Utility Theorem

Let O be the set of outcomes

$O = Z \cup \Delta(O)$

- $Z \rightarrow$ some set of outcomes
- $\Delta(O) \rightarrow$ set of lotteries over finite subsets of X

$[p_1 : x_1, \dots, p_K : x_K]$ → List / dist. of pairs of prob. and outcome with that prob.

with $x_1, \dots, x_K \in X$ and $\sum_{j=1}^{k} p_j = 1$

A preference relation compares the relatively desirability of outcome

- $o_1 \succeq o_2 \rightarrow$ agent weakly prefers o_1 to o_2
- $o_1 \succ o_2 \rightarrow$ agent strictly prefers o_1 to o_2
- $o_1 \sim o_2 \rightarrow$ agent is indifferent b/w o_1 and o_2

A utility function is a function $u: O \rightarrow \mathbb{R}$

- It represents a preference relation $\succeq$ iff

1) $\forall o_1, o_2 \in O : o_1 \succeq o_2 \iff u(o_1) \geq u(o_2)$, and
2) $\forall [p_1 : o_1, \dots, p_k : o_k] \in O : u([p_1 : \ , \dots, p_k : o_k]) = \sum_{j=1}^{k} p_j u(o_j)$

→ If preference relation $\succeq$ satisfies the axioms Completeness, Transitivity, Monotonicity, Substitutability, Decomposability, and Continuity.
Then there exists a utility fn u that represents $\succeq$

Axioms

- $\forall o_1, o_2 \in O : (o_1 \succ o_2) \vee (o_1 \prec o_2) \vee (o_1 \sim o_2)$ → Completeness
- $\forall o_1, o_2, o_3 \in O : (o_1 \succeq o_2) \wedge (o_2 \succeq o_3) \Rightarrow o_1 \succeq o_3$ → Transitivity
- $(o_1 \succ o_2) \Rightarrow [p : o_1, (1-p) : o_2] \succ [q : o_1, (1-q) : o_2]$ → Monotonicity
  ↳ prefer 90% chance of getting \$1000 to 50% chance of getting \$1000
- $p + \sum_{j=3}^{k} p_j = 1$, if $o_1 \sim o_2$
  $[p : o_1, p_3 : o_3, \dots, p_k : o_k] \sim [p : o_2, p_3 : o_3, \dots, p_k : o_k]$ → substitutability
  ↳ If I like apples & bananas equally, then shouldn't care about 30% chance of getting apples or bananas
- $(\forall o \in O : P_{\ell_1}(o) = P_{\ell_2}(o)) \Rightarrow \ell_1 \sim \ell_2$ → Decomposability (No fun in Gambling)
  ↳ prob. that outcome o is selected
  $\ell_1 = [0.5 : [0.5 : o_1, 0.5 : o_2], 0.5 : o_3]$, $\ell_2 = [0.25 : o_1, 0.25 : o_2, 0.5 : o_3]$
  $\ell_1 \sim \ell_2$ ∵
  $P_{\ell_1}(o_1) = 0.5 \times 0.5 = 0.25 = P_{\ell_2}(o_1)$
- $o_1 \succ o_2 \succ o_3 \Rightarrow \exists p \in [0,1] : o_2 \sim [p : o_1, (1-p) : o_3]$ → Continuity

Proof Sketch (Von Neumann & Morgenstern)

- Choose o^+, o^- such that $o^- \preceq o \preceq o^+$ for all o
- Construct $u(o) = p$ such that $o \sim [p : o^+, (1-p) : o^-]$
- Substitutability lets us replace everything with these canonical lotteries, Monotonicity lets us assert the ordering b/w them

→ For a given set of preferences, the utility function is not uniquely defined.

Comparisons of expected values are invariant to positive affine transformations

$$X \succeq Y \Leftrightarrow E[u(X)] \geqslant E[u(Y)]$$
$$\Leftrightarrow cE[u(X)] \geqslant cE[u(Y)]$$
$$\Leftrightarrow cE[u(X)]+b \geqslant cE[u(Y)]+b$$
$$\Leftrightarrow E[\underline{cu(X)+b}] \geqslant E[\underline{cu(Y)+b}]$$

for all $b \in \mathbb{R}$ and $c>0$

Von Neumann & Morgenstern → risk neutral

⇒ <u>Utility theory</u> → requires you to know what the probabilities are.

↳ assumes known, objective probabilities

<u>Other theorems state that</u> ⇒ rational agents must (a) have prob. beliefs (b) update those beliefs as if by conditioning c) max. the expected value w.r.t them.

Summary

- Utility theory proves that agents whose preferences obey certain simple axioms about preferences over lotteries must act as if they were maximizing the expected value of scalar func.
- Rational agents → who satisfy the axioms.

Game Representations

Finite, n-person game: $\langle N, A, u \rangle$ → Game

N → finite set of n players

$A = \langle A_1, \dots, A_n \rangle$ is a set of action sets for each player i

- $a \in A$ is an action profile

$u = \langle u_1, \dots, u_n \rangle$, a utility function for each player, where $u_i : A \to \mathbb{R}$

- <u>Normal form</u> → 2 player game as a matrix
  ↳ table rep. of simultaneous move game

	C	D
C	a,a	b,c
D	c,b	d,d

$c>a>d>b$ → prisoner's dilemma where this is satisfied

Players have exactly opposed interests
→ Exactly two players
→ for all action profiles $a \in A$, $u_1(a) + u_2(a) = c$ → constant sum game
 – Special case: zero sum
→ Thus only need to store utility fn for one player

} Games of pure competition

<u>Matching Pennies</u>

↳ zero sum game

	H	T
H	1,−1	−1,1
T	−1,1	1,−1

<u>Rock Paper Scissors</u> → zero sum simultaneous move game

	R	P	S
R	0,0	−1,1	1,−1
P	1,−1	0,0	−1,1
S	−1,1	1,−1	0,0

Game of Cooperation

Players have exactly same interests

- no conflict: all players want the same things
  $\forall a \in A, \forall i,j, \; u_i(a) = u_j(a)$
- single payoff per cell

What side of road to drive on?

	L	R
L	1, 1	0, 0
R	0, 0	1, 1

↳ cannonical game of cooperation

- Most interesting games combine elements of cooperation and competition

Battle of Sexes

	B	F
B	2, 1	0, 0
F	0, 0	1, 2

→ General sum game

- Repeated Games
  → Play the same game over and over again

- Perfect-Information Extensive Form Game
  → Game unfolds over time
  → Player can see each other's moves
  → Can be written as tree, leaves labelled as payoffs

- Stochastic Game
  → Combines perfect-information extensive form game with repeated game
  → Multiplayer generalization of a Markov Decision Process (MDP)
  state: which game is being played
  actions: set of alternatives in that game
  reward: payoffs in that game
  transition ftn: mapping from players' actions to next state

- Imperfect-Information Extensive form Game
  ↳ generalizes perfect-information extensive-form game games by allowing for imperfect observation of previous player's moves
  → Some actions might be observed perfectly

- Stackelberg Games → player commits to drawing action from a prob. distribution
  ↳ special case of imperfect-info games
  → One player commits to a strategy which is observed by the second player
    → If strategy is randomized, the second player can't see random draws
  → Same game played over & over again, the second player can see the first player's actions and figure out her strategy

- Bayesian Games
  → Uncertainity about payoffs (either one's own or other's)
  prior → over which game is played

Nash Equilibrium

Agents simultaneously make a decision.
Then they receive an outcome depending on profile of actions } Normal form game recap

$G = (N, A, u)$ → utility profile
↳ set of n players; → set of action profile per agent

Optimal Decisions in Games

In single-agent decision theory, the key notion is optimal decision:
→ the action that maximizes the agent's expected utility.

$$a^* = \underset{a \in A}{\operatorname{argmax}} \; E[u(a)]$$

In multi-agent, the optimality notion is ill-defined

$$a_i^* = \underset{a_i \in A_i}{\operatorname{argmax}} \; E[u_i(a_i, a_{-i})]$$

→ issue because same for other agents

- The best strategy depends on actions of others.

