Artificial Intelligence for Economics

Prof. Dripto Bakshi

Humanities and Social Sciences

Indian Institute of Technology Kharagpur

Week – 05

Lecture - 24

Lecture 24 : Network Economics

Welcome to this lecture on artificial intelligence in economics. In this lecture and the next, we will look at network economics. Now, there are two ways to view this problem. First, we look at random network formation, and the other type involves examining strategic network formation. Let's understand what the fundamental difference is between these two. Random network formation is where you have a bunch of players, and the links between the players are formed randomly.

For example, the way COVID spreads. So, COVID spreads through a network. I am infected by you. I'm going to infect a few other people I come in contact with, and so on and so forth.

Now, that's a random network. If I come into contact with a COVID patient, then, with a certain probability, I may also get infected with COVID. So there is a connection between that person and me. He infected me with COVID.

Now, in strategic network formation, on the other hand, it is a different paradigm. In this situation, two players establish a connection. Between them, if it is beneficial for both to do so. Instead of a link being formed randomly between two players, here the players are deciding to form a link. Where do we see such instances? Professional networks and trade networks are where we see instances of strategic network formation. Now, when we analyze strategic networks, there are essentially two issues that concern us. Of course, there are many topics, and in this course, I can't cover everything. I would like to introduce you to the basics and the most fundamental aspects of strategic network formation. The two fundamental issues that are often in conflict, as we will see as we move on, are stability and efficiency.

Stability and efficiency are two important features. So, what is stability? Given a set of features and their utility functions, stability addresses the individual incentives of players to establish or sever ties. Whereas efficiency refers to the maximization of social welfare, it is concerned with maximizing the total payoff or total utility of all the players in the network. We will formally define these terms in a few minutes. Okay, so now let's formalize and get acquainted with the basic notations.

So, what do we have? We have a set of players, labeled $1, 2, \ldots, n$. G_n is the set of all possible networks that can be formed among the set of n players.

So, there are all the possible networks that can be formed. The set of all possible networks is represented by G_n. Now the utility of player I depends on the network that has been formed. In different networks, I am connected to different people, and those different people are connected to all other people. The amount of payoff a certain person receives depends on the network in place, so the existing network determines the payoff for a person who is part of it.

So, if G is a network that is one of all the possible networks, then the UIG represents the utility of player I when the network is G. By the way, utility represents all benefits net of costs, as link formation also involves costs, as we will see. Anyway, the bottom line is that we have a bunch of players who can be thought of as the vertices of a graph. The payoff of any player depends on the graph that exists and the edges of the graph. The capital "G" and "N" denote the set of all possible graphs.

Great! Now, we see that for different networks, players will have different payoffs. That's fine. Now, the question is: what is a pairwise stable network? Now I'm going to introduce this term. Let's understand it. So, let's say we have a network G. So the network could be of this kind, let us say: I have player 1 connected to player 2, connected to player 3, and connected to players 4 and 5, something like this.

So, that is a network. Now, a particular network G is said to be pairwise stable. Remember, we had two fundamental issues that we want to discuss when dealing with networks, specifically strategic networks. So the first thing is stability, and the other one is efficiency. When it comes to stability, this is one notion of stability: pairwise stability.

Now let's understand what pairwise stability is and when a network is called a pairwise stable network. A network is called pairwise stable if, for all edges, whenever there is an edge i, j in the network G, the payoffs of both players i and j are better off with the edge existing than without it. Which in turn means that both i and j have no incentive to delete the edge $\{i, j\}$. Remember, they can delete the edge $\{i, j\}$.

I am connected to you. I can choose not to connect with you. So, if any of them have an incentive to delete the edge $\{i, j\}$, they can. A pairwise stable network is one in which no two players connected by an edge have an incentive to delete that edge in the graph.

That's number one. Also, if there does not exist an edge ij, then the payoff of i and the payoff of j must be considered. If the payoff of i is greater with edge ij existing, then the payoff of j will necessarily be smaller with ij existing, which means there cannot exist an edge that is not present in the network. Such that if an edge comes into being, both players involved in the edge will be better off. Okay, I repeat once more: what does this tell you? It tells you that if an edge does not exist in the graph, in a pairwise table graph or a pairwise table network, then.

Bringing that edge into the network, while keeping everything else constant, cannot make both people or both players involved in the edge better off, nor can it make even one person better off while keeping the other person as good as he was. So, if I introduce an edge that is not present in a pairwise stable network, it must be beneficial for one person and harmful for the other. So, what is a pairwise stable network, to put it very simply? It is a network in which no two players connected by an edge have an incentive to delete the edge, and no two players not connected by an edge have an incentive to establish one between them. Okay.

