Formal Political Theory

Professor Keith Schnakenberg

Pol Sci 505

Overview

Introduction

Motivating example: The Presidential veto

- In the United States and many other places, the President must sign legislation passed by the legislature (or be overridden) before it becomes law.

- Does the presence of a Presidential veto substantially shape the laws that are eventually implemented?

From Chuck Cameron's <u>Veto Bargaining</u>:

> *Consider the extreme rarity of vetoes. Between 1945 and 1992, Congress presented presidents with over 17,000 public general bills.... From this flood of bills, presidents vetoed only 434. In other words, presidents in the post-war period vetoed only 25 public general bills per 1,000 passed....How can a weapon that is hardly ever used shape the content of important legislation under frequently occurring circumstances? (9)*

A simple bargaining model

To specify a model we must:

- Name the relevant players
- Specify what choices those players can make and how those choices lead to different outcomes
- Specify players' preferences over outcomes

Bargaining model

Players Congress and the President.

Choices and order of moves Congress moves first and chooses to pass a left-wing bill (L) or a moderate bill (M). The President moves second, sees what bill is chosen by Congress, and chooses to veto or not.

Outcomes If a bill is passed and not vetoed then that bill becomes law. If a bill is vetoed then it does not become law and the outcome is a status quo policy.

Preferences President prefers the moderate bill to the status quo but prefers the status quo to the left-wing bill. Congress prefers the left-wing bill to the moderate bill and the moderate bill to the status quo.

When does the President veto the bill?

- President can unilaterally kill any bill and implement the status quo.
- $\Rightarrow$ President's choice is between Congress's bill and the status quo.
- President's preferences:

$$\text{Moderate bill} > \text{SQ} > \text{Left-wing bill}$$

- *Rightarrow* President will use the veto when the left-wing bill is passed and not when the Moderate bill is passed.

What bill will Congress pass?

- Congress prefers the left-wing policy to the moderate policy and prefers the moderate policy to the status quo.
- Congress would therefore LIKE to pass the left-wing bill.
- However, Congress knows that the left-wing bill would be vetoed resulting in the status quo but the moderate policy would not be vetoed.
- $\Rightarrow$ Congress passes the moderate bill.

Some quick take-aways

- We predict NO VETOES yet the presence of a veto changes the final policy
- The reason is that Congress is thinking strategically and should try to anticipate the actions of the President
- We could complicate this much further (Congress might make errors, there might be conflicts within Congress, we could allow veto overrides) but this simple model makes this point nicely
- Broader point: We need a theory of the interaction we are studying in order to even begin to think about the meaning of the data we observe.

What are formal models?

This is two questions: What is a model? And, what makes a model formal?

Start with the first question

Models

- Loosely speaking a model is a representation of some target system which the analyst wishes to understand

Some types of models:

- Physical representation (a globe and a map are different models of the earth)
- Verbal analogy (e.g. Wlezien thermostatic model of public opinion)
- Mathematical system (e.g. the models we'll study in this class)

Models and accuracy

- Every model differs from its target system in some deliberate way.
  - Frictionless environments in physics
  - Two-species models in biology
  - Perfectly rational agents in a social scientific model

- A good model should also resemble the target system in some key ways, which will depend on the aspects of the target system the analyst wants to study

What makes a model formal?

Formal model:

- System of symbols and axioms
- Set of rules for manipulating them to derive results

Why?

- Anyone who applies the rules will derive the same results
- This is good! We'd like to accumulate and share knowledge
- We will study mathematical models. This is one type of formal model but not the only type (e.g. computational models are formal models).

Game theory

Most (not all) formal models in political science are game theoretic.

Game theory is the study of mathematical models of strategic interaction among two or more decision-makers.

Questions a game theorist asks about a problem

- Who are the relevant players?
- What decisions can these players make?
- What does each player know at the time that he or she must make a decision? What are the possible outcomes?
- What are players' preferences over these outcomes?

The answers to these questions will be used to create a mathematical system which we call a <u>game</u>.

The rules for making predictions from these games will come in the form of <u>solution concepts</u> which will be the main focus of our technical discussion.

- Questions about syllabus?
- Break.

Individual decision theory

A decision problem

Three components

- Actions that are available to the player
- Outcomes (i.e., consequences as the result of actions and possibly unforeseen/random events). Sometimes this is also referred to as alternatives.
- Preferences (i.e., how the player ranks the outcomes).

Actions

- Actions available to the decisionmaker is a set denoted by A.

- Example: A vote choice model
  - Voter can choose between voting for the Democrat, voting for the Republican, voting for a third-party candidate, or abstaining.
  - $A = \{d, r, t, a\}$

- Example: How much water to drink from a 1 gallon jug
  - Decisionmaker can choose any proportion of the jug from 0 to 1
  - $A = [0, 1]$

Outcomes

- Set of outcomes denoted by X
- Vote choice example:
  - $A = \{d, r, t, a\}$
  - Outcomes: the Democrat wins, the Republican wins, a third party candidate wins: $X = \{D, R, T\}$.

Preferences

- Preferences are defined over outcomes.

- Preferences defined as a binary relation over elements of X.

Binary relation?

- Technically, a set of ordered pairs of elements of X
- Relates any two elements of a set according to some criterion.
- Familiar example: Weak inequality relation "$\geq$" defined on the set of integers, where the expression $x \geq y$ is interpreted as "integer x is at least as big as the integer y."
- ...but back to preferences.

