CS-352 Theoretical Computer Science	October 3, 2012

Lecture 5

Lecturer: Aleksander Madry *Scribes: Chidambaram Annamalai and Farah Charab*

1 How to Play Games?

The topic of this lecture is game theory, a field that aims at understanding interaction of rational parties (agents) in the situation of conflict. Here, "rational" means that the agents will always purse the actions that leads to their maximum benefit, where the "benefit" is captured via a notion of *utility* function evaluating each possible outcome of the interaction. The notion of utility is intentionally left very broad here and, for example, should not be confused with only monetary reward (although such a reward might be a part of it). It can, in particular, incorporate many seemingly "irrational" concepts like altruism or spitefulness. So, once utility function rewards, say, altruism, being altruistic becomes a perfectly rational behavior.

The key motivation of game theory stems from the belief that many real-world situations and problems can be casts in such game-theoretic framework (even though the exact rules of the games might be unclear).

The main tool of game theory in its endeavor are mathematical *games* – certain thought experiments that model (and thus predict) the outcome of interactions of rational agents when they are pursuing different (usually conflicting) goals. Before we define these games formally, let's see a few examples.

2 Game Examples

2.1 Prisoner's Dilemma

This is probably the most well-known game. The situation here is the following. We have two prisoners (players) that committed a crime together and are now interrogated in separate rooms. Each one of them has a choice (two possible actions): confess or remain silent. In case both of them remain silent, the police cannot prove their guilt, so they will both serve one year in prison just based on some minor other offenses that the police is able to charge them with. In case both of them confess, they will both be sentenced for three years in prison (which is a reduced penalty due to their collaboration with the police). Now, if one of them confesses, but the other one remains silent, then the police can use the first one as a witness against the other one. In this way, the latter prisoner will be sentenced to maximum penalty for the crime (five years), and the witness is able to strike a deal that allows him/her to walk away free.

We represent all these possible outcomes and corresponding utilities of the players (that are proportional to the number of years they are sentenced to) in the table below. We use a convention in which the rows are indexed by possible actions of one player and the columns by the possible actions of the second player. In this setting, each entry corresponds to one possible outcome and the utility of both players is just a vector of two numbers in that cell.

	Prisoner 1 C	Prisoner 1 S
Prisoner 2 C	(-3,-3)	(0,-5)
Prisoner 2 S	(-5,0)	(-1,-1)

The question now is: what will each of them do *if they are rational*?

Looking at the utilities, it is intuitively clear that the "best" possible outcome for both of them would be if they both remained silent. However, the problem is that to ensure such an outcome, they would need to cooperate and trust each other. Unfortunately, in the game-theoretic setting the only type of "trust" one can expect is the one that stems from alignment of utilities, i.e., two agents will cooperate only if it is in the best interest of both of them.

As a result, it turns out that the only rational outcome of this game is for both prisoners to confess. To see why, let's look at the reasoning of Prisoner 1. If he/she remains silent, he/she would suffer a utility of -1 in case of Prisoner 2 remaining silent too and -5 if Prisoner 2 decides to confess. However, if he/she confesses, then she/he gets *even better* outcome (0 and -3 respectively) in *both* cases. So, there is no reason to not confess! (After all, he/she doesn't care about the other prisoner, he/she is rational).

Clearly, the same reasoning would apply to Prisoner 2. Thus, as we see, there is absolutely no benefit for them in remaining silent - confessing is a better outcome (i.e., provides better utility) *irregardless* of what the other prisoner does.

This is the first (of many) example when acting rationally might not lead to the truly "best" outcome, since such outcomes might not be "stable" due to at least one of the agents being tempted to deviate to improve his/her own utility even further.

2.2 Pollution Game

In this game, we have n players (countries), and each country has to choose to either pass a legislation that restricts the pollution produced by the country's industry, or to not pass it (and thus pollute). The utility function of each player is shown below.

$$u_i = \begin{cases} -3 - p & \text{if i-th country passes the legislation} \\ -p & \text{if i-th country decides to pollute} \end{cases}$$

where p is the total number of countries that decide to pollute.

So, in other words, each country suffers the penalty of $-p$ (even if it is not one of the polluters) as the pollution produced by other countries affects it, and the -3 term is the penalty that country's industry suffers due to the smaller efficiency caused by anti-pollution restrictions.

