CS-7642: week 12

Correlated Equilibria and CE-Q

Correlated Equilibrium pt. i

Informally: a **randomized** assignment of potentially **correlated** action recommendations to agents, such that **nobody wants to deviate**. [2]

Correlated? In the prototypical “traffic intersection” example, imagine how the traffic light works...you’re told to stop, and so you infer the other car is told to what?...and so you do...what?

	go	wait
go	$-10, -10$	$1, 0$
wait	$0, 1$	$-1, -1$

Correlated Equilibrium pt. ii

Policy π is a **correlated equilibrium** if for any agent *i*, if all the other agents follow the advice of the policy, agent *i* maximizes its expected utility by **also following** the advice of the policy.

We want the expected payoff of current action to be at least as great as when agent *i* **alone** switches to some different action.

for all $i \in N$ and for all $a_i, a_i' \in A_i$,

$$\sum_{a_{-i} \in A_{-i}(s)} \pi_s(a) Q_i(s, a) \geq \sum_{a_{-i} \in A_{-i}(s)} \pi_s(a) Q_i(s, (a_{-i}, a_i'))$$

Piazza Question:

“[Is it] Possible to work through an example of translating a game to a LP? I worked my way through RPS, but still struggling a bit with writing the LPs. **Also anything describing how we leverage LPs in P3 would be great too.**”

What is the role of Linear Programming in CE-Q?

In the multiagent Q-learning algorithm, as agents take each joint action, we execute the following Q-learning update (for each agent in 1...n):

3. for $i = 1$ to n

   (a) $V_i(s') = f_i(Q_1(s'), \ldots, Q_n(s'))$

   (b) $Q_i(s, \vec{a}) = (1 - \alpha)Q_i(s, \vec{a}) + \alpha[(1 - \gamma)R_i + \gamma V_i(s')]$

In order to compute V(s) we will need to use LP (*f* contains your LP procedure).

Steps to Compute $V_i(s)$ in CE-Q; pt. i

Prerequisite: decide on the “flavor” of CE-Q you will use which determines the *objective* function you will use in your LP formulation (utilitarian, egalitarian, republican, libertarian).

Q: What are your unknowns? (the things you’re solving for in the LP)

A: We’re looking for an equilibrium *policy* in *joint action space*. In practice this is a probability distribution over the possible joint actions. So each possible *joint action* would have a probability associated with it.

Steps to Compute $V_i(s)$ in CE-Q; pt. ii

Q: Given our unknowns, what are our constraints (ie what *conditions* must our unknowns pi(s) satisfy)?

A1: ("probability constraints") since each unknown represents a *probability* that some joint action will be taken, what is the *sum* of these unknowns?

A2: ("rationality constraints i.") since each unknown represents a *probability* can we assert something about its *sign*?

A3: cont ...

Steps to Compute $V_i(s)$ in CE-Q; pt. iii

(Rationality constraints continued...) Remember the defining inequality which expresses the condition that (mixed) policies must satisfy to be a CE (translated into CE-*Q* friendly notation):

$$\sum_{a_{-i} \in A_{-i}(s)} \pi_s(a) Q_i(s, a) \geq \sum_{a_{-i} \in A_{-i}(s)} \pi_s(a) Q_i(s, (a_{-i}, a_i'))$$

This expression defines a set of inequalities over our unknowns and provides the remainder of our LP constraints.

Steps to Compute $V_i(s)$ in CE-Q; pt. iv

The final requirement for setting up your CE selection as an LP is to choose an *objective function*.

This is the thing that we are seeking to maximize, minimize, etc.

In your case you will be choosing one of the four objective functions defined Greenwald and Hall.

References

[1] Greenwald, Hall (2003)

[2] Jackson, Leighton-Brown, Shoham http://game-theory-class.org/

[3] https://stanford.edu/~saberi/lecture6.pdf

[4] https://www.cs.cornell.edu/courses/cs684/2004sp/feb20.pdf