Preference relation

- relates any two outcomes in X as more or less desirable to the agent.
- Our notation: xRy is interpreted as "x is at least as good as y for the agent.
- May say "weak preference relation" xRy does not rule out indifference between x and y.

Preferences: Strict and weak

We can define strict preference and indifference using R:

$$xPy \text{ if and only if } xRy \text{ and not} yRx$$
$$xIy \text{ if and only if } xRy \text{ and } yRx.$$

Preference notation review page

Concept	My notation	Alternative notation	Analogy to numbers
Weak preference	R	$\succeq$	$\geq$
Strict preference	P	$\succ$	$>$
Indifference	I	$\sim$	$=$

Rationality

So far we have placed no restrictions on preferences whatsoever

Unfortunately, without more restrictions, things can get weird

Example: Things get weird

Consider our voter. The relevant outcomes are $X = \{d, r, t, a\}$ (democrat, republican, third party, abstain).

Consider the following strict preference relation:

$$\begin{aligned} dPr \\ rPt \\ tPa \\ aPd. \end{aligned}$$

What should the agent choose?

$$dPr$$
$$rPt$$
$$tPa$$
$$aPd.$$

Consider each possibility:

- d? This cannot be best because *aPd*
- r? no, *dPr*
- t? no, *rPt*
- okay then, a! oh no! *tPa*!

There is no best choice!

Okay fine, let's make some assumptions

We want to make a sufficient number of assumptions to guarantee that we can find a maximal choice for the decision-maker

Otherwise we cannot say what the decision-maker would do

We will collectively refer to these assumptions as rationality

Maximal outcomes

Definition

An outcome $x \in X$ is a maximal element of X with respect to the binary relation R if and only if xRy for all $y \in X$.

Reflexivity and completeness

Definition

A binary relation R on X is

1. Reflexive if for all $x \in X$, xRx.
2. Complete if for all $x, y \in X$ such that $x \neq y$, either xRy or yRx (or both).

We need reflexivity and completeness

We want to find assumptions that guarantee that there is always a maximal choice (i.e. there is a maximal choice given any subset of X).

It turns out reflexivity and completeness are necessary.

Our example satisfied reflexivity and completeness, so obviously these two assumptions will not be sufficient.

Lemma

If R violates reflexivity or completeness then there exists some $S \subseteq X$ such that there is no maximal element of S with respect to R.

Proof.

Part 1:
Suppose R violates reflexivity.
Then there exists some $x \in X$ such that xRx does not hold.
This implies that x is not a maximal element of any set
Therefore there is no maximal element for the subset $S = \{x\}$ of X.
Part two: Now suppose that R violates completeness.
Then there exist some $x, y \in X$ such that $\neg xRy$ and $\neg yRx$.
This immediately implies that there is no maximal element of the subset $S = \{x, y\}$ with respect to R. □

We have not yet captured what was wrong with out voting example.

In that example the voter's preferences were inconsistent in a problematic way: I should not strictly prefer d to r and r to t but then strictly prefer t to d.

The next conditions deal with this problem

Consistency conditions

Definition

A binary relation R on X is:

1. Transitive if for all $x, y, z \in X$, xRy and yRz implies xRz.
2. Quasi-transitive if for all $x, y, z \in X$, xPy and yPz implies xPz.
3. Acyclic if for all $\{x, y, z, \ldots, u, v\} \in X$, $xPy \& yPz \ldots \& uPv$ implies xRv.

The are ordered from strong to weak: Transitivity implies quasi-transitivity and acyclicity but not the reverse, quasi-transitivity implies acyclicity but not the reverse

Example: Quasi-transitive but not transitive

Example

Let $X = \{x, y, z\}$ with xPy, yIz, zIx. Then R is quasitransitive (vacuously). However, it is not transitive: yRz and zRx but not yRx. $\square$

Example: Acyclic but not quasitransitive

Example

Let $X = \{x, y, z\}$ with xPy and yPz and xIz. Then R is acyclic (since xRz) but not quasitransitive (since zRx). $\square$

Theorem

Let R be reflexive and complete and let X be finite; then the set of maximal outcomes is nonempty for all $S \subseteq X$ if and only if R is acyclic.

To prove this we need to prove both directions of the “if and only if”: that is, we need to show that acyclicity is sufficent for a maximal choice and also that it is necessary.

Proof of sufficiency

For any $S \subseteq X$ choose $x \in S$. If, for all $s \in S$, xRs then x is a maximal element and we are done.

Otherwise (since R is reflexive and complete) there must exist $y \in S \backslash \{x\}$ such that yPx.

If for all $s \in S$, yRs then y is a maximal element and again the we are done. Otherwise, there exists $z \in S \backslash \{x, y\}$ such that zPy. By acyclicity, it must be that zRx as well.

If for all $s \in S$, zRs then again we are done. Otherwise there exists $w \in S \backslash \{x, y, z\}$ such that wPz. By acyclicity this means wRy and wRx as well.

Because X (and hence S) is finite, we can continue this logic to conclude that there must exist an alternative weakly preferred to all other alternatives in S. $\square$

Proof of necessity

Let $x_1, x_2, \ldots, x_n$ be elements of X and assume

$$x_1 P x_2, x_2 P x_3, \ldots, x_{n-1} P x_n;$$

we wish to show that $x_1 R x_n$ if the set of maximal outcomes is nonempty.

