anarchy (PoA) value as the measurement tool, which is applied to the several models that we have built. Our modeling starts from a simple construction, and two generalizations are applied to the model successively, yielding two levels of generalized models.

Introduction

In Tsinghua university, "involution", or "nei juan" in Chinese, is an often-mentioned word, which refers to the ineffective and burdensome internal competition between university students. Such word also applies to a wide range of other scenarios where competition entails additional costs for all competitors, and is gaining increasing popularity in recent years. Due to the ubiquity of such a term, there are reasons to believe that involution is actually a common phenomenon in everyday life, and a game-theoretical analysis of it may provide some useful insights. In this term paper, we provide some game theory models trying to characterize the essence of involution, and we hope that analysis of such models may reveal some properties of the phenomenon.

The First Attempt: A Purely Competitive Model

Model Construction

Intuitively, involution describes the case where internal competition reduces the social welfare, therefore we should construct a model where the Price of Anarchy (PoA) is high.

Our modeling starts by considering the involution between university students. Suppose n students $s_1, s_2, \ldots, s_n$ attended a class in Tsinghua university, and they are competing for good grades in that class. In our model, students dedicate some amount of work (characterized by a positive rational number) to that class, and get their grades correspondingly.

For the $i^{th}(1 \leq i \leq n)$ student, suppose the amount of work that he[1] spends on this class is w_i ($w_i \in \mathbf{R}^+$), and the grade that he finally gets is g_i. In this section, we only consider the case where the grading is purely competitive, i.e., each student's grade only depends on his relative ranking in the class. Under this assumption, we set that

$$g_i = \frac{w_i}{\sum_{j=1}^n w_j} (\forall i \in \{1, 2, \ldots, n\}).$$

Denote the utility function of s_i with u_i. Obviously, g_i represents the reward of s_i, and w_i represents the cost of s_i. Therefore we set

$$u_i(w_1, \ldots, w_n) = g_i - w_i = \frac{w_i}{\sum_{j=1}^n w_j} - w_i,$$

which completes the model construction.

Formally, our model describes a normal form game G with n players $s_1, s_2, \ldots, s_n$, in which player s_i chooses strategy w_i from $\mathbf{R}^+$ to maximize his utility function u_i.

Analysis of Model

The primary step of analyzing our model is to find all the Nash Equilibria (abbreviated as NE starting from now) of the game, of which we have the following result.

Proposition 1. *There exists an unique NE in the game G.*

Proof. Suppose the strategy profile $w^* = (w_1^*, \ldots, w_n^*)$ is an NE, then w_i^* is the optimal strategy for s_i under w^*, which yields n equations

$$w_i^* = \arg\max_w u_i(w_1^*, \ldots, w_{i-1}^*, w, w_{i+1}^*, \ldots . w_n^*), \forall i.$$

The above equations imply that

$$\frac{\partial u_i}{\partial w_i}(w_1^*, \ldots, w_n^*) = 0, \forall 1 \leq i \leq n, \tag{1}$$

$$\Rightarrow -1 + \frac{\sum_{k \neq i} w_k^*}{\left(\sum_{j=1}^n w_j^*\right)^2} = 0, \forall 1 \leq i \leq n, \tag{2}$$

Denote $S = \sum_{j=1}^n w_j^*$, then equation system 2 implies

$$S - w_i^* = S^2, \forall 1 \leq i \leq n$$

[1] For simplicity of addressing, here we suppose all students are male.

summing over all i yields

$$nS - \sum_{j=1}^{n} w_j^* = nS^2 \Rightarrow S = \frac{n-1}{n} \Rightarrow w_i^* = \frac{n-1}{n^2}, \forall i.$$

By simple calculations, one can verify that $w_1^* = \cdots = w_n^* = \frac{n-1}{n^2}$ is indeed an NE, hence the unique NE of the game. Under such NE, the utility of each student is $u_i = \frac{1}{n^2}$, and the social welfare is $\sum_{j=1}^{n} u_j = \frac{1}{n}$. □

The next proposition characterizes the social optimal value of the game.

