Lecture 3: Decision Trees

Yeganeh M. Hayeri

SYS 660

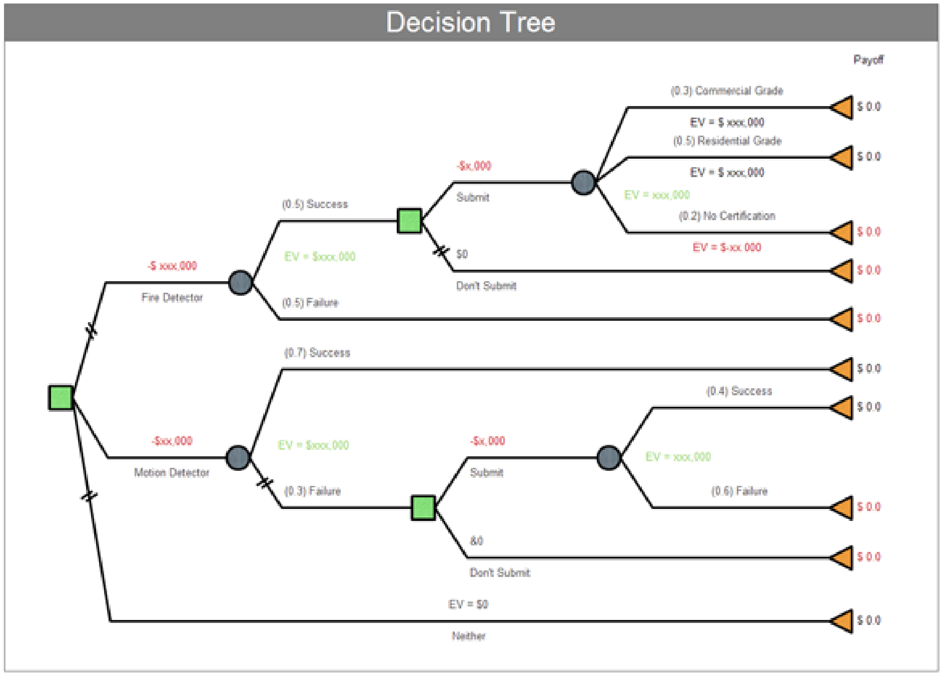
Spring 2019

Today's Lecture

- Decision trees
- Tradeoffs under certainty

Decision Trees

- ***Expected value*** is calculated as the sum of all possible values each multiplied by the probability of its occurrence.
- A ***decision tree*** is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs and utility.



Source: What is a decision tree algorithm? – Dictionary - Wikipedia

Structuring Decisions

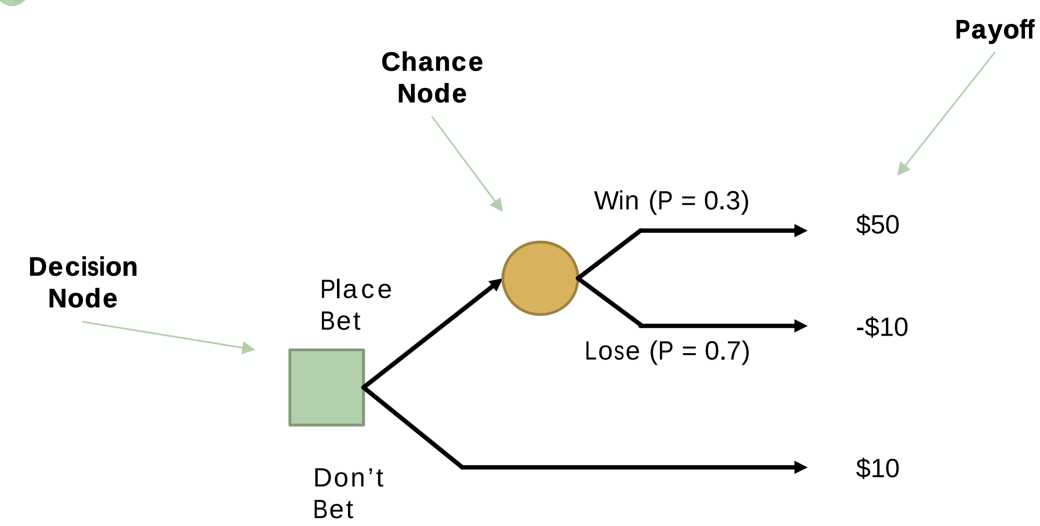
- All about objectives (what you want to achieve)
- Decision: choice between options
  - There is always an option, including status quo
  - Waiting for more information is an option
- Uncertainty: always exists
- Outcomes: possible results of uncertain events.
- Many uncertain events lead to complexity.