Let $S = \{x_1, \ldots, x_n\} \subseteq X$ and suppose that the set of maximal elements of S with respect to R is nonempty.

Therefore. Because $x_{i-1} P x_i$ for $i = 2, \ldots, n$, we have that x_i is not maximal for any $i = 2, \ldots, n$.

Therefore, if the set of maximal elements is nonempty it must be the case that x_1 is maximal which implies in particular that $x_1 R x_i$ for all $i = 2, \ldots, n$; in particular $x_1 R x_n$ as required. $\square$

Rationality

- This result lays the foundation for a minimal set of requirements on preferences to guarantee that the agent has an optimal choice in any given situation.
- As shorthand, we will say an agent is <u>rational</u> if her preferences are reflexive, transitive, and complete.
- But we will typically make stronger assumptions to make the problem more tractable.

Utility

In practice, we would rather not work with binary relations all the time.

We therefore introduce utility, which lets us work with a mathematical function rather than an ordinal relation

Utility

A utility function is a function $u : X \to \mathbb{R}$ that assigns each outcome in X a real number

A utility function represents a preference relation R if the function assigns higher numbers to outcomes that are ranked higher under R

Definition

A utility function $u : X \to \mathbb{R}$ represents the preference relation R if for any pair $x, y \in X$, $u(x) \geq u(y)$ if and only if xRy.

Remarks on utility functions

1. No utility function uniquely represents a preference relation: since preference relations are only ordinal, any two functions preserving the same order represent the same preferences
   - e.g. if u represents the preference relation R, then so does any increasing transformation of u.

2. For the simplest settings with no uncertainty, the ordinal properties of the utility function will not matter for the solution.

3. We next turn to the question: When can preferences be represented by a utility function?

When can preferences be represented by a utility function?

Theorem

Let X be finite. Then any reflexive, complete, and transitive preference relation over X can be represented by a utility function.

Proof by construction

Let R be a reflexive, complete, and transitive preference relation over X.

Because the preference relation is complete and transitive, we can find a least-preferred outcome $\underline{x} \in X$ such that all other outcomes $y \in X$ are at least as good as $\underline{x}$.

Now define the “worst outcome equivalence set”, denote X_1, to include $\underline{x}$ and any other outcome y for which the player is indifferent between y and $\underline{x}$.

Proof by construction (continued)

Then, from the remaining elements $X \backslash X_1$, define the "second worst outcome equivalence set," X_2, and continue in this fashion until the "best outcome equivalence set" X_n, is created.

Because X is finite and R is reflexive, transitive, and complete, such a finite collection of n sets exists.

Now consider n arbitrary numbers such that $u_n > u_{n-1} > \cdots > u_2 > u_1$, and assign payoffs according the rule: for any $k \in \{1, \ldots, n\}$ and $x \in X_k$, $u(x) = u_k$. This payoff function represents the preference relation R and therefore we have proven that such a function exists. $\square$

Example: Two simple decision problems

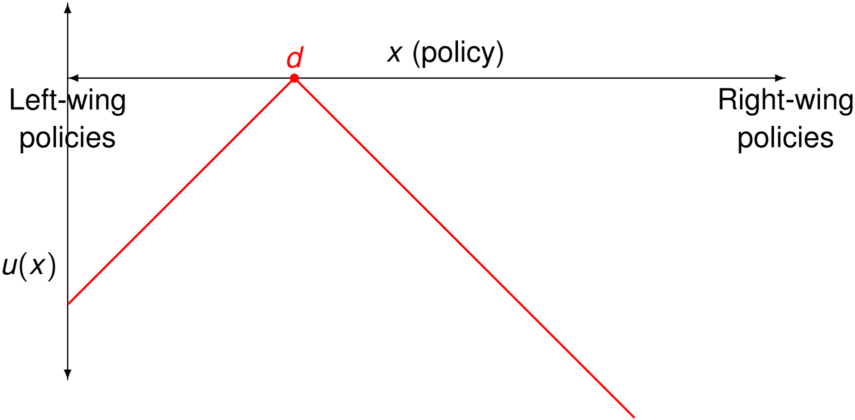
$X = A = \{a, b, c\}$

Utility function	$u(a) = 1, u(b) = 2, u(c) = 3$	$\tilde{u}(a) = 1, \tilde{u}(b) = \tilde{u}(c) = 3.$
Maximal elements	? $\{c\}$	? $\{b, c\}$

Example: Spatial policy choice

- Political scientists often represent policies and policy preferences as points on the real line.
- Let $x \in \mathbb{R}$ be a policy outcome. We think of smaller values of x as being more left-wing policies and larger values as being more right-wing policies.
- Let $d \in \mathbb{R}$ be the decisionmaker's ideal policy (often called an ideal point). This represents the policy the decisionmaker most prefers. The decisionmaker likes policies closer to d more than policies further away from d.
- We can represent these preferences with the utility function:

$$u(x) = -|x - d|$$



It's immediate that the decisionmaker's optimal policy choice is $x = d$.

Example: Allocating effort

- Consider a bureaucrat asked to allocate effort toward policy implementation
- The policy can be implemented poorly with little effort or implemented better with increasing effort
- The possible effort levels are the set of nonnegative real numbers (i.e. $X = [0, \infty)$.)
- The decisionmaker's preferences are represented by the following utility function:
$$u(x) = x - kx^2.$$

Example: Allocating effort (solution)

How do we solve for the maximal effort?