That's pairwise stable. Now, this notion of pairwise stability has some limitations. Well, first, pairwise stability assumes that no player in a pairwise stable network will benefit from severing or deleting an edge.

So, as we saw in the first example, no player has an incentive to delete an edge. In the second example, we observed that we won't have a pair of players who are both better off, or one player being better off while the other player is at least as good as he was, by establishing a new edge. However, they may benefit by deleting a few links at a time; yes, for one edge and one player, it's true. But what if they bring two or three edges at once? That might be profitable.

So, I don't know. Also, pairwise stability assumes that deviations occur by two players at a time. One or two players at a time. But what if there is a complex reorganization? What if ten players decide to change, form, or delete links? So the idea of setterist parabus goes away.

Correction: So the idea of setterist parabus disappears. So, these are essentially the limitations of pairwise stability. Another very conspicuous limitation is that a pairwise stable network may not exist. Let's take a look at this four-player network. So, if I'm here —let's say I start here, at the point where the tick is.

So the payoff is 0, 7, 0, and 7. Now, if these two players form a link, their payoff will increase by 7. So, they will form a link. So if I'm here, if I'm in the tick, I will connect to the hash network. Now, if I'm in the hash network, you can see that these two players will have an incentive to form a link.

Then, they will form a link. So I will come to the star network, let's say, let's call it that. I'm just naming them. But if I'm in the star network, you can see this player; if he decides to snap this link, he's going to get 11.

So, he will do that. So I will end up in, let's say, the square network or something similar. Okay, so tick, hash, star, and square. I will keep moving between these networks, and once I return to the square network, I will go back to the tick network. I will continue cycling between these four networks without settling anywhere. Therefore, none of these four networks is pairwise stable, and a pairwise stable network does not exist here.

Okay, next: an efficient network. Now, what is an efficient network? Given a set of players and a profile of utility functions $u_1, \ldots, u_n$, along with a network G that belongs to the set of all possible networks, a particular network is called efficient if the sum of the utilities of all the players is greatest for that network. The sum of the utilities of all the players in the network is greater than the sum of the utilities of all the players in any other network belonging to G_n. So, this is the network that maximizes the total payoff for all the players in the network. And, of course, an efficient network will exist.

Because what am I doing here? This is the network that provides the highest total utility. However, how many networks are possible? If I have capital N players, then the total number of possible networks will also be finite, given that I have a finite number of players. Therefore, if I have a finite number of players, I also have a finite number of networks, which allows me to compare them and identify the network with the highest total utility. I refer to that network as the efficient network. There can be more than one efficient network, but there will be at least one efficient network. Then, there is the Pareto efficient network. Now, a Pareto efficient network is defined as a specific network G that will be referred to as Pareto efficient.

Again, we have a set of players and a set of utility functions. A particular network G is called Pareto efficient if there does not exist any other network G' such that $u_i(G')$ is greater than or equal to $u_i(G)$ for all i, with a strict inequality existing for at least one player. So, what does this mean? The strict inequality holds for at least one player.

A particular network is called Pareto efficient if no other network exists such that all players are better off in the other network G' than in G, or if all players are as well off in G' as in G, with one player being better off. So, if such a G' does not exist, then G will be called a Pareto-efficient network. I will repeat it once more. A network G is called Pareto efficient if there does not exist any other network G' such that all players are at least as well off in G' as in G, with one player being strictly better off. Great! Now, is there a connection between efficiency and Pareto efficiency? The answer is that efficiency implies Pareto efficiency. That is, if a network is efficient, then it is necessarily Pareto efficient. What's the logic? The logic is very simple and straightforward.

Let's say network G is efficient; this means... G is defined as the value that maximizes the sum of the utility for all the players. The summation of the utility under G is greater than the summation of the utility under G'. This means that if I deviate from G to G', it cannot happen that all players are better off. In other words, if I deviate from G to G', it is necessary that at least one player will be worse off. At least one player will be worse off, or perhaps all players will be equally well off.

In that case, both G and G' are efficient. If I deviate from G to G', and G is a uniquely efficient network, then at least one player will be worse off. It cannot happen that all players are equally well off if at least one player is strictly better off by deviating from G to G'. That's an impossibility, which, in other words, implies that G is also Pareto-efficient. So, if a network is efficient, it is also a Pareto-efficient network. Great! Let's take a quick example to understand it. Again, let's set up a four-player network. You can see that the red network, which is Pareto efficient, is indeed efficient.