Expected Value

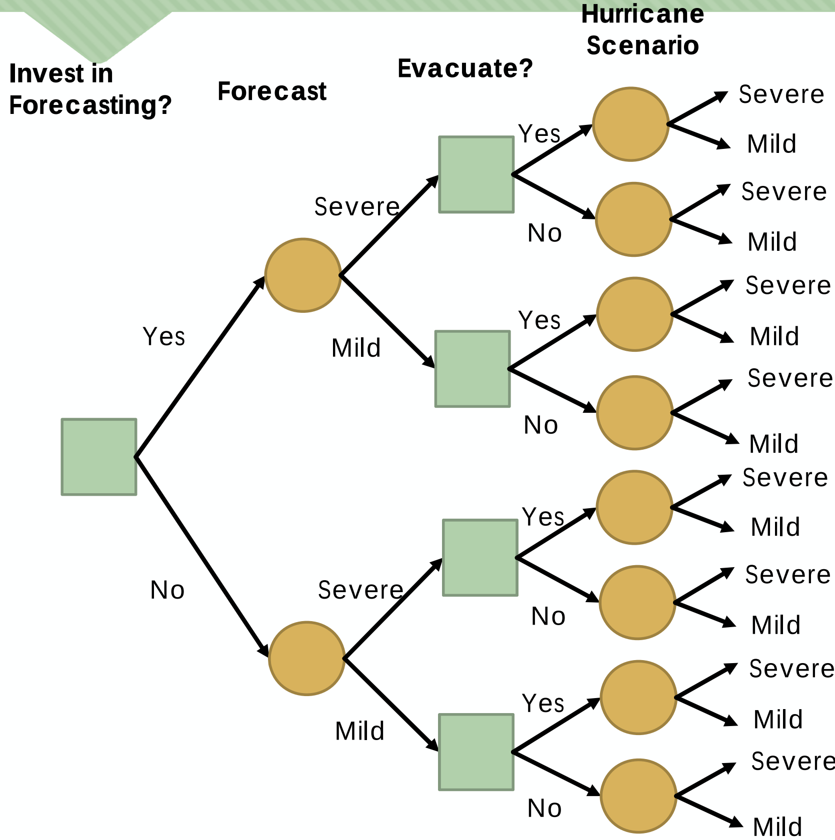
- Expected value is calculated as the sum of all possible values each multiplied by the probability of its occurrence.
- Outcome: x_1, x_2, x_3, …
- Probability: p_1, p_2, p_3,…
- Expected value: $x_1p_1+x_2p_2+x_3p_3+$…….

Decision trees and expected values can serve as effective tools for your decision analysis.

Elements of a Decision Tree



Example Decision Tree



Loss of Life	Damage	Activity Loss
Low	Med	High
Low	Low	Med
High	High	High
Low	Low	Low
Low	Med	High
Low	Low	Med
High	High	High
Low	Low	Low
Low	Med	High
Low	Low	Med
High	High	High
Low	Low	Low
Low	Med	High
Low	Low	Med
High	High	High
Low	Low	Low

Characteristics of a Decision Tree

- All decisions and events are sequenced.
- All possible decision options are exhaustively included.
- All possible outcomes for chance events are exhaustively included.
- All possible combinations of outcomes against the objectives hierarchy are represented by the branches at the end of the tree.
- A decision tree is a form of dynamic program and can be solved recursively to find the “optimal” policy.