The first-order condition for a maximum is that the first derivative of the utility function with respect to x be equal to zero:

$$\frac{\partial u(x)}{\partial x} = 1 - 2kx = 0,$$

which is solved by $x = \frac{1}{2k}$.

The second-order condition to ensure that this is indeed a maximum is that the second derivative is negative:

$$\frac{\partial^2 u(x)}{\partial x^2} = -2k < 0,$$

which holds for all values of x since $k > 0$.

Next step: Decisionmaking under uncertainty

- So far we have only considered problems in which the agent faces no uncertainty

- In reality, decisionmakers are not always completely certain of the payoff consequences of their choices. We need to expand our framework to account for this.

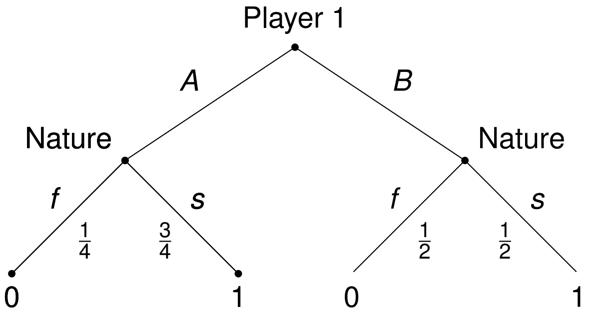
Motivating example: Choosing between two policies

- Two policies, A and B
- Policy A is successful with probability $\frac{3}{4}$ and unsuccessful with probability $\frac{1}{4}$. Policy B is successful with probability $\frac{1}{2}$ and unsuccessful with probability $\frac{1}{2}$.
- Payoff: 1 for a successful policy and 0 for an unsuccessful policy.

Motivating example: Choosing between two policies

- We can think of the decisionmaker as choosing between a lottery that offers a payoff of 0 with probability $\frac{1}{4}$ and of 1 with probability $\frac{3}{4}$ and another lottery that offers a payoff of 1 or 0 each with equal probabilities.

- We may think of this problem using a decision tree with moves by “Nature”:


Player 1
A
B
Nature
Nature
f
s
f
s
1/4
3/4
1/2
1/2
0
1
0
1

Introducting lotteries

Definition (Lottery over finite outcomes)

A simple lottery over outcomes $X = \{x_1, x_2, \ldots, x_n\}$ is defined as a probability distribution $p = (p(x_1), p(x_2), \ldots, p(x_n))$, where $p(x_k) \geq 0$ is the probability that x_k occurs and $\sum_{k=1}^{n} p(x_k) = 1$.

Lotteries over infinite sets

Preferences over lotteries

- Conceptually there is no problem with simply defining preferences over lotteries and imposing the same rationality requirements as before
- In practice, it is more useful for us to think about preferences over lotteries as extensions of preferences over fundamental outcomes, combined with beliefs about probabilities
- To this end, we will think of agents as maximized <u>expected</u> utility

Expected utility

Definition

Let $u(x)$ be the player's utility function over outcomes $X = \{x_1, x_2, \ldots, x_n\}$, and let $p = (p_1, p_2, \ldots, p_n)$ be a lottery over X such that $p_k = \Pr[x = x_k]$. Then we define the player's expected utility from the lottery p as

$$\mathbb{E}[u(x)|p] = \sum_{k=1}^{n} p_k u(x_k).$$

Continuous version

Solving the two uncertain policies problem

- Expected utility from choosing A:

$$\mathbb{E}[u(x)|A] = \Pr[s|A]u(s) + \Pr[f|A]u(f) = \frac{3}{4}1 + \frac{1}{4}0 = \frac{3}{4}.$$

- Expected utility from choose B is

$$\mathbb{E}[u(x)|B] = \Pr[s|B]u(s) + \Pr[f|B]u(f) = \frac{1}{2}1 + \frac{1}{2}0 = \frac{1}{2}.$$

- $\Rightarrow$ A is the maximal choice

Allocating effort with uncertainty

- The bureaucrat allocates effort equal to $a \in [0, 1]$.

- The policy is successful ($x = 1$) with probability $\Pr[x = 1|a] = a$ and unsuccessful ($x = 0$) with probability $1 - a$.

- Bureaucrat's payoff:
$$u(x, a) = x - ka^2$$

Solving for optimal effort level

- Expected utility from any effort level a:

$$\begin{aligned}\mathbb{E}[u(x,a)|a] =& a(1-ka^2)+(1-a)(0-ka^2)\\ =& a-ka^2.\end{aligned}$$

- Differentiating and solving for the FOC gives: $a=\frac{1}{2k}$.

- Accounting for the fact that effort is constrained to be in [0, 1]: the agent gives maximum effort of $a=1$ for $k\leq\frac{1}{2}$ and gives $a=\frac{1}{2k}$ if $k>\frac{1}{2}$.

Remarks on expected utility

- The cardinal properties of the utility function matter now!
- The properties of u capture risk preferences. For continuous action sets, concavity of u implies risk aversion, convexity of u implies risk acceptance.
- This is not an innocuous assumption and there are other possible ways to represent preferences over lotteries. Real people sometimes do violate expected utility.

Dynamic decision problems

- We have analyzed only static problems
- Sometimes, we are interested in how a decisionmaker makes multiple interconnected choices at different points in time
- We will look in these cases for an entire plan that maximizes expected utility at every step.

Policy then effort choice

First, decisionmaker chooses between two policies:

- Safe policy: Implemented with no effort
- Risky policy: Better than the safe policy when it's successful, worse when it is not. Effort helps increase the likelihood of success.