Proposition 2. *The social optimal value of G is 1, i.e.,*

$$\sup_{(w_1,\ldots,w_n)} \sum_{j=1}^{n} u_j(w_1, \ldots, w_n) = 1.^2 \tag{3}$$

Proof. First note that

$$\sum_{j=1}^{n} u_j(w_1, \ldots, w_n) \leq \sum_{j=1}^{n} \frac{w_j}{\sum_{k=1}^{n} w_k} = 1.$$

On the other hand, consider an infinite sequence of strategy profiles $\{(w_{1,t}, \ldots, w_{n,t})\}_{t \in \mathbf{Z}^+}$ satisfying $w_{i,t} = \frac{1}{t}, \forall i$, then one can easily verify that

$$\lim_{t\to\infty} \sum_{j=1}^{n} u_j(w_{1,t}, \ldots, w_{n,t}) = \lim_{t\to\infty} 1 - \frac{n}{t} = 1,$$

□

By proposition 1 and 2, the price of anarchy of the game is

$$PoA = \frac{\sup_{(w_1,\ldots,w_n)} \sum_{j=1}^{n} u_j(w_1, \ldots, w_n)}{\min_{(w_1,\ldots,w_n)\in S^{NE}} \sum_{j=1}^{n} u_j(w_1, \ldots, w_n)} = \frac{1}{1/n} = n$$

here S^{NE} denotes the set containing all NE of the game.

Through our analysis, one can see how the internal competition reduces everyone's gain: when every player maximizes his utility, the social welfare is greatly reduced, and the PoA becomes n. We hope that such a model captures the structure of involution: when students compete for a higher grade, their effort cancels out, and results in a lower utility for everyone.

Generalizing the model by adding non-competitive elements

Model Construction

One might argue that the naive model that we built in the last section is unable to capture some properties of the involution due to its purely-competitiveness. That is, we assume that the grading of students only depends on their relative ranking, while teachers in real life usually consider other factors during grading students, like their performances. For example, suppose one teacher teaches the same course for two classes A and B, and students in class A performs generally better than students in class B, then it is reasonable to assume that students in class A will get generally higher grades. To put it another way, a possible interpretation of "involution" could be " competition hindering cooperation", therefore a purely-competitive model is not enough for characterizing it, and non-competitive elements must be introduced.

To generalize the model, we still consider the scenario described in the previous section, where n students $s_1, \ldots, s_n$ attend the same class in Tsinghua university, but the following changes are made to the model: consider there exists a "basic requirement of the class" b, and the grade g_i that student s_i gets is now defined as

$$g_i = \frac{w_i}{b + \sum_{j=1}^{n} w_j}$$

while the utility of s_i being defined as

$$u_i(w_1, \ldots, w_n) = ag_i - w_i = \frac{aw_i}{b + \sum_{j=1}^{n} w_j} - w_i.$$

Also $a > b$ is required.

After adding the parameter b, the grade of s_i depends not only on his relative ranking among students, but also the basic requirement of the class b. When the requirement becomes stricter, b increases, all g_i decreases.

Also note that another parameter a is now added into the utility function, it describes how students values between their grade and their work. A large a indicates that students think that grades are worth much working, and vice versa. In our previous model, $a = 1$.

Analysis of Model

Denote the generalized version of game G described in this section with H, then results similar to proposition 1 and 2 still hold.

Proposition 3. *There exists an unique NE in the game H.*

Proof. Suppose $w^* = (w_1^*, \ldots, w_n^*)$ is an NE in the game H, we could solve w^* following the same process as in the previous section.

First note that equation 1

$$\frac{\partial u_i}{\partial w_i}(w_1^*, \ldots, w_n^*) = 0, \forall 1 \leq i \leq n$$

still holds, therefore

$$-1 + \frac{a\left(b + \sum_{k\neq i} w_k^*\right)}{\left(b + \sum_{j=1}^{n} w_j^*\right)^2} = 0, \forall 1 \leq i \leq n, \tag{4}$$

Denote $S = b + \sum_{j=1}^{n} w_j^*$, we get

$$a(S - w_i^*) = S^2, \forall 1 \leq i \leq n$$

[2]Note that here the social optimal value is defined as the supremum, instead of maximum, of all achievable social welfares. This is because the strategy space of our model $((\mathbf{R}^+)^n)$ is continuous, therefore the maximum of all social walfares may not always exist.