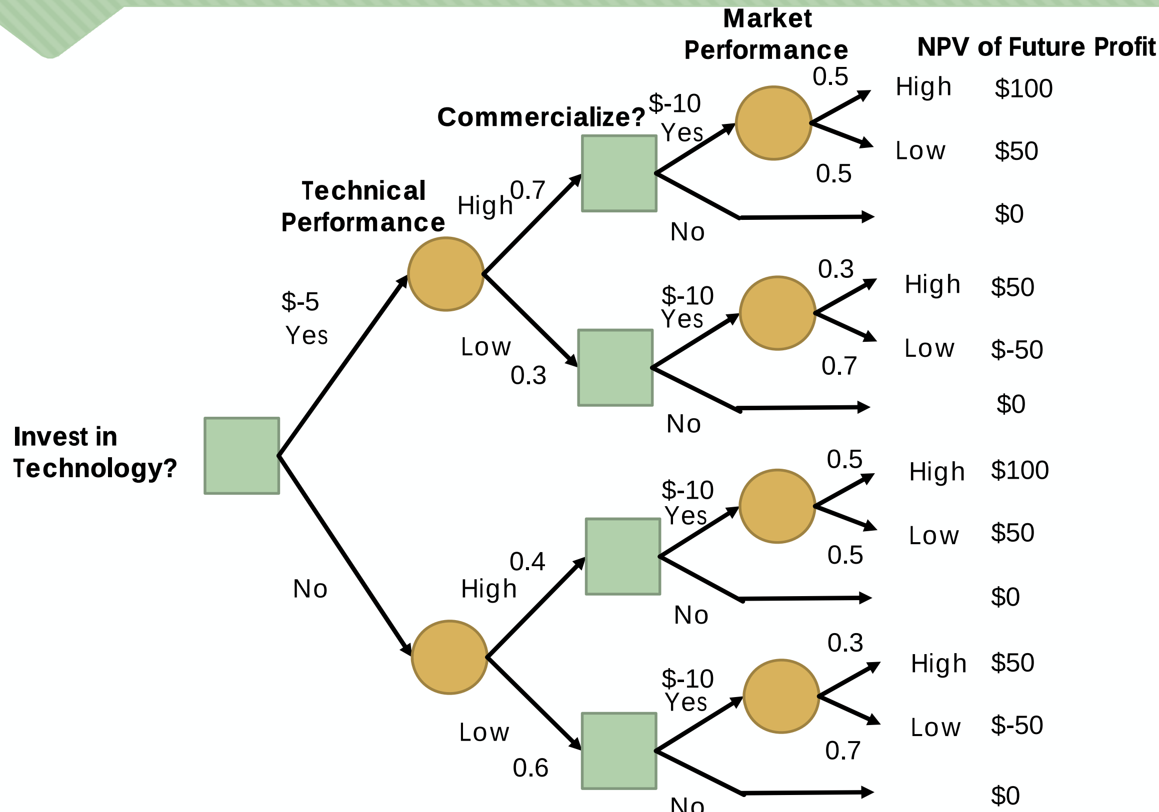
Comparison

- Properly constructed influence diagrams and decision trees are isomorphic
  - It is possible to convert one into the other
- Decision trees contain more detail but have a tendency to “explode” graphically
- Each has advantages and disadvantages
  - Influence diagrams may be better for communicating with others
  - Decision trees may be better for detailed analysis like sensitivity analysis
- One possible approach is to start with an influence diagram to explore the decision context and then convert it to a decision tree for detailed analysis