Why? Look at the total utility: $3+3+3+3=12$. So, that is 12. No other network has a total utility of 12.

So, this network is both efficient and Pareto-efficient. It is efficient, and as we saw in the previous slide, it is also Pareto efficient. If it is Pareto efficient, it need not be efficient, as we can see.

Look at the green network. The green network is Pareto efficient. Deviating from the green network, we cannot end up in a situation where all the players are equally well off, with at least one player being strictly better off. Therefore, the green network is Pareto efficient. If we move from the green network to the blue network, players one and two are worse off.

In the blue network, they have 2.5, whereas they had 3.25 in the green network. If I move from the green network to the red network, again, both 1 and 2 are worse off. They had 3.25 in the green network.

In the red network, they have 3, which is less than 3.25. Fine. So, we can see that the green network is indeed Pareto-efficient.

But is the green network or the red network pairwise stable? The only pairwise stable network is the blue network. See, no two players or any player has an incentive to snap an edge or a tie, and no two players have an incentive to form a new edge. By the way, the efficient network need not be pairwise stable. Look at the red network; it is efficient, but players 1 and 2 have an incentive to form a new edge between them.

The red network will transition to the green network because the red network is not pairwise stable. Here, two players, 1 and 2, will benefit by forming an edge between them.

They will end up getting 3.25 and 3.25 instead of 3 and 3. Similarly, the green network is also not pairwise stable. Because 3 and 4 will have an incentive to form an edge between them, they had 2 and 2.

By forming an edge, they will have 2.5. 2.5, okay, great. So, we see that efficiency and pairwise stability can be at odds. So, here the pairwise stable network is not efficient. The efficient network is not pairwise stable. So we can be torn between the two. Great! So now we look at a particular kind of network that is characterized by something called a distance-based utility function.

Then we will try to see. Whether they constitute what are the efficient and pairwise stable networks in this situation, when we have distance-based utility functions. Let's go for it. So, what is the basic idea behind distance-based utility functions? The basic idea is that a player derives utility from direct and indirect connections, right? The utility deteriorates with the distance between the individuals.

So let's say B is a mapping from the finite set $\{1,\ldots,n\} \to R$, which is a payoff function. What is B. L_{ij}? Denotes the benefit I derive as a result of connecting with J, where L_{ij} is the minimum path distance between i and j. That's how B is defined; B is a function.

Also, there is a cost to forming a link; that particular cost is denoted by C. So, what is the utility of player I? Given a network G, the utility of player I is given by the following. What utility does he derive from connecting with player J? It is $B(L_{ij})$. So, what is the total benefit he is getting from all his connections? Well, it is simply given by the summation of J belonging to N, where $j \neq i$, of $B(L_{ij})$. Okay? Great. How many links is he forming, minus? Well, it is equivalent to his degree. So, D_i is the degree of player i, which is the number of edges associated with player i. This represents the number of edges in which he is involved, specifically in the formation of D_i edges. So, $C(D_i)$ is the total cost incurred by player i, and this represents the total benefit.

This is the total cost: benefit minus cost. That's the utility of player i in a network G, okay? Obviously, $B_k > B_{k+1}$. What is B? B is the benefit, and what is taken as an argument? It takes in L_{ij}, which is the shortest distance between i and j; the shorter the shortest distance, the greater the utility.

That goes without saying. Great. So, if this is the situation, let's look at what the efficient networks are. Given the distance-based utility function, the unique efficient network structure is represented by this. If C, the cost of forming a link, is less than $B_1 - B_2$, where B_1 is the benefit a player receives from another player when the shortest distance between them is 1, then.

.. B_2 is the benefit a player receives from connecting with another player when the shortest distance between the two players is two. So, if C is less than $B_1 - B_2$, then.

.. Then a complete network is the most efficient network. What is a complete network? Where everyone is connected. If C lies in this interval, then a star encompassing all nodes is an efficient network. Also, if C is greater than this, then an empty network is an efficient network.

Let's try to prove this.

Let's look at number one. Let's try to prove number one. So in an efficient network, what am I trying to do? I'm trying to maximize the total utility of all the players involved in the network. So, if C is less than $B_1 - B_2$, what does it mean? So this implies that $B_1 - C$ is greater than B_2. Now, what does this imply? If you have any two players who are not connected, what is the maximum utility a player can derive? So, if I have a player i and a player j. If they are not connected, what is the maximum utility i can derive from j? The answer is B_2, right? That's the maximum utility i can derive from j without being directly connected. Now, if he directly connects, then what will be the payoff from this association between i and j? What is the maximum benefit i can derive? As in, by being associated with j directly, the answer is $B_1 - C$.

