Last Time

Sampling and Inference in Bayes Nets

↑

Input: BN (G, θ)

Output: samples for every variable

Input: BN, values of some

Output: Distributions of others

This Time

Utility Theory

Simple Games

Preferences

$A \succ B$

$A \sim B$

$A \succeq B$

Lottery $[S_1 : p_1, S_2 : p_2, \ldots, S_n : p_n]$

Assumptions

Completeness: one of $A \succ B$, $B \succ A$, or $A \sim B$ holds

Transitivity: if $A \succeq B$ and $B \succeq C$ then $A \succeq C$

Continuity: if $A \succeq C \succeq B$ $\exists\, p \in [0,1]$ s.t.
$[A : p, B : 1-p] \sim C$

Indepence: if $A \succ B$ then for any C and p
$[A : p, C : 1-p] \succeq [B : p, C : 1-p]$

If assumptions

$\exists U$ s.t. $U(A) > U(B)$ iff $A \succ B$

$U(A) = U(B)$ iff $A \sim B$

Preference Elicitation

$U(\underline{S}) = 0$ $U(\overline{S}) = 1$

find p such that $[\overline{S} : p, \underline{S} : 1-p] \sim S$

$U(S) = p$

Pitfalls

Humans are often not rational
- Cognitive Bias
  - Certainty Effect
  - Framing Effect
- Quantitative measures ($) don't directly work

$100 |
$50 ||

Indifferent Risk Neutral
Prefer $100 50% Risk Seeking
Prefer $50 100% Risk Averse

Simple Games

"Normal Form" "Matrix"

Joint Action Space

$\bar{a} \in A$

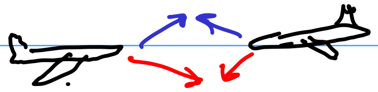
$A = A^1 \times A^2 \times \dots A^k$ $\quad I = \{1, \dots k\}$

Joint Reward Function

$R(\bar{a}) = (R^1(\bar{a}), R^2(\bar{a}) \dots R^k(\bar{a}))$

2

		climb	descend
1	climb	(-5,-5)	(-1,0)
	descend	(0,-1)	(-4,-4)



Joint policy $\pi^i(a)$ probability that agent i plays action a
strategy profile

Pure Strategy if $\pi^i(a) \in \{0,1\}$
Mixed Strategy otherwise

Breakout Room

What is the best policy?

Group2
1. Descend
2. Ascend

Correlated

Group 1: Communication
1. East always ascend

Nash Equilibrium

Response

$-i$ shorthand for $(1, \dots, i-1, i+1, \dots t)$ $I \setminus i$

$\bar{a} = (a^i, a^{-i})$

$R(\bar{a}) = R(a^i, a^{-i})$

$\bar{\pi} = (\pi^i, \pi^{-i})$

Best Response

π^{-i}

π^i s.t. $U^i(\pi^i, \pi^{-i}) \geq U^i(\pi^{i'}, \pi^{-i}) \quad \forall \pi^{i'}$

Nash Equilibrium

A N.E. is a joint policy π in which all agents follow a best response

OR

A N.E. is a joint policy in which no agent has an incentive to unilaterally swith their policy.

(-5,-5)	(-1,0)
(0,-1)	(-4,4)

$\pi^1(\text{climb}) = 1$
$\pi^2(\text{climb}) = 0$

$\pi^1(\text{climb}) = 0$
$\pi^2(\text{dimb}) = 1$

All finite-action games have atleast 1 N.E.

R.P.S.

		R	P	S
	R	(0,0)	(-1,1)	(1,-1)
1	P	(1,-1)	(0,0)	(-1,1)
	S	(-1,1)	(1,-1)	(0,0)

2

Uniformly Random $\pi^i(R) = \pi^i(P) = \pi^i(S) = 1/3$

defender

		climb	descend
attacker	climb	(3,-5) →	(-1,0)
	descend	(0,-1)	(4,-4)

$$\pi^* = \begin{cases} \pi^d(\text{climb}) = 0.5 \\ \pi^a(\text{descend}) = 1.0 \end{cases}$$

---

finding NE is PPAD-complete : No known polynomial-time solutions

$$\underset{\pi, U}{\text{minimize}} \sum_i \left(U^i - U^i(\pi)\right) \qquad 0 \text{ if NE}$$

$$\text{subject to} \quad U^i \geq U^i(a^i, \pi^{-i}) \quad \forall\, i, a^i$$

$$\sum_{a^i} \pi^i(a^i) = 1 \quad \forall\, i$$

$$\pi^i(a^i) \geq 0 \quad \forall\, i, a^i$$

Correlated Equilibrium

For N.E. $\underline{\pi(\bar{a}) = \prod \pi^i(a^i)}$

$$\sum_{a^{-i}} R^i(a^i, a^{-i})\, \pi(a^i, a^{-i}) \geq \sum_{a^{i'}} R^i(a^{i'}, a^{-i})\, \pi(a^{i'}, a^{-i})$$

Finding a correlated equilibrium involves solving an L.P.

Every N.E. is a C.E., but not every C.E. is a N.E.

Things that might converge

- Iterated best response : randomly choose 1 player
  solve for best response
  cycle through players
- Fictitios Play : Use M.L. estimates of othe agents policies, use to choose policy that you play

Humans

Level k response

- Level-0 players : random policy
- Level-k players calculate best response to level k-1 policy

Combine with softmax

$50 |
$2 ||