Solving decision Trees

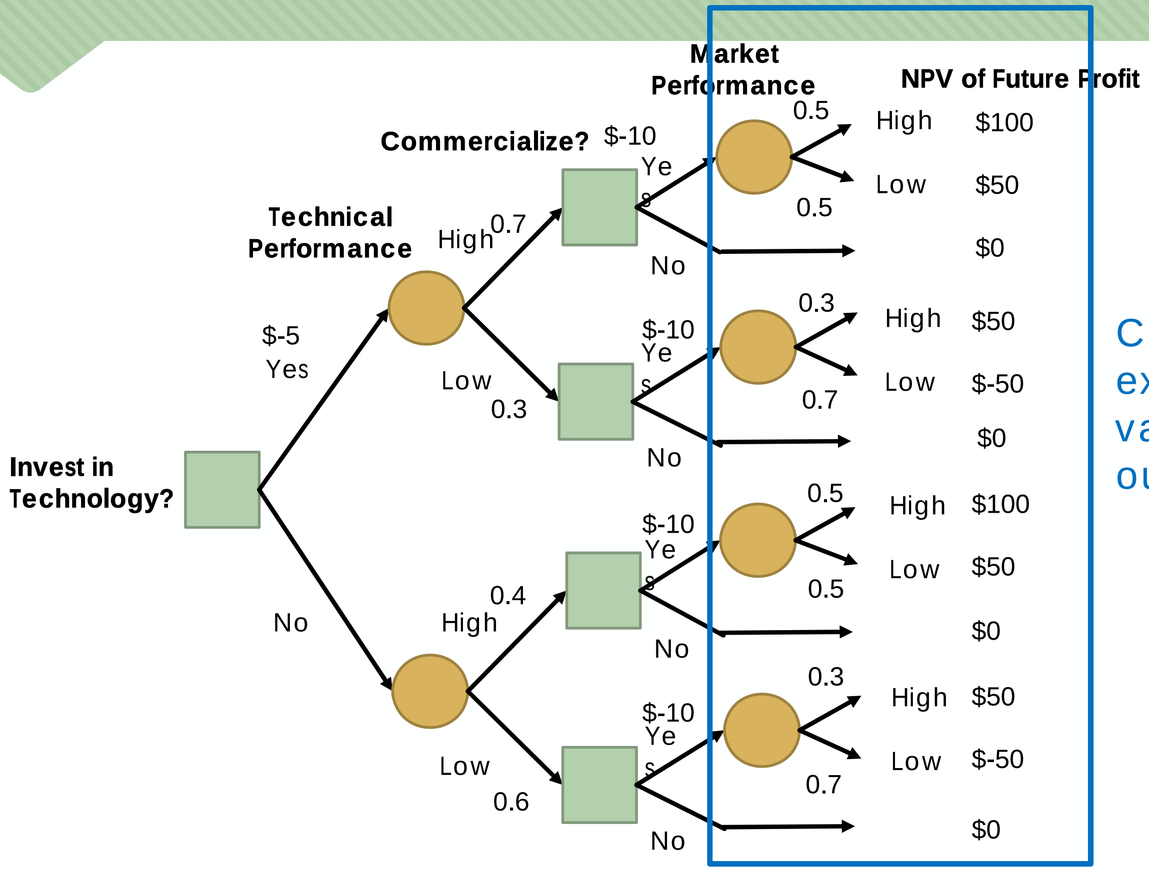
- Decision trees are a form of dynamic programming
- They can be solved by “rolling back” the tree
- This means starting at end of the tree and working back to the start
- In other words, solve the last decision first for each possible situation
- This determines the expected outcome for each of those situations
- Repeat the process for the next to last decision and so on until you reach the first decision
- The result will be an optimal policy for each possible path through the tree

Example: Invest in Technology



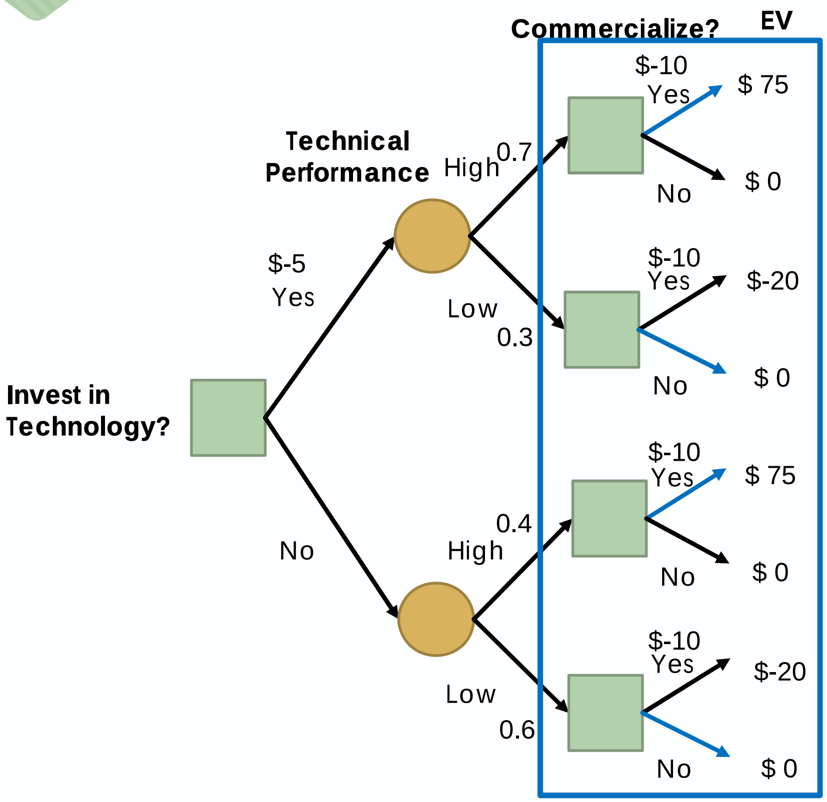
Assumption: All $ values are in present value