Solution Concepts

→ From an outside perspective, some outcomes of a game be considered better
→ Solution concepts pick sets of things that are interesting/good/desirable.
→ we cannot compare utilities of the agents to each other, because of affine invariance.

Pareto Optimality → weak

Suppose outcome o is atleast as good as o′ for every agent i
- Further there is some agent who strictly prefers o to o′
- In this case, o seems defensibly better than o′

all comparisons within one single agent and done for all agents.

→ o Pareto dominates o′ whenever $o \succeq_i o'$ for all $i \in N$ and $o \succ_i o'$ for some $i \in N$

→ An outcome o^* is Pareto optimal if no other outcome Pareto dominates it.

→ Question is asked for a single agent and is done for all.
→ A game can have multiple pareto optimal outcomes.

Example:
o′ = Everyone gets pie
o = Everyone gets pie & Alice gets also cake
} o → seems better than o′ because everyone is okay with o and o′.

	Coop.	Defect
Coop.	−1, −1	−5, 0
Defect	0, −5	−3, −3

Nash Equilibrium

↳ more individual approach

$a_{-i} = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$ → every agent except for agent i
$a = (a_i, a_{-i})$

Best Response → which actions are better from an individual agent's viewpoint.

$$BR_i(a_{-i}) = \{a_i^* \in A_i \mid u(a_i^*, a_{-i}) \geq u(a_i, a_{-i}) \; \forall a_i \in A\}$$

↳ i's best response to a_{-i}

In Nash Equilibrium, stable outcome one where no agent regrets their action.

An action profile $a \in A$ is a (pure strategy) Nash Equilibrium iff.

$$\forall i \in N: \ a_i \in BR_i(a_{-i}) \quad \rightarrow \text{for every } i$$

- A game can have more than one pure strategy Nash equilibrium.
  ↳ Nihilistic Game e.g. → all payoffs (0,0)
- Can be the case where there is no P. Nash eq. →
  ↳ rock paper scissors (Yes)

	A	B
A	0,0	0,0
B	0,0	0,0

	Coop.	Defect
Coop.	−1, −1	−5, 0
Defect	0, −5	−3, −3

→ Nash Equilibrium

Pareto dominates the Nash equilibrium

- Matching Pennies has no pure strategy

Mixed Strategies

→ consider randomize

A strategy s_i for agent i is any probability dist. over the set A_i, where each action a_i is played with prob. $s_i(a_i)$.

strategy

Pure strategy: $s_i(a_i) = 1$ for some a_i only one action played

Mixed strategy: $s_i(a_i) < 1$ for all a_i for all a_i randomized over multiple actions

Set of i's strategies: $S_i = \Delta(A_i)$

Strategy profiles: $S = S_1 \times \dots \times S_n$

→ The utility of mixed strategy profile is its expected utility.

Assume that agents are decision theoretically rational

→ Agents randomize independently

$$\Delta(A_i) \times \dots \times \Delta(A_n)$$

$$u_i(s) = \sum_{a \in A} Pr(a|s)\, u_i(a) \quad \text{Prob. of profile } a \text{ given strategy profile}$$

utility of strategy profile

$$= \sum_{a \in A} \Big(\prod_{j \in N} s_j(a_j) \Big) u_i(a)$$

→ $Pr(a|s) = \prod_{j \in N} s_j(a_j)$ → independence assumption

The set of i's best responses to a strategy profile $s_{-i} \in S_{-i}$ is

$$BR_i(s_{-i}) = \{ a_i^* \in A_i \mid u_i(a_i^*, s_{-i}) \geq u_i(a_i, s_{-i}) \ \forall a_i \in A_i \}$$

A strategy profile is Nash Eq. iff,

$$\forall i \in N, \ s_i' \in S_i: \ u_i(s) \geq u_i(s_i', s_{-i})$$

Equivalently,

prob. that action a_i is played under strategy s_i

$$\forall i \in N, \ a_i \in S_i: \ s_i(a_i) > 0 \iff a_i \in BR_i(s_{-i})$$

When at least one s_i is mixed, s is a mixed strategy Nash eq.

Theorem ⇒ Every game with finite number of players and action profiles has atleast one Nash eq.

Is it possible/rational for an agent to play anything other than Nash eq.? Yes

↳ Nash Eq. only optimal if believe that other agents will play their parts of the same Nash equilibrium.

Computing Mixed Nash Equilibrium

	B	F
B	2,1	0,0
F	0,0	1,2

- Let player 2 play B with p, F with $1-p$
- If player 1 best-responses with mixed strategy, player 2 must make her indifferent b/w F and B.
  - ↳ because if there is preference, then it is not a mixed strategy as best-response.

$u_1(B) = u_1(F)$

$2p + 0(1-p) = 0p + 1(1-p)$

$p = 1/3$

↳ If p is −ve or bigger than 1 then there is no way of picking p so that player is indifferent.

- Likewise, player 1 must randomize to make player 2 indifferent.
  - Player 1 is willing to randomize because she is being indifferent.

Let player 1 play B with q, F with $1-q$

$u_2(B) = u_2(F)$

$q + 0(1-q) = 0q + 2(1-q)$

$q = \frac{2}{3}$

⇒ Each randomize b/w all the things such that it makes the other person indifferent.

Strategies: $(\frac{2}{3}, \frac{1}{3}), (\frac{1}{3}, \frac{2}{3})$ are Nash equilibrium.

Maxmin Strategy → defensive

Player's i's maxmin strategy is a strategy that maximizes i's worst-case payoff, in the situation where all other players $(-i)$ happen to play the strategies which cause greatest harm to i.

The maxmin value of the game for player i is that minimum amount of playoff guaranteed by a maxmin strategy.

$\arg\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$ → maxmin strategy

maxmin value $= \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$ → maxmin value

Could play ⇒ conservative agent maximizing worst-case payoff
paranoid agent believing everyone is out to get him.

Minmax Strategy → road-rage of GT

Player i's minmax strategy against player $-i$ in a 2-player game is a strategy that minimizes $-i$'s best-case payoff and the mimax value for i against $-i$ is payoff.

$\arg\min_{s_i} \max_{s_{-i}} u_{i}(s_i, s_{-i})$ → strategy minmax

$\min_{s_i} \max_{s_{-i}} u_{-i}(s_i, s_{-i})$ → value of minmax

Could play ⇒ burn the world, revenge, threat

Theorem (Von Neumann) ⇒ In any finite, two-player zero-sum game, in any Nash equilibrium each player receives a payoff that equals both his minmax value and maxmin value.