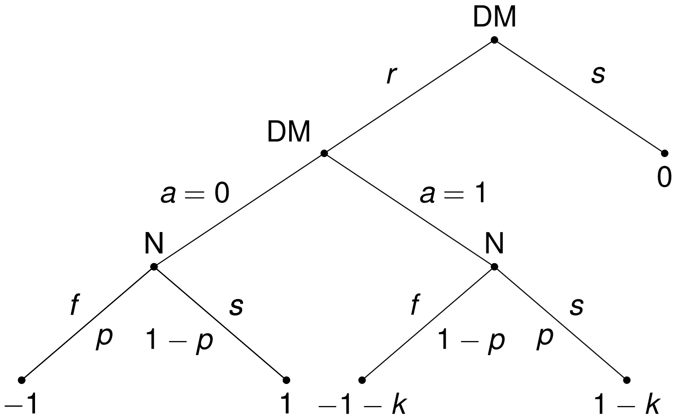
Second, decisionmaker chooses effort $a \in \{0, 1\}$.

The probability that the risky policy succeeds is $p > \frac{1}{2}$ if $a = 1$ and $1 - p$ if $a = 0$.

The policymaker's payoff is 0 from the safe policy and

$$u(r, a) = \begin{cases} 1 - ka & \text{if the policy succeeds} \\ -1 - ka & \text{if the policy fails} \end{cases}$$

from the risky policy, where $k > 0$ is a cost of effort.


DM
r
s
DM
0
a = 0
a = 1
N
N
f
s
f
s
p
1 − p
1 − p
p
−1
1
−1 − k
1 − k

Solving by backward induction

Start with the effort decision:

$$\mathbb{E}[u|r, a = 0] = p(-1) + (1 - p)1 = 1 - 2p.$$

$$\mathbb{E}[u|r, a = 1] = p(1 - k) + (1 - p)(-1 - k) = 2p - 1 - k.$$

$\Rightarrow$ policymaker makes an effort if

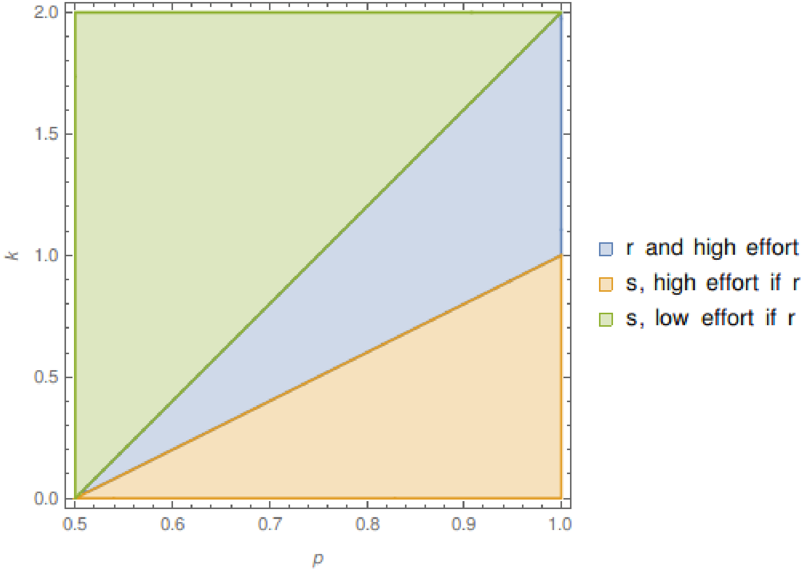
$$2p - 1 - k \geq 1 - 2p$$

Back one step: The policy decision

- Case 1: $k \leq 4p - 2$.
  - DM will make an effort if she chooses r, so her expected utility for choosing r is $2p - 1 - k$.
  - r is optimal in this case if $2p - 1 - k > 0$
  - $\Rightarrow$ if $2p - 1 < k \leq 4p - 2$ then the decisionmaker's optimal choice is s. If $k \leq 2p - 1$ then the optimal choice is r.

- Case 2: $k > 4p - 2$.
  - The agent would not make an effort after choosing r.
  - Then her expected utility for choosing r is $1 - 2p < 0$
  - Thus, the decisionmaker should always choose s in this case.

Solution pictured



Decision theory wrap up

- Rationality requirements: Acyclicity or transitivity, completeness, reflexivity
- Basic problems: Choose the maximal choice from a set
- Wrinkles: Uncertainty, dynamics

The rest of the class: Multiple decisionmakers (games)

Static games of Complete Information

A policy problem: Defending against natural disasters

Consider the mayor of a town: 

She must allocate resources to prepare for a potential natural disaster at two sites.

The sites:

		
	City Hall	The Lookout
Chances of disaster	75%	33%
Damages from disaster	50	200

Resource allocation lessens the damage in some way proportional to the amount of resources

- This is a decision-theoretic problem and is fairly easy to solve

- We know the probabilities and the outcomes, we just compute the expected damage and allocate funds to optimize the Mayor's objective function

For instance, if her objective is simply to minimize expected damages she should allocate all of the resources for the Lookout.

Related problem: Defending against a terrorist attack

Why is this problem so different?

Attacks are strategic!

The Mayor needs to think about what the bad kitties are going to do. Just as importantly...

The Mayor should assume that the cats are thinking about what she is going to do!

The solution must be different!