Once again, it turns out that the only stable solution with respect to all the players being rational, is that everyone decides to pollute. To see that, let's focus on player i and fix the decisions of all the other players and let p' be the number of them that decided to pollute. This country is facing two possible utilities: $-3 - p'$, if it decides to pass the legislation; or $-(p' + 1)$ if it decides to pollute. Clearly, since each players wants to maximize his/her utility, it always has an incentive to pollute.

Thus, here again we see vast sub-optimality of the "rationally" stable outcome: the utility of each country is $-n$ versus -3 that would result from all the countries adopting the policy.

2.3 Tragedy of the Commons

Here, we have n users and an Internet link of capacity 1 to be shared by all of them. In this setting, the action of user i is just to choose a share of the link x_i that he/she demands. Now, the utility function is as shown below.

$$u_i(x_1, \ldots, x_n) = \begin{cases} 0 & \text{if } \sum_j x_j \geq 1 \\ x_i(1 - \sum_j x_j) & \text{otherwise} \end{cases}$$

To understand the utility structure, note that the benefit of a user from his/her share is proportional not only to the size of this share, but also to the remaining unclaimed bandwidth $(1 - \sum_j x_j)$ of the link. (One can think that this term models the latency of the link.) So, clearly, to utilize the link best it would be important to make sure that the users are not overusing the capacity and thus leave enough unclaimed bandwidth to allow everyone to get a good benefit out of their shares. But what will happen if the users are behaving rationally?

Let us consider only situations in which the users do not claim the whole capacity of the link. Otherwise, the scenario is not too interesting as no one derives any positive utility.[1]

Let us focus on a particular user i and let us fix the claims of all the other users. We look at what is the condition on the choice of x_i that would make the user i not want to deviate. Let $t = \sum_{j \neq i} x_j \leq 1$ be the sum of claims of the other users. Clearly, in terms of t and x_i the utility of player i is $x_i(1 - t - x_i)$, which is maximized if $x_i = \frac{1-t}{2}$. Since this condition holds for each user i, it follows that the only "rationally" stable solution in this setting is a one with each $x_i = \frac{1}{n+1}$.

So, once more we see that: the best utilization of the link would result from having each x_i equal to $\frac{1}{2n}$ – this gives a total utility of $\frac{1}{4}$ – while the best "rationally" stable solution achieves total utility of only

$$\frac{n}{n+1}\left(1 - \frac{n}{n+1}\right) = O\left(\frac{1}{n}\right),$$

as the users trying to increase their marginal gain will overuse the link.

Note, however, that the above example is different from the previous ones in one important aspect. Here, the best action of the player depends on what the other players do, while previously there was always an action that was the most desirable irregardless of other player's actions. (We will come back to this later.)

2.4 Penalty Shot Game

Our final example is a game between a shooter trying to shot a penalty kick, and a goalkeeper that tries to defend it. Each one of them can choose two possible actions: shot/jump either left or right. Clearly, if they both choose the same side then it is goalkeeper's gain as he/she can save, otherwise the shooter managed to fool the goalkeeper and scores the goal. The following table summarizes this game.

	Shooter R	Shooter L
Goalkeeper R	(1,-1)	(-1,1)
Goalkeeper L	(-1,1)	(1,-1)

[1] Note that this restriction is *with* loss of generality, as there actually are stable solutions with users overclaiming the link. Namely, if the sum of the claims $\sum_j x_j$ of all the users is at least two, i.e., it is much larger than the available bandwidth, then each of the users gets zero utility out of it, but has no incentive to deviate as he/she alone is not able to unblock the link (and thus make his/her utility positive).

What are the "rationally" stable outcomes here?

It is not hard to see that for any choice of actions by both players, it is always the case that one of them has incentive to deviate. This would suggest that there is no such stable outcome here.

However, the intuition tells us that there should be a stable state: a one in which the players are allowed to *randomize* their actions. Indeed, if we consider such randomized outcomes then an outcome in which each player chooses an action with the same probability is a stable one. Also, it is not hard to see that this is such only stable outcome – if distributions over the actions of one player is even slightly biased, the other player has an incentive to just choose a deterministic strategy that takes advantage of this bias – e.g., if shooter's distribution is biased to the right, then the goalkeeper has an incentive to always go with right too – and we already know that deterministic choices do not give rise to a stable outcome.