1- Each player's maxmin value is equal to his minmax value. By convention, the maxmin player for player 1 is called the value of the game.

2- Any maxmin strategy profile, minmax strategy profile is a Nash equilibrium. All Nash equilibrium have same payoff vector.

3- For both players, the set of maxmin strategies coincides with set of minmax strategies.

Iterated Removal of Dominated Strategies

→ strategy where no matter what the other player does, my choice is always better

- No equilibrium can involve a strictly dominated strategy
- Remove it and end up with strategically equal game
- Allows us to remove another strategy that wasn't dominated before

→ can do this over & over to termination

• Used as a pre-processing step before computing an equilibrium
→ In traveler's dilemma, doing this can end up with just (180,180) → equilibrium

Rationalizability → In 2-player games, those that survive iterated removal of dominated strategies

• Rather than ask what's irrational, ask what is a best response to some beliefs about opponent.

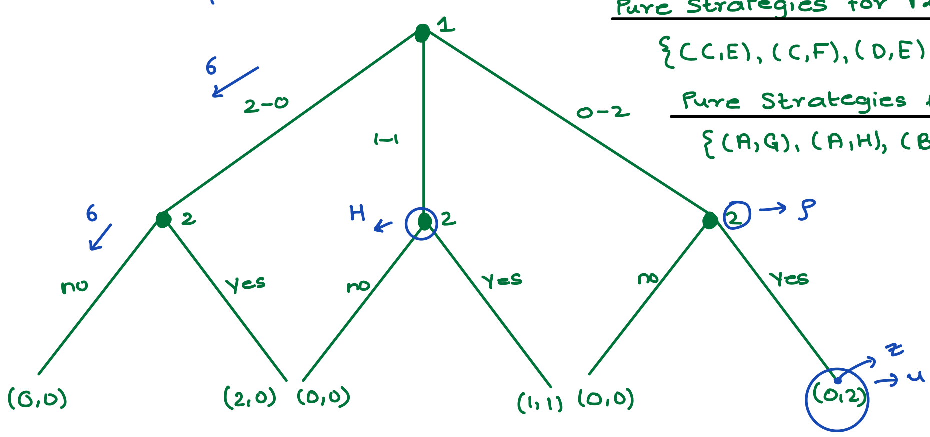
→ Equilibrium strategies are always rationalizable

Furthermore, in two-player games, rationalizable ⟺ survives iterated removal of dominated strategies

Extensive Form Games

Normal-form games do not have any notion of sequence: all actions happen simultaneously.

The extensive form is a game representation that explicitly includes temporal structure (tree representation).



Pure Strategies for P2

{(C,E), (C,F), (D,E), (D,F)}

Pure Strategies for P1

{(A,G), (A,H), (B,G), (B,H)}

Perfect-Information → Every agent sees all actions of the other players (including any special "Chance" player)

A finite perfect-information game in extensive form is a tuple $G = (N, A, H, Z, \chi, \rho, \sigma, u)$ where

N → set of n players

A → set of actions

H → set of non-terminal choice nodes

Z → set of terminal nodes (disjoint from H)

$\chi : H \to 2^A$ is the action func. (actions available for choice nodes)

$\rho : H \to N$ is the player function (turn of the player)

$\sigma : H \times A \to H \cup Z$ is the successor function (what to do next, deterministic)

$u : (u_1, \dots, u_n)$ → profile of utility fns

$u_i : Z \to \mathbb{R}$ for each player i

Pure Strategies (in perfect-information in EFG)

$\prod_{h \in H | \rho(h) = i} \chi(h)$ → cross-product of actions available to i at each of their choice nodes

→ associate an action with every choice-node, even those that will be never be reached.

Induced Normal Form

- Any pair of pure strategies uniquely identifies a terminal node, which identifies a utility for each agent.

	C,E	C,F	D,E	D,F
A,G	3,8	3,8	8,3	8,3
A,H	3,8	3,8	8,3	8,3
B,G	5,5	2,10	5,5	2,10
B,H	5,5	1,0	5,5	1,0

→ induced normal form

- Any perfect-information extensive-form game defines a corresponding induced-normal form game.

Mixed Strategy → prob. distribution over pure strategies

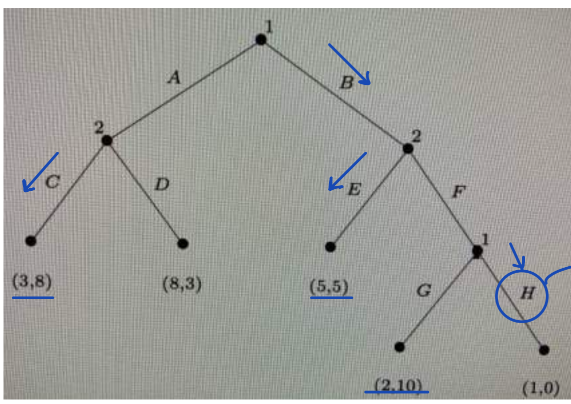
Best Response → set of mixed strategies that is utility maximizer against some other set of mixed strategies.

Theorem (Zermelo) ⇒ Every finite perfect-information game in extensive form has atleast one Nash equilibrium (pure)

every normal form game has NE, and every extensive form game has induced normal game.

Solve by backward induction

- Start from the bottom of the tree, no agent needs to randomize because deterministic best response.
- Replace those nodes with resulting utility vector.
- Repeat until an action is assigned for all choice nodes.



→ less compelling

→ Any one implausible?

- (A,H), (C,F)
- (B,H), (C,E)

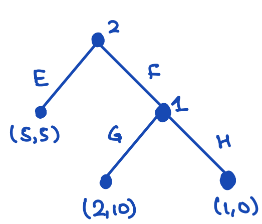
} both are implausible

Subgame Perfection

- Some equilibria are better than others

subgame → The subgame of G rooted at h is the restriction of G to the descendants of h.

The subgames of G are the subgames of G rooted at h for every choice node h ∈ H.



(subgame example)

Strategy Profile $s \in S$ is a subgame perfect equilibrium of G iff, for every subgame G′ of G, the restriction of s to G′ is a Nash eq. of G′.

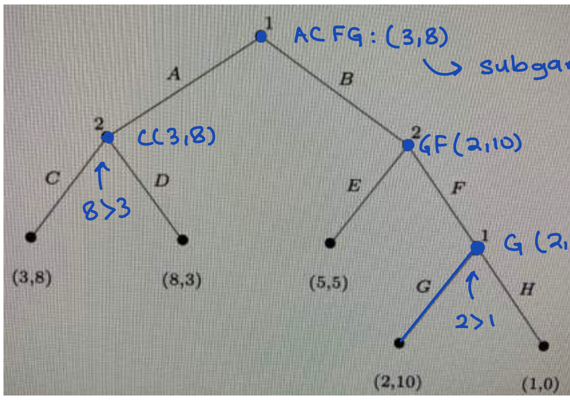
⇒ So, any equilibrium computed by backward induction will be subgame perfect.

↳ because it goes back up as the backward induction is done.

→ So every finite perfect-info game in EF has atleast one subgame perfect equilibrium.