summing it over all i yields

$$a(nS-(S-b))=nS^2$$

$$\Rightarrow S=a\cdot\frac{n-1+\sqrt{(n-1)^2+(4bn)/a}}{2n}$$

$$\Rightarrow w_i^*=\frac{(a-2b)n-a+\sqrt{a^2(n-1)^2+4abn}}{2n^2},\forall i.$$

For the same reason as in the previous section, one can verify that the strategy profile that we solved above is an unique NE of the game, and its corresponding social welfare is

$$\begin{aligned}
&\sum_{j=1}^{n}u_j(w_1^*,\dots,w_n^*)\\
&=\frac{a(S-b)}{S}-S+b\\
&=an+b-(n+1)S\\
&=\frac{(n^2+1)a+2bn-(n+1)\sqrt{a^2(n-1)^2+4abn}}{2n}
\end{aligned}$$

□

Proposition 4. *The social optimal value of H is $a+b-2\sqrt{ab}$.*

Proof.

$$\begin{aligned}
&\sup_{(w_1,\dots,w_n)}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)\\
=&\sup_{(w_1,\dots,w_n)}\frac{a\sum_{j=1}^{n}w_j}{b+\sum_{j=1}^{n}w_j}-\sum_{j=1}^{n}w_j\\
=&\sup_{(w_1,\dots,w_n)}a+b-\left(\frac{ab}{b+\sum_{j=1}^{n}w_j}+\left(\sum_{j=1}^{n}w_j+b\right)\right)\\
=&\,a+b-2\sqrt{ab}
\end{aligned}$$

The social optimum is achieved if and only if $\sum_{j=1}^{n}w_j=\sqrt{ab}-b$ (note that $a>b$). □

Now we can measure how the involution damages the social welfare by calculating the PoA.

Proposition 5. *The price of anarchy of game H satisfies*

$$n\cdot\frac{a+b-2\sqrt{ab}}{a-3b}\geq PoA\geq n\cdot\frac{\sqrt{a}-\sqrt{b}}{\sqrt{a}+\sqrt{b}}.$$

Proof. Define S^{NE} in the same way as the previous section, then by proposition 3, the social welfare of NE in game H satisfies

$$\begin{aligned}
&\min_{(w_1,\dots,w_n)\in S^{NE}}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)\\
=&\frac{(n^2+1)a+2bn-(n+1)\sqrt{a^2(n-1)^2+4abn}}{2n}\\
=&\frac{(n^2+1)a+2bn-(n+1)\sqrt{(an+2b-a)^2+4b(a-b)}}{2n}\\
\leq&\frac{(n^2+1)a+2bn-(n+1)(an+2b-a)}{2n}\\
=&\frac{a-b}{n}
\end{aligned}$$

Then by proposition 4,

$$\begin{aligned}
PoA&=\frac{\sup_{(w_1,\dots,w_n)}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)}{\min_{(w_1,\dots,w_n)\in S^{NE}}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)}\\
&\geq\frac{a+b-2\sqrt{ab}}{(a-b)/n}\\
&=n\cdot\frac{\sqrt{a}-\sqrt{b}}{\sqrt{a}+\sqrt{b}}
\end{aligned}$$

The upper bound could be proven similarly:

$$\begin{aligned}
&\frac{a-b}{n}-\min_{(w_1,\dots,w_n)\in S^{NE}}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)\\
=&\frac{n+1}{2n}\cdot\frac{4b(a-b)}{\sqrt{(an+2b-a)^2+4b(a-b)}+(an+2b-a)}\\
\leq&\frac{n+1}{2n}\cdot\frac{4b(a-b)}{(an+2b)+(an+2b-a)}\\
\leq&\frac{n+1}{n^2}b
\end{aligned}$$

$$\begin{aligned}
PoA&=\frac{\sup_{(w_1,\dots,w_n)}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)}{\min_{(w_1,\dots,w_n)\in S^{NE}}\sum_{j=1}^{n}u_j(w_1,\dots,w_n)}\\
&\leq\frac{a+b-2\sqrt{ab}}{(a-b)/n-(n+1)b/n^2}\\
&\leq n\cdot\frac{a+b-2\sqrt{ab}}{a-3b}
\end{aligned}$$

□

Proposition 5 uses PoA to measure the damage of involution, and one can see that it has the relative magnitude of $PoA=\Theta(n)$. For example, in the simple model G built in the previous section, $a=1,b=0$, and both the upper and lower bound of PoA in proposition 5 collapses to n, which coincides with our calculation in the previous section.