3 Abstract Setting

After seeing the examples, we are ready to make our game-theoretic framework precise. Formally, a game consists of n players (*agents*), each of whom has a set $\mathcal{S}_i$ of possible strategies or *actions* that they can choose. Each choice $s = (s_1, \ldots, s_n) \in \mathcal{S}$ of strategies by all the players is an *outcome* or *strategy profile* s of the game, where $\mathcal{S}$ is a Cartesian product of all the $\mathcal{S}_i$. Each such outcome induces, for each player i, *utility* $u_i(s)$, where each *utility function* $u_i : \mathcal{S} \to \mathbb{R}$ is given as a part of the game description.

Note that this definition allows us to also easily incorporate the randomized strategies (that we, e.g., needed in case of the Penalty Shot game). We just need to substitute each $\mathcal{S}_i$ (and thus also $\mathcal{S}$) with its "simplex closure" $\overline{\mathcal{S}}_i$ (and $\overline{\mathcal{S}}$) that contains all the convex combinations of the actions from $\mathcal{S}_i$ (and $\overline{\mathcal{S}}$ is again just a Cartesian product of $\overline{\mathcal{S}}_i$). We will sometime call the deterministic strategies (from $\mathcal{S}_i$) *pure*, and the randomized ones (from $\overline{\mathcal{S}}_i$) *mixed*.

4 Classification of Stable Outcomes

Once we formally defined the game, we can proceed to classifying the types of stable outcomes that arise from rational behavior of the players. First, however, we will introduce one piece of useful notation. For a given outcome s, by s_{-i} we denote the vector s with i-th coordinate removed. So, s_{-i} represents the strategies of all the players other than i.

Definition 1 *An outcome $s^* \in \overline{\mathcal{S}}$ is a* (weakly) dominant strategy *for a given game iff for any player i and any outcome $s' \in \overline{\mathcal{S}}$*

$$u_i(s_i^*, s'_{-i}) \geq u_i(s').$$

Furthermore, we say that s^* is a strictly dominant strategy *if the above inequality is strict for every i and $s' \in \overline{\mathcal{S}}$ with $s'_i \neq s_i^*$.*

In other words, each s_i^* is a choice of strategy from which there is no incentive to deviate *no matter what the rest of the players is doing.* Note that we had such a (strictly) dominant strategies in case of Prisoner's Dilemma and Pollution game above. Also, it is easy to see that the strictly dominant strategy has to be unique (while there might be many weakly dominant one).

In a sense, existence of dominant strategy is something extremely desirable from the game-theoretic point of view. After all, any rational player is strongly compelled to play such a strategy, even if the other players are not behaving rationally. Unfortunately, as we have already seen in our examples, not every game possesses such a strong notion of stable outcome. So, to encompass the other (weaker) notions, we introduce the following definition.

Definition 2 *An outcome $s^* \in \overline{\mathcal{S}}$ is a* (mixed) Nash equilibrium *for a given game iff for any player i and any strategy $s_i' \in \overline{\mathcal{S}}_i$*

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i', s_{-i}^*).$$

Moreover, we say that s^* is a pure Nash equilibrium *if all strategies s_i^* are deterministic, i.e., $s^* \in \mathcal{S}$.*

Informally, a Nash Equilibrium is an outcome in which no player can achieve a better utility by *unilaterally* changing his/her strategy. Note that the key word here is "unilaterally", i.e., the player has no incentive to change his/her action only if he/she believes that the rest of the players will actually follow s^*. This is a crucial difference between Nash equilibrium and dominant strategy.

5 Braess' Paradox

We have already seen some examples in which rationality of players leads to vastly suboptimal outcomes. Here is one more example of how selfish players' behavior can, quite counter-intuitively, produce a net loss for all the players involved.

Consider a simple road network that connects two cities s and t via two different roads with two segments each. Each segment has a delay that is either proportional to the flow of traffic on that road, or always equal to 1 - cf. Figure 1 a).

Imagine now that we have a large number of (rational) players that want to simultaneously get from s to t. We will think of them as a flow of one going from s to t. Clearly, in this setting the only Nash equilibrium is routing half of the flow on the upper path and another half on the lower one. Each of these paths incurs a delay of $\frac{3}{2}$ so there is no incentive for anyone to switch. Therefore, the average delay here is $\frac{3}{2}$ and, in fact, this is the optimal average delay that one can achieve (even when everyone cooperates).