Backward Induction → finds the subgame perfect equilibrium

↳ replace subgames with their equilibrium values



⟹ (A,G), (C,F)

Imperfect Information Extensive Form Games

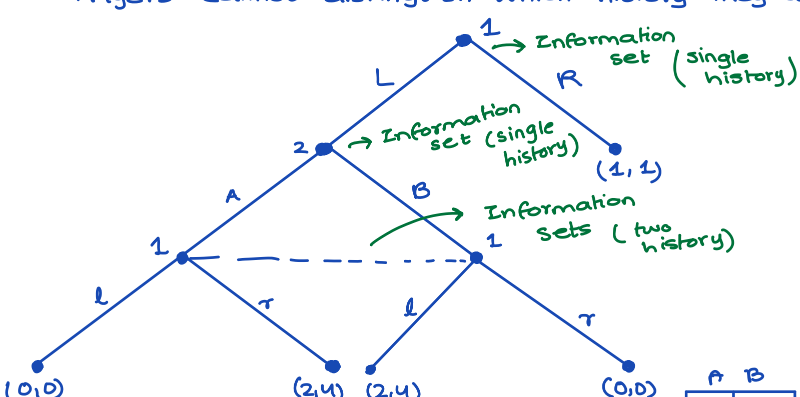
↳ model games with sequential actions some of which are hidden

$G = (N, A, H, Z, \chi, \rho, \sigma, u, I)$

- $N \rightarrow$ set of n players
- $A \rightarrow$ set of actions
- $H \rightarrow$ set of non-terminal choice nodes
- $Z \rightarrow$ set of terminal nodes (disjoint from H)
- $\chi : H \rightarrow 2^A$ is the action func. (actions available for choice nodes)
- $\rho : H \rightarrow N$ is the player function (turn of the player)
- $\sigma : H \times A \rightarrow H \cup Z$ is the successor function (what to do next, deterministic)
- $u : (u_1, \dots, u_n) \rightarrow$ profile of utility fns
  - $u_i : Z \rightarrow \mathbb{R}$ for each player i
- $I : (I_1, \dots, I_n)$, where for each $i \in N$, $I_i = (I_{i,1}, \dots, I_{i,k_i})$ is an equivalence relation on (i.e., partition of) $\{h \in H \mid \rho(h) = i\}$ and

For every h, h' such that $h \in I$ and $h' \in I$ for some $i \in N$ and $I \in I_i$,
$\rho(h) = \rho(h') = i$ and $\chi(h) = \chi(h')$

The elements of partition are often called information sets.
Players cannot distinguish which history they are in within an information set.



induced normal form ←

	A	B
L,l	0,0	2,4
L,r	2,4	0,0
R,l	1,1	1,1
R,r	1,1	1,1

Pure Strategies

Need something to take information set / choice node and map it to the action.

$$\prod_{I \in I_i} \chi(I) \rightarrow \text{actions available at each information sets}$$

where $\chi(I) = \chi(h)$ for any arbitrary $h \in I$

Mixed Strategies

Distribution over pure strategies.

Best Response

A set of strategies maximizing utility given another set of strategy.

Q- Can you represent an arbitrary perfect information EFG as an imperfect information EFG? Yes. Every single information set will contain one choice node.

→ Any normal form game can be represented as imperfect information EFG.

→ Players can play in any order

because none of the player that goes second can see what the first player did.

↳ one player plays first, second player doesn't get to see what first player did and it plays and game ends.

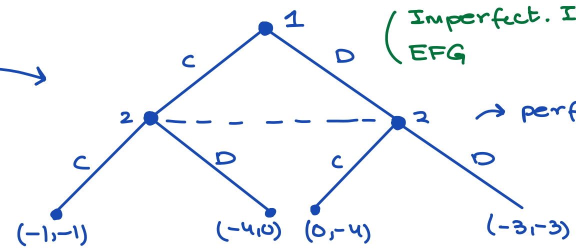
Q- What happens if we run this NFG → EFG translation on induced normal form game of any arbitrary EFG? → lose all the structure and nothing is gained.

↳ cannot recover EFG structure

both will have same Nash Eq.

	C	D
C	-1,-1	-4,0
D	0,-4	-3,-3

G (NFG)



(Imperfect. Information) EFG

→ perfect-recall

Behavioural vs. Mixed Strategies

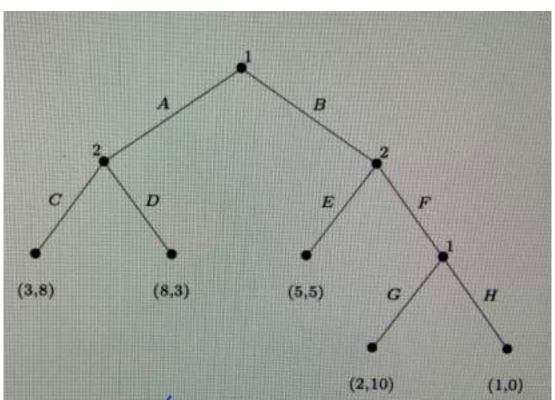
Mixed strategy s_i (in an imperfect-information EFG) is any distribution over an agent's pure strategies:

$$s_i \in \Delta(A^{I_i})$$

Behavioural strategy b_i is a mapping from an agent's information sets to a distribution over the actions at that information set, which is sample independently each time the agent arrives at the information set

$$b_i \in [\Delta(\chi(I))]_{I \in I_i}$$

↳ randomize at the information set



perfect-information

Behavioural Strategy: ([.6: A, .4: B], [.6: G, .4: H])

Mixed Strategy: [.6: (A, G), .4: (B, H)]

Are they equivalent? No.

Two strategies are equivalent if they induce same dist. over outcomes for all strategies of other player.

These are not equivalent.

Person 2: chooses F then these two give different results.

Can you construct a mixed strategy that is equivalent to the behavioral strategy above? Yes.

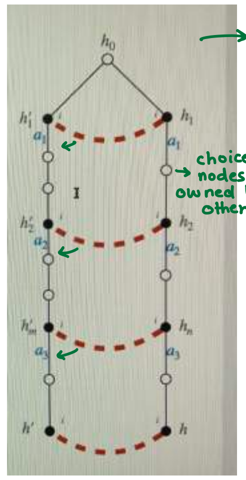
Can you construct a behavior strategy that is equivalent to the mixed strategy above? Yes.

game has to be perfect-recall for this to be true

Perfect Recall

If for any pair of choice nodes in same information set and for any path $h_0, a_0, h, a_1, \ldots, h_n, h$ from the root of the game to h and for any path $h_0', a_0', h_1', a_1', \ldots, h_m', h'$ from the root of the game to h', it must be the case that:

1) $n = m$ (same nodes)
2) for all $0 \le j \le n$, if $\rho(h_j) = i$, then exists an information set $I' \in I_i$ such that $h_j, h_j' \in I'$ (both are in same information set)
3) for all $0 \le j \le n$, if $\rho(h_j) = i$, then $a_j = a_j'$ (same action)



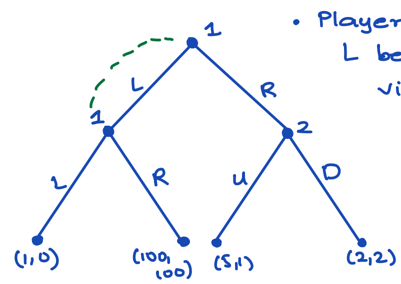
→ never forget where we are in the tree

If I can't tell where I am in the tree then that is not because I forgot something.