Example: Invest in Technology



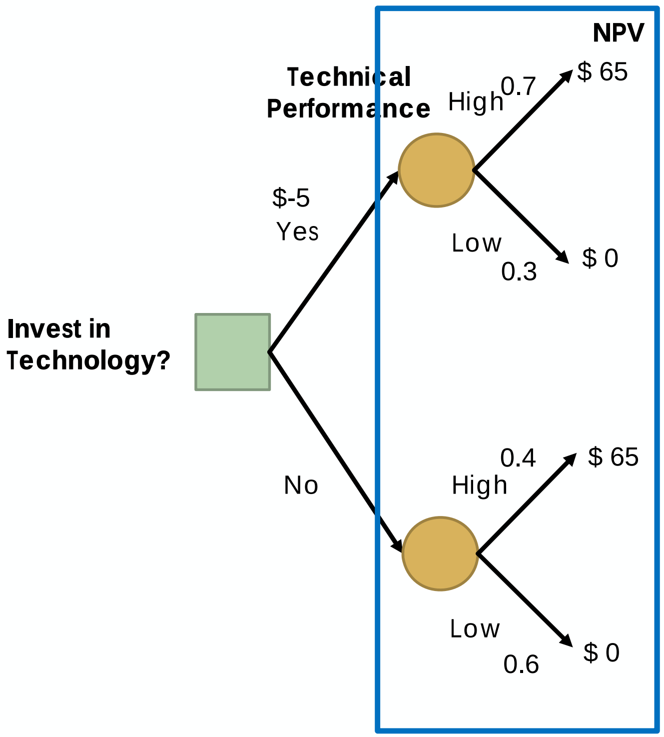
Calculate the expected values of the outcomes

Example: Invest in Technology



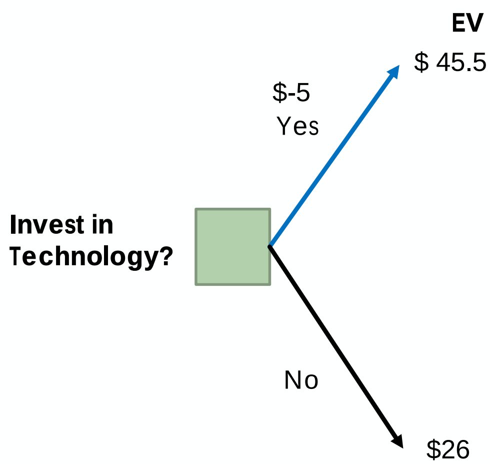
Select the best option at each decision node