- Suppose Mayor Goodway devotes all of her resources to protecting the Lookout, meaning no damage will occur
- Then the Kitten Catastrophe Crew will for sure not attack the lookout and will instead target City Hall
- But then the original allocation of resources is no longer optimal!
- This is a game theoretic problem. Politics and policy is full of them. The rest of this class is about how to solve them.

Strategic form games

Definition

A game in strategic (or normal) form has three elements:

1. The set of players $N = \{1, 2, \ldots, n\}$, with a particular player denoted by $i \in N$.
2. A set of pure strategies S_i for each players, with the set of all pure strategies denoted $S = \{S_1, S_2, \ldots, S_n\}$.
3. A set of payoff functions $\{u_1, u_2, \ldots, u_n\}$ that give play i's von Neumann-Morgenstern utility $u_i(s)$ for every profile $(s_1, s_2, \ldots, s_n)$ of pure strategies.

Common knowledge

In a game of complete information, all aspects of the game are common knowledge

An event E is common knowledge if everyone knows E, everyone knows E, everyone knows that everyone knows E, and so on ad infinitum.

Important for how we understand games: I choose my actions anticipating that you understand the game, that you believe that I understand the game, that you believe that I believe that you understand the game, and so on.

Important example

A battle of wits

Example: Paper, Rock, Scissors

- Players: Player 1 and Player 2 ($N = \{1, 2\}$)
- $S_1 = S_2 = \{\text{Rock}, \text{Paper}, \text{Scissors}\}$
- Rock beats Scissors, Paper beats Rock, and Scissors beats Paper. Each player gets 1 for a win, -1 for a loss, and 0 for a draw.

Matrix Form

		Player 2		
		Rock	Paper	Scissors
Player 1	Rock	$0, 0$	$-1, 1$	$1, -1$
	Paper	$1, -1$	$0, 0$	$-1, 1$
	Scissors	$-1, 1$	$1, -1$	$0, 0$

Strategies

- A mixed strategy σ_i for player $i \in N$ is a probability distribution over i's pure strategies S_i.
- i.e. $\sigma_i(s_i)$ is the probability that i's strategy assigns to s_i.
- Let Σ_i be the set of all of player i's mixed strategies

- Each pure strategy can be represented as a degenerate case of a mixed strategy: a strategy in which i always plays some action s is simply a mixed strategy assigning $\sigma_i(s) = 1$ and $\sigma_i(s') = 0$ for all $s' \neq s$.

Back to Rock, Paper, Scissors

Examples of mixed strategy profiles:

$$\begin{aligned}\sigma_1(\text{Rock}) =& \frac{1}{3}, \sigma_1(\text{Paper}) = \frac{1}{3}, \sigma_1(\text{Scissors}) = \frac{1}{3}\\ \sigma_2(\text{Rock}) =& \frac{1}{2}, \sigma_2(\text{Paper}) = \frac{1}{4}, \sigma_2(\text{Scissors}) = \frac{1}{4}.\end{aligned}$$

$$\begin{aligned}\sigma_1(\text{Rock}) =& 1, \sigma_1(\text{Paper}) = 0, \sigma_1(\text{Scissors}) = 0\\ \sigma_2(\text{Rock}) =& 0, \sigma_2(\text{Paper}) = 0, \sigma_2(\text{Scissors}) = 1,\end{aligned}$$

(ok just to write $s_1 = \text{Rock}, s_2 = \text{Scissors}$.)

Payoffs and expected payoffs given strategies

- $u_i(s)$ is player i's payoff when every player's action follows the strategy profile s.
  - e.g. $s = (\text{Rock}, \text{Scissors})$ is a particular strategy profile in the RPS game, for which we have $u_1(s) = 1$ and $u_2(s) = -1$

- Convention: For a profile s, let $s = (s_i, s_{-i})$ for some player i, where $s_{-i} \in S_{-i}$ is the strategy used by every player other than i

- Hence, $u_i(s) = u_i(s_i, s_{-i})$

- Similarly, (σ_i, σ_{-i}) can denote a particular mixed strategy profile

Expected payoffs from arbitrary mixed strategies

Player i's expected utility from choosing the pure strategy $s_i \in S_i$ when her opponents choose the mixed strategy $\sigma_{-i} \in \Sigma_{-i}$ is

$$
\begin{aligned}
U_i(s_i, \sigma_{-i}) &= \sum_{s \in S_{-i}} \Pr[s|\sigma_{-i}] u_i(s_i, s_{-i}) \\
&= \sum_{s \in S_{-i}} \left(\prod_{j=1}^{n} \sigma_j(s_j) \right) u_i(s_i, s_{-i}).
\end{aligned}
$$

Expected payoffs from arbitrary mixed strategies (cont)

Player i's expected utility from playing the pure strategy σ_i when her opponents play σ_{-i} is therefore

$$U_i(\sigma_i, \sigma_{-i}) = \sum_{s_i \in S_i} \sigma_i(s_i) U(_i s_i, \sigma_{-i}), \tag{1}$$

by the law of iterated expectations.