⇒ G is a game of perfect recall if every player has perfect recall in G.

→ A game of perfect-information is a game of perfect-recall.

Imperfect Recall



• Player 1 doesnot remember whether they played L before or not. In this case, because the visit same information set multiple times.

↳ don't have to always visit same information set for imperfect recall.

Mixed Strategy → has ability to do the same thing at a given info. set

Behavior Strategy → allows randomization at the given info. set.

Mixed Strategy Equilibrium

→ convert to induced normal form and there is always a MSE.

Equilibrium in Behavioural Strategy → not guaranteed always

Model as Imperfect Recall?

→ when actual agents may forget previous history
  • cases where the agent strategies are executed by proxies

→ approximation technique

Theorem (Kuhn's) ⇒ If game has perfect recall, then any mixed strategy has equivalent behavior strategy and vice versa.

Repeated Games

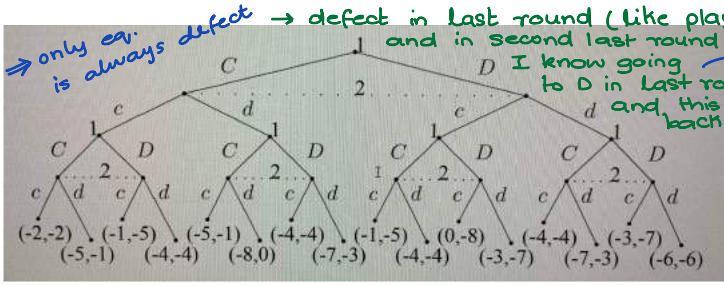
Play the same NFG over and over game
- each round is called stage game

Finitely Repeated Game

↳ written as extensive form game with imperfect information
- at each round players don't know what other have done
- overall payoff ftn is additive: sum of payoffs in stage games

$\Rightarrow$ only eq. is always defect $\rightarrow$ defect in last round (like playing Prisoner's dilemma 1) and in second last round

I know going to D in last round and this backs up. $\rightarrow$ playing prisoner's dilemma twice



	C	D
C	$-1,-1$	$-4,0$
D	$0,-4$	$-3,-3$

$\rightarrow$ normal form game

- Repeating a Nash strategy in each stage game will be an equilibrium in behavioural strategies (called stationary strategy).
- When NFG has dominant strategy, then we can apply backward induction.

Infinitely Repeated Game

$\hookrightarrow$ infinite deep game tree

can't have payoffs at the terminal node

can't add stage game payoffs (divergent series)

Given an infinite sequence of payoffs $r_1, r_2, \ldots$, for player i, the average reward of i is

$$\lim_{k \to \infty} \sum_{j=1}^{k} \frac{r_j}{k}$$

Discounted Reward $\rightarrow$ $0 < \beta < 1$

$$\sum_{j=1}^{\infty} \beta^j r_j$$

- near term well-being is more important
- agent cares about future just as much as present, but with probability $1-\beta$ the game will end.

Pure Strategy in Infinitely Repeated Game

$\rightarrow$ choice of action at every decision point

action at every stage game $\rightarrow$ infinite number of actions

set of pure strategies

Nash Equilibria

- with infinite number of equilibria, can't construct an induced normal form
- Nash's theorem only applies to finite game
- Could be infinite number of pure-strategy equilibria
- Can characterize a set of payoffs that are achievable under equilibrium, without having to enumerate equilibria

Folk Theorem

Consider any n player game $G = (N, A, u)$ and payoff vector $r = (r_1, r_2, \ldots, r_n)$

$\hookrightarrow$ real number payoff to each player of game

$$v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$$

$\hookrightarrow$ i's minmax value: the amount of utility i can get when when $-i$ play minmax against i.

guaranteed for i

- A payoff profile r is enforceable if $r_i \geq v_i$ $\rightarrow$ everyone gets atleast their minmax value
- A payoff profile r is feasible if there exist rational, non-negative values α_a such that for all i, we can express r_i as $\sum_{a \in A} \alpha_a u_i(a)$ with $\sum_{a \in A} \alpha_a = 1$

$\hookrightarrow$ expressed as weighted sum

- payoff profile is feasible if it is a convex, rational combination of outcomes in G.

Theorem (Folk's)

Consider any n-player game G and any payoff vector $(r_1, r_2, \ldots, r_n)$

- If r is the payoff in any Nash eq. of the infinitely repeated G with average rewards, then for each player i, r_i is enforceable.
- If r is both feasible and enforceable, then r is the payoff in some Nash eq. of the infinitely repeated G with average rewards. $\rightarrow$ there is a Nash eq. with an average payoff vector r iff. r is enforceable & feasible.

feasible & enforceable $\longrightarrow$ Nash equilibrium

$\hookrightarrow$ never accept anything lower than what others are forcing you to get when BR.

Bayesian Games

→ randomness about choice of game, other person's and/or own utility ftn.

(Def #1)

→ a set of games that differ only in their payoffs, a common prior defined over them, and a partition over the games for each agent.

(N, G, P, I) where

$N \to$ set of agents

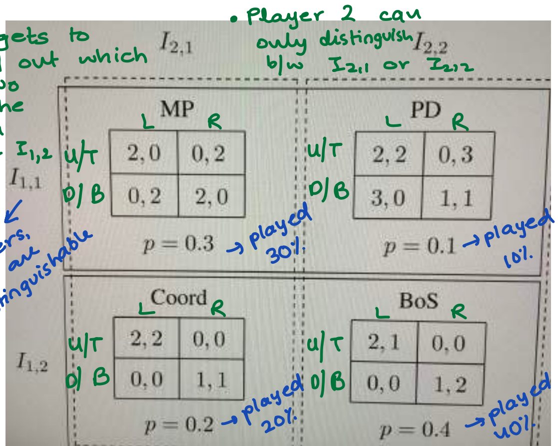
$G \to$ set of games with N agents each such that if $g, g' \in G$ then for each agent $i \in N$ the strategy space in g is identical to strategy space in g'

how likely each of the game is. ← $P \in \Pi(G) \to$ common prior over games, where $\Pi(G)$ is a set of all prob. dist. over G → update prior based on partition

$I = (I_1, \dots, I_N) \to$ set of partitions of G, one for each agent

• P1 gets to find out which of two sets he is in $I_{1,1}$ or $I_{1,2}$

• Player 2 can only distinguish b/w $I_{2,1}$ or $I_{2,2}$



L R U/T D/B

→ played 30%

→ played 10%

→ played 20%

→ played 40%

for row players, rows are indistinguishable

→ Players can see their info set before making a move

→ Game is sampled based on a common prior

→ Information is updated when game is selected

→ Two players get to see different things

Epistemic Types (Def #2)

↳ directly represent uncertainity over utility function using the notion of epistemic type

Bayesian Game = (N, A, Θ, p, u)

$N \to$ set of agents

$A \to (A_1, \dots, A_n)$, A_i set of actions for i

$\Theta \to (\Theta_1, \dots, \Theta_n)$, where Θ_i type space of i

over joint types not over games ← $p \to \Theta: [0,1]$ common prior over types

$u \to (u_1, \dots, u_n)$, $u_i: A \times \Theta \to \mathbb{R}$ utility ftn for i

so agents have posterior abt other types

↳ depends on action taken and type of everyone

• Agent can't see type of others, can only see own type

• when players are deciding which action to take, they have to decide w/o fully knowing the game.