Example: Invest in Technology



Calculate the expected values of the outcomes

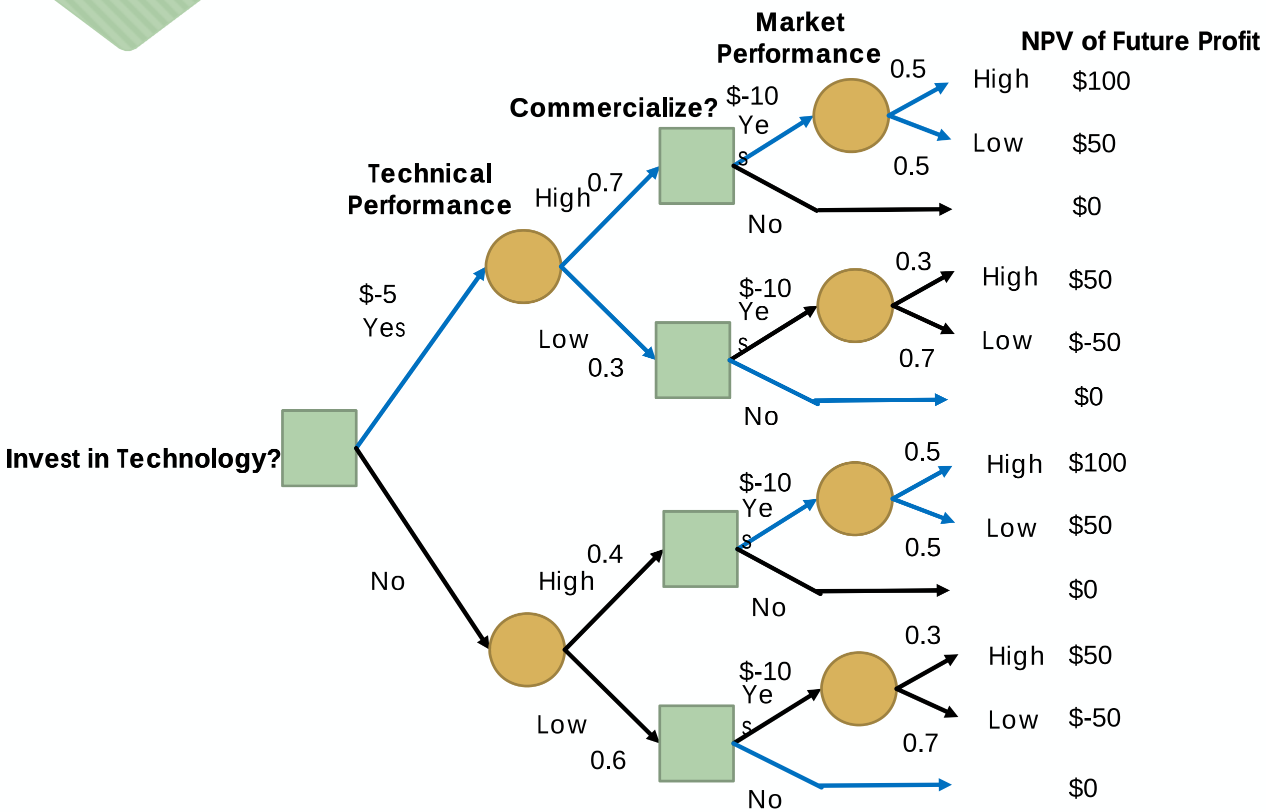
Example: Invest in Technology



Select the best option

You should invest in the technology for an expected NPV of $40.5

Example: Optimal Policy



Objective Measurement Scales

- Not every objective can be measured in dollars
- Ideally, you would define one numerical measurement scale for each objective
  - These are called attributes
  - Think back to the last lecture, what type of measurement scale would we prefer?
- What if there is no natural measurement scale for a particular objective?
- Option1: Choose a surrogate measure
  - E.g., maximize student test scores in place of maximize student learning
- Option 2: Construct an ordinal attribute scale
  - E.g., define descriptive characteristics for poor performance, adequate performance, and good performance

Dealing with Multiple Objectives

What happens when a decision has multiple objectives and each is assessed on a different measurement scales?

We call these objectives non-commensurate.

Making tradeoffs among non-commensurate objectives is the core mission of decision analysis.

Decision Tree – In class Example

Alternatives

Decision node

Bring umbrella

Leave umbrella

Conditions

Chance node

Rain

No Rain

Outcomes

Metric

Minutes of Lost work

Tradeoffs

- Making tradeoffs is the heart of decision analysis.
- Tradeoffs are necessary when you have multiple, conflicting objectives such that improving performance against one objective necessitates reducing it against another (e.g. maximize performance, minimize cost)
- Tradeoffs are what make decisions hard.
- There is no one, objectively correct answer; the best decision is the one that best satisfies the decision maker's values.

Example

Choosing a car:

Imagine that you are looking to buy a new car and you care about two things:

1- You like to drive fast.

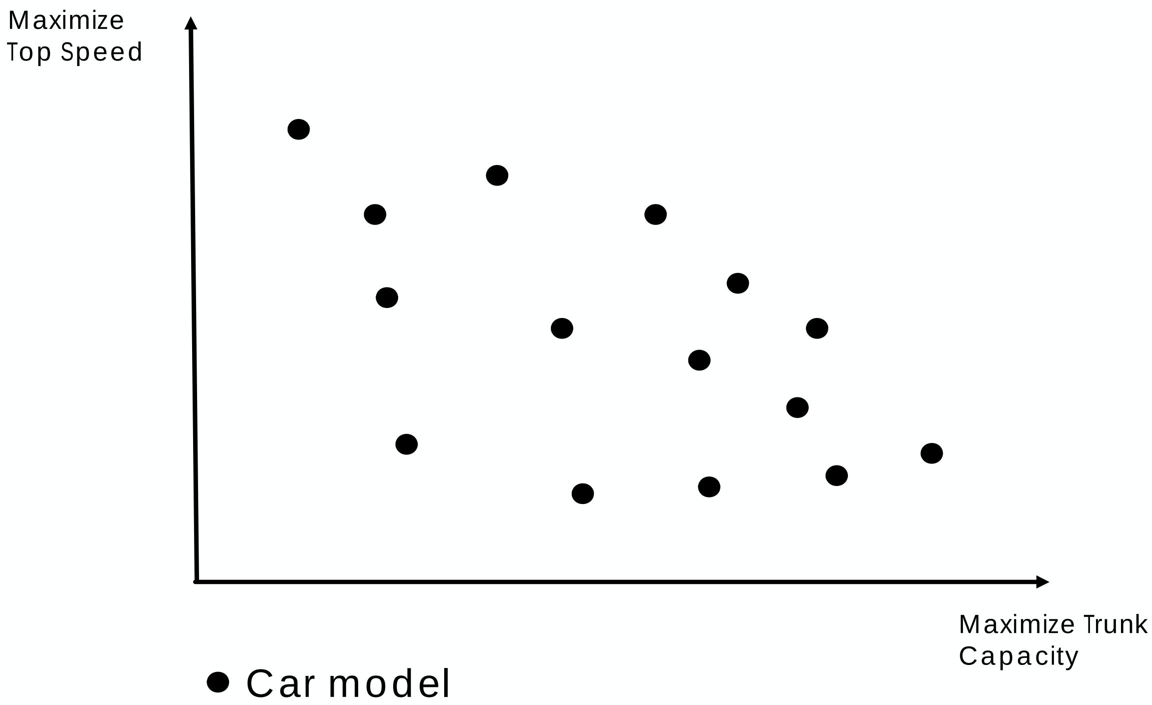
2- You also take a lot of long road trips, so you need a big trunk.