Here, B_1 is the benefit, and C is the cost of forming the link between i and j. If this is the benefit of forming a direct link with j, this is the maximum benefit I can achieve through an indirect association with j. Now, if $B_1 - C$ is greater than B_2, Then it is always rational for player i and player j to form a direct link between them, which means that any pair of players you choose will have an incentive to form a direct link between themselves. In other words, we will have a complete network. Not only is it an incentive; maybe I placed that incorrectly.

What is my objective here? My objective is to maximize total welfare. I will just correct myself on this. If B_2 is the total, then what is the total benefit? j and i both derive $2B_2$. i get B_2, and j gets $2B_2$ if they are not connected.

That is the maximum indirect benefit of the network. And if there is a link between them, what is the maximum benefit? Both of them will receive $B_1 - C$ so the total is $2B_1 - C$. That's the total utility of i and j. If $B_1 - C$ is greater than B_2, then $2B_1 - C$ is greater than $2B_2$. In other words, adding the edge increases the total utility of both players. So, by that logic, we'll have a complete network, meaning we'll have an edge existing between any pair of nodes or any pair of players in the network.

Okay, let's move on to number two. What is number two saying? Number two states that if C lies in this interval, then the efficient network is a star. A star looks like this. With one central player and other peripheral players.

So, this is my central player. The other players are my secondary players. So this kind of structure is the uniquely efficient network. Why?

Let's see. Let's try to prove this. If I have one player at the center and a number of peripheral players, let's say there are n players. What is the total utility of this network? Let us try to understand it. What is the purpose of the central guy? Well, it is. So, you start with the total utility if there is a star network. So, what is u star? What is the utility of the central player? Well, he has formed $n-1$ links, and from each association—let's say with this association—what is his benefit? What is the benefit he is deriving from this association, the one I have marked—this one, where it is b1 minus c? That's the benefit the central player is getting from this particular peripheral player, and there are $n-1$ such peripheral players. So, this is the total payoff for the central player. What about the peripheral players? What about this player? Let's say, how much is he getting? Well, he's directly connected to the central player, so he's getting B_1.

He is connected indirectly to $n-2$ other peripheral players via the central player, and the distance between any two peripheral players is B_2. So, it is simply $n-2B_2$. So, this is the payoff that this particular peripheral player is receiving, and how many such peripheral players are there? $n-1$. So, this is the total payoff of all the players combined in a star network. Let us simplify this a little. So, if I take the $n-1$ out, what do we end up with? $2B_1$ plus, oh, by the way, there is also a $-C$. Remember, because every peripheral player also forms a link, every peripheral player also forms a link, so that's $-C$. Therefore, this is $2B_1+n-2B_2-2C$.

That's what we are left with. If I just take the 2 outside, I'm left with B_1+n-2 divided by $2B_2-C$ is greater. So, this is applicable if the network is a star network. So, I have a net with N players. If it is a star network, this is the payoff I end up getting. Now, why is this payoff the maximum I can receive? Can't there be another network where the total utility is larger than this? Let us create another network.

Let us say I have another network with m links, where m is greater than n minus 1, I will explain where we are going. Let's say I take another network with m links; we'll call this network G. Now, what is UG? Remember, there are m links, so there are m direct connections. If it's a direct connection, every player connected via this link will receive B_1-C.

Therefore, what is the total utility due to these links? Well, it is simply... To calculate B_1-C, there are m such links, so we multiply by m. Therefore, this is the total utility from direct links, plus the utility from indirect links. Now, how many pairs of players are there who are not directly connected? There are m links and n players. So, the total number of possible pairs of players is $\frac{n(n-1)}{2}$.

So, $n-C-2-m$ players are those who are not directly connected by an edge. So, if they are even indirectly connected, what is the maximum benefit due to the indirect connection? It is B_2. So, this is B_2. Okay, so this is the maximum benefit that I can get, so UG will be greater than or equal to this.

Okay, great. Now I'll have to prove what u_G-u^* is. By the way, let's proceed with this: what was u^*? u^* was this, and this is u_G, so what is u_G-u^*? I am sorry, but this will be less than that, right? So, this is the maximum indirect benefit that is possible. So, let me do it this way. So, this will be the benefit from direct links; this is the maximum benefit from indirect links.