Moreover, note that when b increases, by proposition 5, PoA of H decreases, which might provide us with some useful insights: at those courses teaching easier contents, the involution becomes only fiercer, each student acquires less knowledge while dedicating more effort to the ineffectual competition. One might say that the involution compensates the easiness of course contents.

Another observation is that PoA increases when a increases, whose implication is quite straightforward. The involution worsens when students are hardworking (like those students in Tsinghua). On the contrary, students benefits more when they are lazy!

Further Generalization: Introducing Asymmetry to The Model

Model Construction

So far, all models that we have built are symmetric: the utilities and strategies are all symmetric among different players. One might wish to consider the individualistic differences between different students, which could be characterized by the further generalized model describe in this section.

Consider a generalized version I of game H, where the n students $s_1, \ldots, s_n$ have the following utility functions

$$u_i(w_1, \ldots, w_n) = a_i g_i - w_i = \frac{a_i w_i}{b + \sum_{j=1}^n w_j} - w_i, \forall 1 \leq i \leq n.$$

Here we substitute the parameter a in game H with n parameters $a_1, \ldots, a_n$, where each $a_i > b$ is a positive real number representing the tolerance of working of student s_i. Other settings in H remain unchanged.

Analysis of Model

Proposition 6. *There exists an unique NE in game I.*

Proof. Define $w^* = (w_1^*, \ldots, w_n^*)$ in same way as in game G and H. Similar to previous calculations, we have

$$-1 + \frac{a_i \left(b + \sum_{k \neq i} w_k^*\right)}{\left(b + \sum_{j=1}^n w_j^*\right)^2} = 0, \forall 1 \leq i \leq n, \qquad (5)$$

Define $S = b + \sum_{j=1}^n w_j^*$, we get

$$S - w_i^* = S^2/a_i, \forall 1 \leq i \leq n$$

summing it over all i yields

$$nS - (S - b) = AS^2$$

$$\Rightarrow S = \frac{(n-1) + \sqrt{(n-1)^2 + 4Ab}}{2A}$$

where $A = \sum_{j=1}^n 1/a_j$.

Therefore $w_i^* = S - \frac{S^2}{a_i}$ where $S = \frac{(n-1)+\sqrt{(n-1)^2+4Ab}}{2A}$. For the same reason as in previous sections, one can verify that the $(w_1^*, \ldots, w_n^*)$ that we solved above constitutes the only NE of the game. □

Now consider the utility of s_i under the NE:

$$u_i(w_1^*, \ldots, w_n^*) = \frac{a_i w_i^*}{b + \sum_{j=1}^n w_j^*} - w_i^* = a_i - S - w_i^*$$

$$= \frac{S^2}{S - w_i^*} - S - w_i^* = \frac{{w_i^*}^2}{S - w_i^*}$$

For simplicity of notation, define $u_i^* = u_i(w_1^*, \ldots, w_n^*)$. W.L.O.G, suppose $a_1 \geq \cdots \geq a_n$, then since $u_i^* = \frac{{w_i^*}^2}{S - w_i^*}$ we have

$$w_i^* = S - \frac{S^2}{a_i} \Rightarrow w_1^* \geq \cdots \geq w_n^* \Rightarrow u_1^* \geq \cdots \geq u_n^*.$$

Which means that students with higher a_i gains higher utility in the NE. Recall that a_i records student s_i's ability of tolerating hardworking. So our models shows that those students with higher willingness of competing (whom people usually call by "King(Queen) of Involution", or "juan wang" in Chinese) indeed become winners of the competition.

Similar to proposition 4 and 5, the social optimal value and price of anarchy of the game I could also be calculated, and results similar to the ones we have derived in game H could be obtained. Since the calculations are cumbersome and results are similar, the details are omitted here for the sake of conciseness.

Hereby we end our construction of models trying to characterize the involution phenomenon. The mathematical tools that are utilized for building and analyzing these models are rather naive, but we hope that simple models might also be able to reveal some interesting properties of our studying objects. A more refined analysis of the phenomenon could only be leaved for the future.