can be written this way as well

row agent = U, D

Col. agent = L, R

a_1	a_2	θ_1	θ_2	u_1	u_2
U	L	$\theta_{1,1}$	$\theta_{2,1}$	2	0
U	L	$\theta_{1,1}$	$\theta_{2,2}$	2	2
U	L	$\theta_{1,2}$	$\theta_{2,1}$	2	2
U	L	$\theta_{1,2}$	$\theta_{2,2}$	2	1
U	R	$\theta_{1,1}$	$\theta_{2,1}$	0	2
U	R	$\theta_{1,1}$	$\theta_{2,2}$	0	3
U	R	$\theta_{1,2}$	$\theta_{2,1}$	0	0
U	R	$\theta_{1,2}$	$\theta_{2,2}$	0	0
D	L	$\theta_{1,1}$	$\theta_{2,1}$	0	2

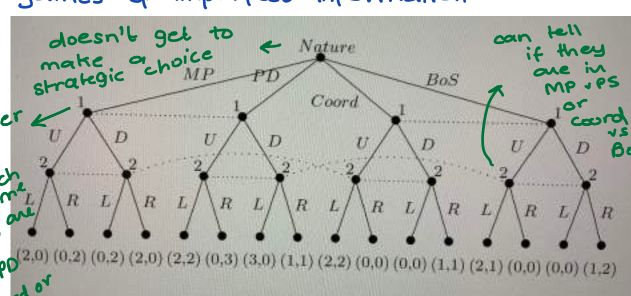
first agent, second agent

→ repeat for others similarly

Extensive Form with Chance Moves

• Add agent 'Nature' following a known mixed strategy based on prior

• Reduce Bayesian games to extensive form games of imperfect information



doesn't get to make a strategic choice

can tell if they are in MP vs PS or coord vs. BoS

player can tell which game they are in MP, PD or coord or BoS

Pure Strategies

→ a mapping from every type agent observes it has to the action it would play if he had that type

$s_i: \Theta_i \to A_i$ → action for each type

Mixed Strategy

→ a mapping from all the types agent might have to distribution over action choices

$s_i: \Theta_i \to \Pi(A_i)$ → prob. dist. over the actions

$s_j(a_j \mid \theta_j)$

→ the prob. under mixed strategy s_j that agent j plays a_j given they observed type θ_j

Ex-interim

→ an agent knows type of own but not of others

$$EU_i(s|\theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} \underbrace{p(\theta_{-i}|\theta_i)}_{\text{belief about other agent's types}} \sum_{a \in A} \underbrace{\left(\prod_{j \in N} s_j(a_j|\theta_j)\right)}_{\text{prob. of actions}} \underline{u_i(a, \theta_{-i}, \theta_i)}$$

i must consider every θ_{-i} and every a in order to evaluate $u_i(a, \theta_i, \theta_{-i})$

→ agent's strategies given by mixed strategies

u_i weighted by
prob. that a would be realized given
all player's mixed strategies and types
prob. that other players type would be θ_{-i} given own type θ_i

Ex-Ante

→ the agent knows nothing about anyone's type

$$EU_i(s) = \sum_{\theta_i \in \Theta} p(\theta_i) EU_i(s|\theta_i)$$

$$EU_i(s) = \underbrace{\sum_{\theta_i \in \Theta} p(\theta_i)}_{\text{common prior}} \sum_{a \in A} \left(\prod_{j \in N} s_j(a_j|\theta_j)\right) \underline{u_i(a, \theta)}$$

↳ sum and weigh by the prior over types

utility weighted by all prob. under mixed strategies

Ex-post

→ an agent knows all agent's type

$$EU_i(s, \theta) = \sum_{a \in A} \left(\prod_{j \in N} s_j(a_j|\theta_j)\right) u_i(a, \theta)$$

- The only uncertainity here concerns the other agent's mixed strategies

Best Response

↳ best response to mixed profile s_{-i} of i

$$BR_i(s_{-i}) = \arg\max_{s_i' \in S_i} EU_i(s_i', s_{-i})$$

↳ based on i's ex-ante expected utility

- Equivalent to performing indep. max of i's ex-interim expected utility conditioned on each type that i could have.

expectation over
→ action's of other players
↳ types of other players

Nash Equilibrium

→ maximize each type's expected utility

A Bayes-Nash eq. is a mixed strategy profile s that satisfies

based on Ex-interim

$\forall i \quad \underbrace{s_i \in BR_i(s_{-i})}_{\text{map from types to actions}}$ → every body is best responding to each other

- Can also construct an induced normal form for Bayesian games
- the numbers in the cell correspond to ex-ante expected utilities
  ↳ as long as the strategy space is unchanged, best responses don't change b/w ex-ante & ex-interim cases

Ex-post Equilibrium

$$\forall \theta, \forall i \quad s_i \in \arg\max_{s_i' \in S_i} EU_i(s_i', s_{-i}, \theta)$$

agents do not to have accurate beliefs about type distribution

* can construct induced normal form for Bayesian game

→ the numbers in cell correspond to ex-ante expected utilities

→ As long as the strategy space is unchanged, best responses don't change b/w ex-ante & ex-interim cases.

Analyzing Bayesian Games

A Sheriff's Dilemma

↳ a sheriff faces an armed suspect and they each must (simultaneously) decide whether to shoot the other or not.

- suspect is either a criminal with prob. p or not with $(1-p)$
- sheriff would rather shoot if suspect shoots, but does not if the suspect does not
- the criminal would shoot regardless of sheriff, as the suspect would be caught if does not shoot
- an innocent suspect would not shoot regardless of sheriff

Sheriff ↓ (columns); suspect → (rows); θ_I, $1-p$; innocent type of suspect

Good	Shoot	Not
Shoot	$-3,-1$ ($-2>-3$)	$-1,-2$ ($0>-1$)
Not	$-2,-1$	$0,0$ ✓

strictly dominant strategy for the good player
If the player sees that he is innocent, the would not shoot since that is his dominant strategy.

Sheriff ↓ (columns); Sus. → (rows); θ_G, p

Bad	Shoot	Not
Shoot	$0,0$ ($0>-2$)	$2,-2$ ($2>-1$)
Not	$-2,-1$	$-1,1$

criminal (guilty) suspect
strictly dominant to shoot so can remove not shoot

Sheriff ↓; suspect →; θ_G, p

Bad	Shoot	Not
Shoot	$0,0$	$2,-2$

So can remove shoot row for the suspect

Sheriff ↓; suspect →; θ_I, $1-p$

Good	Shoot	Not
Not Shoot	$-2,-1$	$0,0$

$-1(1-p)$ → if sheriff shoots

Sheriff's best response?