Back to rock, paper scissors

Consider this strategy:

$$\sigma_1(\text{Rock}) = \frac{1}{3}, \sigma_1(\text{Paper}) = \frac{1}{3}, \sigma_1(\text{Scissors}) = \frac{1}{3}$$
$$\sigma_2(\text{Rock}) = \frac{1}{2}, \sigma_2(\text{Paper}) = \frac{1}{4}, \sigma_2(\text{Scissors}) = \frac{1}{4}.$$

$$\begin{aligned} U_1(\text{Rock}, \sigma_2) &= \sigma_2(\text{Rock})0 + \sigma_2(\text{Paper})(-1) + \sigma_2(\text{Scissors})1 \\ &= \frac{1}{2}0 + \frac{1}{4}(-1) + \frac{1}{4}(1) \\ &= 0, \end{aligned}$$

$$\begin{aligned} U_1(\text{Paper}, \sigma_2) &= \sigma_2(\text{Rock})1 + \sigma_2(\text{Paper})0 + \sigma_2(\text{Scissors})(-1) \\ &= \frac{1}{2}1 + \frac{1}{4}(0) + \frac{1}{4}(-1) \end{aligned}$$

Which strategies are better?

We now need to develop a notion of how agents should choose strategies

The choice is not obvious because we have to take into account beliefs about what other players will do, knowing that these other players are also trying to choose the best strategies

A set of conditions that allow us to choose the best strategy of every player will be called a solution concept. We will work through several.

Dominance

- We will start with a very strong way in which one strategy might be better than another

- We do this by assuming nothing about what *i* believes the other players will do: can we still say that some strategies are better than others?

- Sometimes, yes: we will say a strategy <u>dominates</u> another one if that strategy is better no matter what the other players do

Dominance in pure strategies

Definition

Let $s_i \in S_i$ and $s_i' \in S_i$ be possible strategies for player i. s_i' is strictly dominated by s_i if

$$u_i(s_i, s_{-i}) > u_i(s_i, s_{-i})$$

for all $s_{-i} \in S_{-i}$.

Note: We may also refer to weak dominance, which occurs when the inequality above holds weakly for all $s_{-i} \in S_{-i}$ and strictly for some particular s_{-i}.

Remarks on dominance

- Dominance avoids the difficulties associated with beliefs about other players' strategies: a rational player should never play a dominated strategy, regardless of what they believe the other players will do
- Disadvantage: Often we will have a pair of strategies for which no strategy strictly dominates the other (consider rock, paper, scissors), in which case dominance does not make a prediction

Strictly dominant strategies

Definition

$s_i \in S_i$ is a strictly dominant strategy if every other strategy of i is strictly dominated by it:

$$u_i(s_i, s_{-i}) > u_i(s_i', s_{-i})$$

for all $s_i' \in S_i$ and for all $s_{-i} \in S_{-i}$.

In words: Regardless of what the other players do, s_i is the best choice

Dominant strategy equilibrium

This leads us to our first solution concept for games:

Definition

The strategy profile $s^D \in S$ is a strict dominant strategy equilibrium if $s_i^D \in S_i$ is a strict dominant strategy for all $i \in N$.

Example: Prisoner's Dilemma

The story:

- Two people are taken in for questioning
- Each can stay quiet ("cooperate" with the other person) or provide evidence against the other person ("defect")
- There is not enough evidence to convince on the principal charge, but enough to convict each on a lesser charge
- Both people offered a deal: defect on the other and avoid jail
- If both defect, the testimony is no longer needed so both are convicted, perhaps with some slight leniency in sentencing for helping prosecutors
- If one defects and the other cooperates, the cooperator gets the maximum sentence and the other gets off

A matrix form for the prisoner's dilemma

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$-1, -1$	$-3, 0$
	Defect	$0, -3$	$-2, -2$

Dominance in the PD

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$-1, -1$	$-3, 0$
	Defect	$0, -3$	$-2, -2$

Does Player 1 have a dominant strategy?

If Player 2 cooperates, Player 1 is better off defecting.

If player 2 defects, Player 1 is better off defecting.

That is, Player 1 is better off defecting no matter what Player 2 does. A dominant strategy!

Notice that the same is true for Player 2.

If Player 1 cooperates, Player 2 is better of defecting.

If Player 1 cooperates, player 2 is better off defecting

Therefore, Player 2 also has a dominant strategy to defect.

Thus, (Defect, Defect) is a strict dominant strategy equilibrium: for both players, Defect is a strictly dominant strategy.

Mixed strategy dominance

- We defined dominance for pure strategies but we should extend the idea to mixed strategies
- Fact 1: If s_i' is strictly dominated by s_i then any mixed strategy that assigns positive probability to s_i' is also strictly dominated.
- Fact 2: A strategy s_i' may be strictly dominated by some mixed strategy even if it is not strictly dominated by any pure strategy.

Mixed strategy dominance example

		Player 2		
		L	C	R
Player 1	U	$5,1$	$1,4$	$1,0$
	M	$3,2$	$0,0$	$3,5$
	D	$4,3$	$4,4$	$0,3$

No pure strategy is dominated by any other pure strategy (check)
For player 2, I claim that L is dominated by $(0, \frac{1}{2}, \frac{1}{2})$ (randomizing evenly between C and R).
We can check this for each pure strategy of player 1:

	L	$(0, 1/2, 1/2)$	
U	1	$\frac{1}{2}4 + \frac{1}{2}0 =$	2
M	2	$\frac{1}{2}0 + \frac{1}{2}5 =$	2.5
D	3	$\frac{1}{2}4 + \frac{1}{2}3 =$	3.5

Dominated strategy

Definition

Let $\sigma_i \in \Delta S_i$ and $s_i' \in S_i$ be possible strategies for player i. We say s' is strictly dominated by σ_i if

$$u_i(\sigma_i, s_{-i}) > u_i(s_i', s_{-i})$$

for all $s_{-i} \in S_{-i}$.

s_i' is a strictly dominated strategy if there exists a strategy $\sigma_i \in \Delta S_i$ such that σ_i strictly dominates s_i'.

