Conditionals in Game Theory

Ilaria Canavotto, University of Maryland
Eric Pacuit, University of Maryland

Lecture 3

ESSLLI 2022

V. Capraro and J. Halpern. *Translucent Players: Explaining Cooperative Behavior in Social Dilemmas*. Proceedings of the 15th conference on Theoretical Aspects of Rationality and Knowledge, 2015.

Prisoner's Dilemma

		Bob	
		c	d
Ann	c	3,3	0,4
	d	4,0	1,1

Social Dilemmas

1. There is a unique Nash equilibrium s^N, which is a pure strategy profile;
2. There is a unique welfare-maximizing profile s^W, again a pure strategy profile, such that each player's utility if s^W is played is higher than his utility if s^N is played.

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in dollars from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.
4. The person that wrote the smaller number will receive that amount plus $2 (as a reward), and the person that wrote the larger number will receive the smaller number minus $2 (as a punishment).

Suppose that you are randomly paired with another person from class. What number would you write down?

Expected Utility, Best Response

Suppose that $G = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ is a game in strategic form.
For $a \in S_i$ and $p \in \Delta(S_{-i})$, a is a best response to p when: for all $a' \in S_i$,

$$\sum_{s_{-i} \in S_{-i}} p_i(s_{-i}) u_i(a, s_{-i}) \geqslant \sum_{s_{-i} \in S_{-i}} p_i(s_{-i}) u_i(a', s_{-i})$$

Expected Utility, Best Response

Suppose that $G = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ is a game in strategic form.
For $a \in S_i$ and $p \in \Delta(S_{-i})$, a is a best response to p when: for all $a' \in S_i$,

$$\sum_{s_{-i} \in S_{-i}} p_i(s_{-i}) u_i(a, s_{-i}) \geqslant \sum_{s_{-i} \in S_{-i}} p_i(s_{-i}) u_i(a', s_{-i})$$

Implicitly assumes that i's beliefs about what other agents are doing do not change if i switches from s_i, the strategy he was *intending* to play, to a different strategy.

$p_i^{s_i, s_i'}$: i's beliefs if he intends to play s_i but instead deviates to s_i'

$p_i^{s_i,s_i'}$: i's beliefs if he intends to play s_i but instead deviates to s_i'

Strategy $a \in S_i$ is a best response for i with respect to the beliefs $\{p_i^{a,a'} : a' \in S_i\}$ if for all strategies $a' \in S_i$

$$\sum_{s_{-i} \in S_{-i}} p_i^{a,a}(s_{-i}) u_i(a, s_{-i}) \geqslant \sum_{s_{-i} \in S_{-i}} p_i^{a,a'}(s_{-i}) u_i(a', s_{-i})$$

A player is **translucently rational**— if he best responds to his beliefs.

Translucency will be used to determine $p_i^{a,a'}$:

Suppose that G is a two-player game, player 1 believes that, if he were to switch from a to a', this would be detected by player 2 with probability α, and if player 2 did detect the switch, then player 2 would switch to b.

Translucency will be used to determine $p_i^{a,a'}$:

Suppose that G is a two-player game, player 1 believes that, if he were to switch from a to a', this would be detected by player 2 with probability α, and if player 2 did detect the switch, then player 2 would switch to b.

Then $p_i^{a,a'}$ is $(1-\alpha)p_i^{a,a} + \alpha p'$, where p' assigns probability 1 to b: that is, player 1 believes that with probability $1-a$, player 2 continues to do what he would have done all along (as described by $p_i^{a,a}$) and with probability α, player 2 switches to b.

Explaining Cooperation

Say that an player i has type (α, β, C) if i intends to cooperate and believes that

1. if he deviates from that, then each other agent will independently realize this with probability α;
2. if a player j realizes that i is not going to cooperate, then j will defect; and
3. all other players will either cooperate or defect, and they will cooperate with probability β.

	C	D
C	$b-c,\ b-c$	$-c,\ b$
D	$b,\ -c$	$0,\ 0$

Proposition In the Prisoner's Dilemma, it is translucently rational for a player of type (α, β, C) to cooperate if and only if $\alpha\beta b \geqslant c$.

J. Halpern and R. Pass. *Game theory with translucent players*. International Journal of Game Theory, 47:3, pp. 949 - 976, 2018.

Given a strategic-form game $G = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$, a model of G is a triple

$$\langle W, f, (P_i)_{i \in N}, \sigma \rangle$$

where W is a non-empty set of states, $\sigma : W \to \Pi_{i \in N} S_i$, and:

$$\text{For each } i \in N,\ P_i : W \to \Delta(W).$$

- For all $w \in W$, $P_i(w)([\sigma_i(w)]) = 1$.
- For all $w \in W$, $P_i(w)(\{v \mid P_i(v) = P_i(w)\}) = 1$.

Given a strategic-form game $G = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$, a model of G is a triple

$$\langle W, f, (P_i)_{i \in N}, \sigma \rangle$$

where W is a non-empty set of states, $\sigma : W \to \Pi_{i \in N} S_i$, and:

$$\text{For each } i \in N,\ P_i : W \to \Delta(W).$$

- For all $w \in W$, $P_i(w)([\sigma_i(w)]) = 1$.
- For all $w \in W$, $P_i(w)(\{v \mid P_i(v) = P_i(w)\}) = 1$.
- f associates with each state w, player i and strategy a a state $f(w, i, a)$ where player i plays a. If $f(w, i, a) = w'$, then
  - $\sigma_i(w') = a$.
  - If $\sigma_i(w) = a$, then $w' = w$.

$$P^c_{i,a}(w)(w') = \sum_{\{w'' \in W \ \mid \ f(w'', i, a) = w'\}} P_i(w)(w'')$$

$$P^c_{i,a}(w)(w') = \sum_{\{w''\in W \ \mid \ f(w'',i,a)=w'\}} P_i(w)(w'')$$

- $P^c_{i,a}$ is i's counterfactual beliefs at state w: what i believes would happen if she switched to s at w
- $P_{i,a}(w)^c([a]) = 1$
- It may *not* be the case that $P^c_{i,a}(w)([P^c_{i,a}(w), i]) = 1$: players do not in general know their counterfactual beliefs in state w
- A model is a *strongly appropriate counterfactual structure* if at every state w, every player i knows his counterfactual beliefs.

$$B_i(E) = \{w \mid P_i(w)(E) = 1\}$$

$$B_i^*(E) = \{w \mid \text{for all } s' \in S_i, P^c_{i,s'}(w)(E) = 1\}$$

Characterize solution concepts in terms of the players beliefs, common beliefs, counterfactual beliefs and common counterfactual beliefs.