$-2p$ → if sheriff doesnot shoot

$-1(1-p) = -2p$

$-1+p = -2p$

$\cancel{+}1 = \cancel{+}3p$

$3p = 1$

$p = 1/3$

$p > 1/3$ → suspect is bad and better to shoot else do not shoot
If $p = 1/3$, they are indifferent

for the sheriff

Bayesian Equilibria in Innocent Type
↳ always not shoot
Bayesian Equilibria in Guilty Type
↳ always shoot

for the suspect

→ Strategic uncertainity about what other player is going to do

→ payoff uncertainity about what the best response will be (value to their actions)

Compute Mixed Strategy Nash Equilibriu

1) ↓P2; P1 →

	RL	RR
Up	$2.4^*, 3.2^*$	$0, 2.4$
Down	$1.8, 0.4$	$1^*, 1.2^*$

* Best responses

P1 play Up with p(Up) and down with (1 − p(Up))

$u_2(RL) = u_2(RR)$

$3.2 \times p + 0.4 \times (1-p) = 2.4p + 1.2(1-p)$

$3.2p + 0.4 - 0.4p = 2.4p + 1.2 - 1.2p$

$2.8p + 0.4 = 1.2p + 1.2$

$2.8p - 1.2p = 1.2 - 0.4$

$1.6p = 0.8$

$\Rightarrow p(Up) = \frac{0.8}{1.6} = 1/2$

P2 play RL with p(RL) and RR with 1 − p(RL)

$u_1(Up) = u_1(Down)$

$2.4 \times p(RL) + 0 \times (1-p(RL)) = 1.8\, p(RL) + 1 \times (1-p(RL))$

$\Rightarrow p(RL) = 0.625$

2)

P1 ↓ / P2 →	L	R
T	80, 40	40, 80
B	40, 80	80, 40

P1 play T with p(T) and B with p(1−p(T))

$u_2(L) = u_2(R)$

$40 \times p(T) + 80(1-p(T)) = 80 \times p(T) + 40(1-p(T))$

$40p + 80 - 80p = 80p + 40 - 40p$

$-40p + 80 = 40p + 40$

$-40p - 40p = 40 - 80$

$\cancel{-}80p = \cancel{-}40$

$p = 40/80 = 1/2$

$p(T) = \frac{1}{2}$, $p(B) = 1/2$ } row player randomizes 50-50 to make column player indifferent.

P2 play L with p(L) and R with 1-p(L)

$u_1(T) = u_1(B)$

$80 \times p(L) + 40 \times (1-p(L)) = 40 \times p(L) + 80 \times (1-p(L))$

$80p + 40(1-p) = 40p + 80(1-p)$

$80p + 40 - 40p = 40p + 80 - 80p$

$40p + 40 = -40p + 80$

$40p + 40p = 80 - 40$

$2\cancel{80}p = \cancel{40}\,1$

$p = 1/2$

So, the strategies are $(1/2, 1/2), (1/2, 1/2)$

3)

P1 ↓ / P2 →	L	R
U/T	320, 40	40, 80
D/B	40, 80	80, 40

P2 plays L with p(L) and R with 1-p(L)

$p(T) = p(B)$

$320 \times p + 40 \times (1-p) = 40p + 80(1-p)$

$320p + 40 - 40p = 40p + 80 - 80p$

$(320 - 40)p + 40 = -40p + 80$

$280p + 40 = -40p + 80$

$280p + 40p = 80 - 40$

$320p = 40$

$p = \frac{40}{320} = 1/8 = 0.125$

Strategy: $(\frac{1}{2}, \frac{1}{2}), (\frac{1}{8}, \frac{7}{8})$

4)

P1 ↓ / P2 →	L	R
U/T	44, 40	40, 80
D/B	40, 80	80, 40

P2 plays L with p(L) and R with 1-p(L)

$p(T) - p(B)$

$44p + 40 \times (1-p) = 40p + 80(1-p)$

$(44-40)p + 40 = -40p + 80$

$4p + 40p = 80 - 40$

$44p = 40$

$p = 40/44$

$p = 0.9$

Strategy: $(\frac{1}{2}, \frac{1}{2}), (0.9, 0.1)$

- A strategy profile consisting of dominant strategy for every player must be a Nash equilibrium.
- An equilibrium in strictly dominant strategies must be unique.

Pareto Optimality

↳ from the outside observer

outcome o is at least as good as o′ for everyone and some agent prefers o ≻ o′ (strictly)

→ so reasonable to say that o is better than o′

→ we say that o Pareto-dominates o′.

- A game can have more than one Pareto-optimal outcomes
- Every game must have atleast one Pareto-optimal outcome.
- Every outcome in 0 sum game is Pareto-optimal.

⇒ An outcome o^* is Pareto-optimal if there is no other outcome that Pareto-dominates it.

Mechanism Design

Bayesian Game Setting

→ a Bayesian game with no actions, but instead with a set of outcomes over which agents have utilities that depend on their types.

(N, O, Θ, p, u) where

- N: finite set of n agents
- O: set of outcomes
- $\Theta = \Theta_1 \times \dots \times \Theta_n$: set of possible joint type vectors
- p: common prior (prob. dist over Θ)
- u: utility set $u_i : O \times \Theta \to \mathbb{R}$

Mechanism → creates a set of actions and this can be sequential, simultaneous, anything.

A mechanism (for Bayesian game (N, O, Θ, p, u)) is a pair (A, M) where
- $A = A_1 \times \dots \times A_n$ where A_i set of actions available to agent $i \in N$
- $M : A \to \Pi(O)$ maps each action profile to a distribution over outcomes

The designer gets to specify,
- the action sets for agent (though these maybe constrained by the env)
- the mapping to outcomes, over which agents have utility
- can't change outcomes; agent's preferences or type spaces

The problem is to pick a mechanism that will cause rational agents to behave in a desired way, specifically maximizing the mechanism designer's own utility or objective func.

- each agent holds private information, in Bayesian game sense
- often we are interested in settings where agent's action space is identical to their type space, and an action can be interpreted as a declaration of the agent's type.
- perform an optimization problem, given that the values of (some of) the inputs are unknown
- choose the Bayesian game out of a set of possible Bayesian games that maximizes some performance measure.

Auctions

↳ mechanisms for allocating resources among self-interested agents
→ outcome consists of allocating scarce resources

English Auction → auctioneer starts the bidding at some "reservation price"
bidders shout ascending price
highest bidder gets the good

First price Auction (silent auction) → bidders write down bids on pieces of paper
auctioneer awards the good to bidder with highest bid
the bidder pays amount of bid

Second price Auction → same like first pay auction
winner pays second-highest bid amount

All-pay Auction → everybody pays, winner gets awarded

Representation as Bayesian Game

Agents: bidders

Actions: bid amounts b_i

Types: valuations v_i

Common prior: agent's types drawn independently (IPV)

Risk attitude: risk-neutral, money & utility relationship is linear

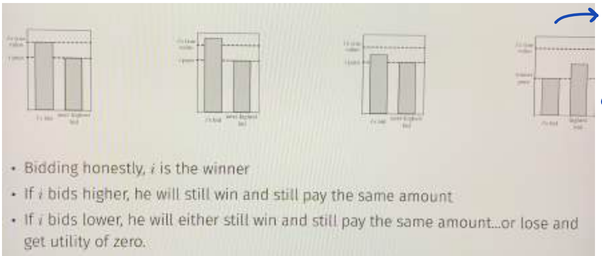
Allocations & payments: determined based on vector of bid amounts b

Utility fn: If i is awarded $v_i - p_i$ else 0
↳ (IPV, risk-neutral case)

Outcomes: allocations & payments
↳ cartesian prod. of who gets what & what everyone pays

Second Price Analysis

Theorem ⇒ Bidding one's value (truthful value) is a dominant strategy in a second price aution



→ case # 1: bidding honestly, i is the winner

case # 2: bidding honestly, i is the loser

↳ can be thought of conversely as of the first.