So, u_G will be less than or equal to this. That is the maximum; the right-hand side represents the maximum utility that a network with m links can have. So, what is u_G-u^*? So this will be less than or equal to 2m; this will be $2B_2$. Let me explain this a little more, as I think I messed up a bit here. Let us say that if there are two players who are directly connected, each player receives a payoff of B_1-C.

Due to direct links, it is $2B_1-C$ for each link, and there are m links. So this is $2B_1-C$.

For any pair that is indirectly connected and at a minimum distance of 2, what is the payoff? To each, it is B_2. So when they put it together, it is $2B_2$. So, the total payoff is ${}^nC_2 - m$; that's the total number of pairs that are not directly connected.

So that is multiplied by $2B_2$. So there will be a 2 here, which I initially forgot. I apologize for that. So this is ${}^nC_2$; ${}^nC_2$ is $\frac{n(n-1)}{2} - m$.

-u^*. What was u^*? Let us take a look; it was $(2n-1)B_1 + (n-1)(n-2)B_2 - 2 - 1$ are okay. Now let us try to simplify this a little bit more and see what we end up getting. So, this is equal to what we end up getting if we take $B=1$. So, we have 2m here. If we take b² common, what do we get? So, we have $n(n-1) - 2m$, and here I have $n-1$ and $n-2$.

And if I take c, then I have $2n-1-2n$. That's what we end up with. Now let's try to see it. So, if I take 2 out, this is $m-n-1$. What do we have here? If I take $n-1$ common, then I have $2n-1-2m$. Again, here I have $2cn-1-m$. So, what do we have here? So, if I have $2B_1 - 2B_2$ and we see that f, I take $m-n-1$ out; what does it lead to? $B_1 - B_2 - C$. In the proposition, we have seen that if C is greater than $B_1 - B_2$, then C is greater than $B_1 - B_2$ here. So, this and this are necessarily negative since C is greater than $B_1 - B_2$. What about this? I have taken m to be greater than $n-1$, so this is necessarily positive. Therefore, the entire expression becomes less than or equal to 0, which means u_G is less than or equal to u^*.

If m is greater than or equal to $n-1$. This is what the proof becomes. It's fairly simple, right? By the way, what if m is less than n minus one? What happens next? Well, let's try to see it. Let's say we have two stars that are mutually exclusive. So, they are two separate components. So, let's say the first star has k_1 and the second star has k_2.

Then, what is the total utility of this network, G? So we have seen that for a particular star, it is given by this. So, $\frac{k_1 - 1}{2} + \ldots$ And if there is a unified star, then u^*, well, this will be $, B_1 + \frac{k_1 + k_2 - 2}{2} \times B_2 - C + \ldots$.

.. Ah, I am sorry, yeah, this. It can be easily proved that u^* is greater than u_G. This means that if I have two disconnected components, even if both are stars, remember that in any component, the maximum utility occurs when it is a star. Therefore, if we have two disconnected components, the unified star will still have a higher utility. Now think about it: if there are n players and I have fewer than n minus one links, then there will be at least one player left out who will be a solitary player. This, in turn, means that for m less than n minus one, I am in this two-component paradigm, and I can prove that the star yields the best; a unified star yields more than two components. So, in other words, for m less than n minus 1 links, this implies two components because it necessarily implies that it cannot be a completely connected graph.

Okay, then the unified star necessarily leads to a higher utility, so it is definitely the uniquely efficient network for $m<n$ and $m \leq n-1$. I have provided a rigorous proof for $m>n-1$. This is the logic. And what happens to the last part of the theorem? What occurs if C is greater than $B_1+\frac{n-2}{2}B_2$? So, you see that the highest utility a player can achieve, the best network is when it is a star.

Here's a corrected version of your sentence: And this is the payoff. In other words, I can say that for all g belonging to G_n, this utility, which we obtained as $2n-1$, is less than C . If C is greater than this, then it turns out that the total utility will be negative, meaning u_G will be negative, or in other words, u_G is less than or equal to zero. However, Why will I have a negative payoff in the network? If I want to maximize the total utility of all the players and I see that I'm getting a negative payoff, I can always choose a network where there is zero payoff, which is the empty network.

So, if C is strictly greater than this, I will end up with negative total utility networks. So my efficient network will become an empty network. So, this brings us to the last statement of the theorem: the empty network is my unique efficient network if C is greater than that threshold. Great, so we have looked into the efficient networks for when we have distance-based utility functions. In the next lecture, we will try to determine if we have distance-based utility functions and what the pairwise stable networks are. Is there a conflict between pairwise stability and efficiency? We'll also examine another type of network model, called the co-authorship network model. But that will be in the next lecture. See you at the next lecture. Thank you.