Iterated Elimination of Strictly Dominated Strategies

- Strict dominated strategy equilibrium has the same strength and weakness: it relies only on player rationality
- This is a strength because we do not have to make strong assumptions about what players believe
- This is a weakness because it very often means we cannot make a prediction
- Furthermore: if players are rational, we might think that players should believe that other players are rational. This lets us take one step further: players will not play dominated strategies AND they will not believe that other players will play dominated strategies.

Iterated Elimination of Strictly Dominated Strategies

- Adding common knowledge of rationality suggests an iterative procedure that we can use to eliminate strategies (called IESDS):
  1. Eliminate strictly dominated strategies from the original game
  2. Consider the new game formed after eliminating those strategies: are there any strictly dominated strategies? If so delete them.
  3. Continue this process until there are no strictly dominate strategies remaining.

- Any strategy profile that survives IESDS is called an iterated-elimination equilibrium.

IESDS example

		Player 2		
		Left	Middle	Right
Player 1	Up	$1,0$	$1,2$	$0,1$
	Down	$0,3$	$0,1$	$2,0$

		Player 2	
		Left	Middle
Player 1	Up	$1,0$	$1,2$
	Down	$0,3$	$0,1$

		Player 2	
		Left	Middle
Player 1	Up	$1,0$	$1,2$

		Player 2
		Middle
Player 1	Up	$1,2$

Best response

- Both of the solution concepts we have used so far are based on eliminating actions that players should never play

- Alternatively we might ask: What strategies might players choose to play and under what conditions?

- To move in that direction we ask: What strategy should a player choose if she believes the other players are playing some strategy σ_{-i}? This is the concept of best response.

- For now, this will be a preview for next week.

Supplemental slides

Lotteries for infinite sets

Definition

A simple lottery over an interval $X = [\underline{x}, \overline{x}]$ is given by a cumulative distribution function $F : X \to [0, 1]$ where $F(\hat{x}) = \Pr[x \leq \hat{x}]$ is the probability that the outcome is less than or equal to $\hat{x}$.

Back to finite lotteries

Continuous expected utility

Statistics review:

- A density function f for a continuous random variable is a function satisfying

$$\Pr[a \leq x \leq b] = \int_a^b f(x)dx. \tag{2}$$

  That is, the probability that a random variable x falls into the interval $[a, b]$ is found by finding the area under the density curve between points a and b.

- $\Rightarrow$ CDF can be defined as

$$F(\hat{x}) = \int_{\underline{x}}^{\hat{x}} f(x)dx \tag{3}$$

  if $\underline{x}$ is the smallest possible value of x.

- $\Rightarrow$ if F is differentiable then

$$f(x) = \frac{dF(x)}{dx}. \tag{4}$$

Continuous expected utility

Definition

Let $u(x)$ be the player's payoff function over outcomes in the interval $X = [\underline{x}, \overline{x}]$ with a lottery given by the cumulative distribution $F(x)$, with density $f(x)$. Then we define the player's expected payoff as

$$\mathbb{E}[u(x)] = \int_{\underline{x}}^{\overline{x}} f(x)u(x)dx.$$

Back to finite expected utility

Example: Two policies with continuous uncertainty

- Decisionmaker must again choose between two policies A and B.
- The set of possible outcomes from each policy choice is $X = [0, 1]$.
- Preferences: $u(x) = x$.
- The lotteries induced by each policy are CDFs: $F_A(x) = x^2$ and $F_B(x) = x$ for $0 \leq x \leq 1$.

Expected utility for policy A

- First we can derive the PDF:

$$f_A(x) = \frac{dF_A(x)}{dx} = 2x.$$

- The expected utility of choosing policy A is therefore:

$$\mathbb{E}[u(x)|A] = \int_0^1 f_A(x)xdx = \int_0^1 2xxdx = \frac{2}{3}.$$

Expected utility for policy B

$$f_B(x) = \frac{dF_B(x)}{dx} = 1.$$

$$\mathbb{E}[u(x)|B] = \int_0^1 f_B(x)xdx = \int_0^1 1xdx = \frac{1}{2}.$$

$\Rightarrow$ A maximizes expected utility.

Example: Choosing effort with continuous uncertainty

- We will think of a politician choosing effort to produce good outcomes in order to gain reelection
- The politician chooses a level of effort $a \in A = [0, 1]$.
- The politician's vote share in the next election is drawn from a uniform distribution on the interval $[0, a]$. (Thus, the pdf of outcomes is $f(x) = 1/a$ for $x \in [0, a]$)
- Payoff is:

$$u(x, a) = \sqrt{x} - a.$$

where x is vote share

Solving the politician's problem

Expected utility for an effort level a:

$$\mathbb{E}[u(x,a)|a] = \int_0^a \frac{\sqrt{x}}{a} dx - a \tag{5}$$

$$= \frac{2}{3}\sqrt{a} - a. \tag{6}$$

Maximizing with respect to a. FOC:

$$\frac{1}{3\sqrt{a}} = 1$$

$\Rightarrow a = \frac{1}{9}$.

SOC: $-\frac{1}{6a^{\frac{3}{2}}} < 0$ ✓