First Price Analysis → both bidders half their value

Theorem ⇒ In FPA, with two bidders (risk neutral) whose valuations are drawn independently and uniformly drawn at random from $[0,1]$, $(\frac{1}{2}v_1, \frac{1}{2}v_2)$ is a Bayes-Nash equilibrium strategy profile.

If n risk-neutral agents then unique symmetric equilibrium is given by strategy profile $\left(\frac{n-1}{n}v_1, \ldots, \frac{n-1}{n}v_n\right)$

Revenue Equivalence

The k^{th} order statistic of a distribution: the expected value of the k^{th}-largest of n draws.

For n IID draws from $[0, v_{max}]$, the k^{th}-order statistic $\frac{n+1-k}{n+1} v_{max}$

↓ from uniform dist.

In second price auction, the seller's expected revenue is

$$\frac{n-1}{n+1} v_{max}$$

- every symmetric game has symmetric equilibrium
- in a symmetric equilibrium of this auction game, higher bid ⇔ higher valuation

} conditions of revenue-equivalence theorem

Computing Bayes Nash Equilibria

$\theta_{1,1}$

$(p=0.3)$ $\theta_{2,1}$ MP

	L	R
U	2,0	0,2
D	0,2	2,0

$\theta_{2,2}$ PD $(p=0.1)$

	L	R
U	2,2	0,3
D	3,0	1,1

$\theta_{1,2}$

$(p=0.2)$ Coord

	L	R
U	2,2	0,0
D	0,0	1,1

BoS $(p=0.4)$

	L	R
U	2,1	0,0
D	0,0	1,2

c ↓

r →

	LL	LR	RL	RR
UU	2,1	1,0.7	1,1.2	0,0.9
UD	0.8,0.2	1,1.1	0.4,1	0.6,1.9
DU	1.5,1.4	0.5,1.1	1.7,0.4	0.7,0.1
DD	0.3,0.6	0.5,1.5	1.1,0.2	1.3,1.1

$EU_1(UU, LL) = 2(0.3) + 2(0.1) + 2(0.2) + 2(0.4)$
$= 2$

$EU_2(UU, LL) = 0(0.3) + 2(0.1) + 2(0.2) + 1(0.4)$
$= 1$

Ex-ante expected utility

Behavioral Game Theory

Best response: Max. utility action is always played

Quantal response: High-utility action is played often than low-utility actions

$$QBR(s_{-i};\lambda)(a_i) = \frac{\exp(\lambda u_i(a_i, s_{-i}))}{\sum_{a_i' \in A} \exp(\lambda u_i(a_i', s_{-i}))} \quad \text{(softmax)}$$

λ → sensitivity to differences in utility

Quantal Response Eq.

→ can have various eq.

→ A strategy profile s is a quantal response eq. (QRE) with precisions λ if every agent is simultaneously quantally responding to the profile of the other agents' strategies i.e.

$$\forall i \in N: s_i = QBR_i(s_{-i}; \lambda)$$

- not clear how agents would arrive at a QRE
- NFG have always a QRE per λ

Note that agents still have rational expectations: they are responding to the real strategies of the other agents

- can relax rational expectations

Level-k Model

Every agent performs some finite number of steps of strategic reasoning:

level-0: some default, non-strategic distribution of play (often uniform)

level-1: Best response to level-0 players

Level-2: Best response to level-1 or to level-1 and level-0 players

⋮

level-k: Best response to level $k-1$ or to levels $\{0, 1, \dots, k-1\}$

A strategy profile s is the prediction of a level-k model with parameters $\alpha_0, \alpha_1, \dots, \alpha_k$ and level-0 strategies $\pi_{i,0}^{Lk}$ if

$$\forall i \in N: s_i = \sum_{k=0}^{K} \alpha_k \pi_{i,0}^{Lk}$$

where $\pi_{i,o}^{Lk} = BR_i(\pi_{-i,k-1}^{Lk})$

$\pi_{i,0}^{Lk}$ → level-0 strategy for each agent

- Every agent has fixed level representing the number of steps of strategic reasoning they can perform
  - They assume that every agent performs exactly one step fewer

$\alpha_0, \dots, \alpha_K$ → frequency of levels

Cognitive Hierarchy

- Agents know that other agents might perform any number of steps of reasoning (less than their own)
- Agents know the relative proportions of all lower levels

A strategy profile s is the prediction of a cog. hier model with parameters $\alpha_0, \alpha_1, \dots, \alpha_k$ and level-0 strategies $\pi_{i,0}^{CH}$ if

$$s_i = \sum_{k=0}^{K} \alpha_k \pi_{i,k}^{CH}$$

$$\pi_{i,k}^{CH} = BR(\pi_{-i,k-1}^{CH}) \quad \text{for all } k > 0$$

$$\pi_{i;0:k}^{CH} = \frac{\sum_{j=0}^{k} \alpha_j \pi_{i,j}^{CH}}{\sum_{j=0}^{k} \alpha_j}$$

- All of these (level-k / CH / QCH) assume that $\pi_{i,0}$ is the uniform assumption.
- Or hand-pick default strategy
- Implausible & don't scale
- Ideally, want level-0 strategy to learn from data.

Quantal Cognitive Hierarchy

A strategy profile s is the prediction of a cog. hier. model with parameters $\alpha_0, \alpha_1, \dots, \alpha_k$ and level-0 strategies $\pi_{i,0}^{QCH}$ if

$$s_i = \sum_{k=0}^{K} \alpha_k \pi_{i,k}^{QCH}, \quad \pi_{i,k}^{QCH} = QBR_i(\pi_{-i,k-1}^{QCH}; \lambda) \quad \pi_{i;0:k}^{QCH} = \frac{\sum_{j=0}^{k} \alpha_j \pi_{i,j}^{QCH}}{\sum_{j=0}^{k} \alpha_j}$$

Level-0 Model: Linear4 → stick it on top of other iterative model (QCH e.g.)

$\pi_{i,0}$ is a Linear4 Level-0 model with parameter $\vec{w}$ if

$$\pi_{i,0}(a_i) \propto \sum_{f \in F} w_f f(a_i)$$ → different weights for every feature

where $F = \{ f^{maxmax}, f^{maxmin}, f^{eff}, f^{fair}, f^{unif} \}$ and

$f^{maxmax}(a_i) = 1$ iff $a_i \in \arg\max_{a_i' \in A} \max_{a_{-i}' \in A_{-i}} u_i(a')$ → max. of my and others actions

$f^{maxmin}(a_i) = 1$ iff $a_i \in \arg\max_{a_i' \in A} \min_{a_{-i}' \in A_{-i}} u_i(a')$ → max of my own & min of others

$f^{eff}(a_i) = 1$ iff $a_i \in \arg\max_{a_i' \in A} \max_{a_{-i}' \in A_{-i}} \sum_{j \in N} u_j(a')$

$f^{fair}(a_i) = 1$ iff $a_i \in \arg\min_{a_i' \in A} \min_{a_{-i}' \in A_{-i}} \max_{j, j' \in N} |u_j(a') - u_{j'}(a')|$

$f^{unif}(a_i) = 1$ for all a_i