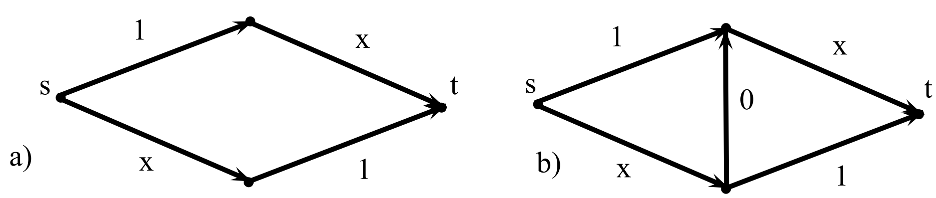

Figure 1: Illustration of Braess' paradox.

Now, consider the same network after augmenting it with a new road that has zero delay and connects the midpoints of the two paths as in Figure 1 b). We would expect that adding a new road will only increase the average delay. However, paradoxically, something exactly opposite happens! It is not hard to see that the addition of this new road changed the Nash equilibrium for this network to a one when everyone goes through this added edge and thus experiences a delay of 2. So, adding a new road to our network made its throughput worse with respect to rationally behaving commuters.

6 Modeling Considerations

Although the model of games we presented above is quite compelling, it is not perfect. In particular, it relies on some implicit assumptions that might or might not be reasonable. We discuss them briefly below.

6.1 Existence and Convergence Issues

Even though in our examples of games we were able to point out a stable outcome, it is not clear at all at this point if we are able to do it for every game that we are presented with. So, one can wonder how useful the concepts of our equilibria are if maybe only very special games exhibit it. Fortunately, as we will see in the next lecture, one can show that for a large class of games such equilibria indeed exist.

This is quite reassuring – we at least can be confident that the objects we are studying are guaranteed to exist – but still leaves out an outstanding question: why do we expect that interactions of rational agents will always (reasonably fast) converge to them?

Answer to this question is even more unclear as we do not know how to (algorithmically) find equilibria of an arbitrary game. In fact, we have some evidence that suggests that it might be a very hard task. So, what is so magical about interactions of rational agents that allows them to always efficiently converge to such an equilibrium?

Also, even if the agents indeed converge to an equilibrium, there might be many possible choices and each one of them might favor a different agent. How do the agents "agree" on which equilibrium to converge to?

These questions are not addressed at all in the model. One just assumes that the convergence happened and studies the static case.

6.2 Perfect Rationality

An important underlying assumption in all the game-theoretic considerations - especially, in the context of Nash equilibrium - is that all the players are perfectly rational. The potential problem with that is not only of philosophical nature – after all, as we already mentioned, a lot of idiosyncrasies of human nature can be incorporated via appropriate utility function – but also a practical one. We can only imagine how complex the real-world games are. How can we expect anyone to be able to analyze them and come up with a perfectly rational way to play them?

6.3 Risk Aversion

One of the very questionable modeling choices in game theory – called by some "the biggest lie of game theory" – is the assumption that rational players will focus on maximization of their expected utility.

To see why this is questionable, imagine a lottery in which, after buying a ticket to participate, a fair coin is tossed repeatedly until heads turns up. Then, you receive $2^{\#\text{tosses}}$ CHF as your reward. Now, if the price of the ticket for this lottery was $1,000,000$ CHF, would you buy it?

It follows from a simple calculation that the expected reward in this game is:

$$\mathbb{E}(\text{payoff}) = \sum_{i \geq 1} \frac{1}{2^i} 2^i = \infty,$$

and hence it is always a rational choice to participate in this lottery. However, most people would decline to do that since they are *risk averse*. That is, humans are reluctant to take actions that with some – even very low probability – could cost them a lot, even if the expected payoff is large. (Of course, again, risk aversion can be incorporated into utility, but one has to do it explicitly.)

6.4 Collusion

Finally, yet another important implicit game-theoretic assumption is that players cannot communicate or make payments to each other, i.e., collusion is forbidden.

To see why this is crucial, consider Prisoner's Dilemma game in which both players strike beforehand a binding deal that stipulates that if a player decides to remain silent then he/she receives a payment of $1,000,000$ CHF from the other player. (Here, $1,000,000$ CHF is meant to correspond to an amount of money that is large enough to outweigh spending additional one or two years in prison.) Now, if we analyze the game again with this new ingredient then actually the dominant strategy of both players is to remain silent (and thus get the best joint outcome)! (Also, note that this dominant strategy does not even result in transfer of money, as both players owe the same amount to each other. All it took is for this reward just to exist - even though it was never really executed.)

As we might imagine, preventing collusion of players in the real-world is pretty impossible, which makes many protocols that are perfectly sound from purely game-theoretic point of view, miserable fail in practice. (There are known high-profile cases that happened in the past featuring such failures.).