You would expect to make a tradeoff when choosing a car because, typically the car with the largest trunk is not also the fastest car.

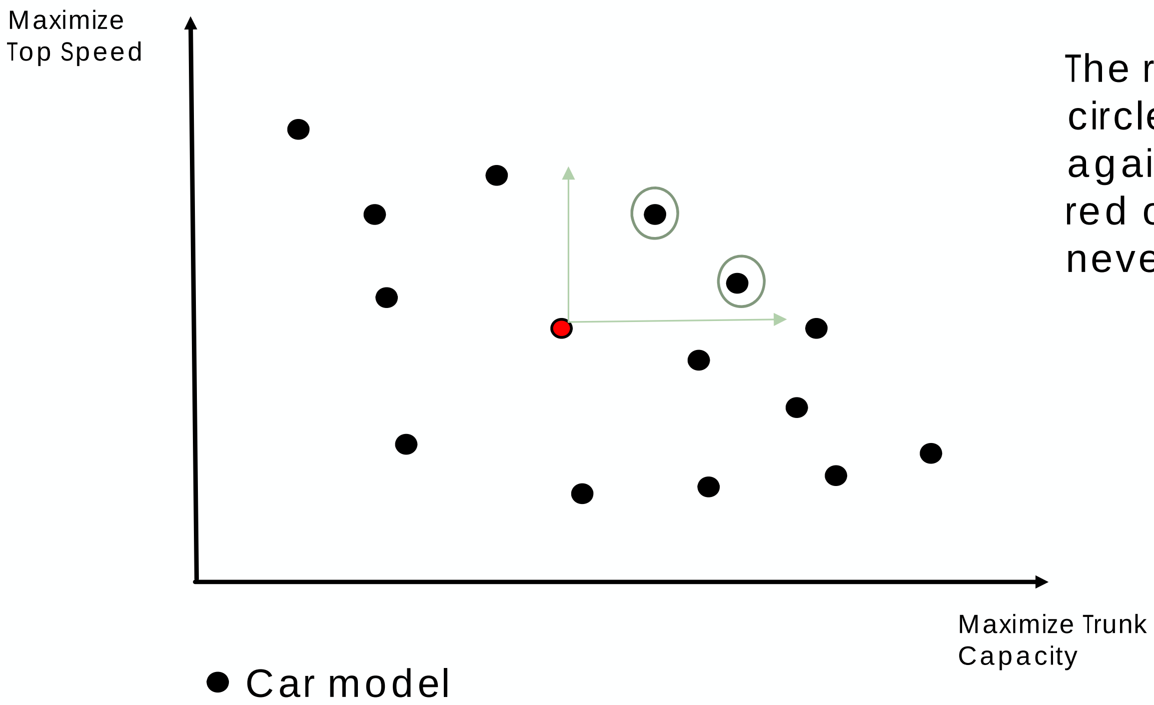
So you have 2 conflicting objectives: 1) Maximizing speed, 2) Maximizing trunk capacity.

The first thing you decide to do is plot all of your car options against both objectives.

Example – Cont.

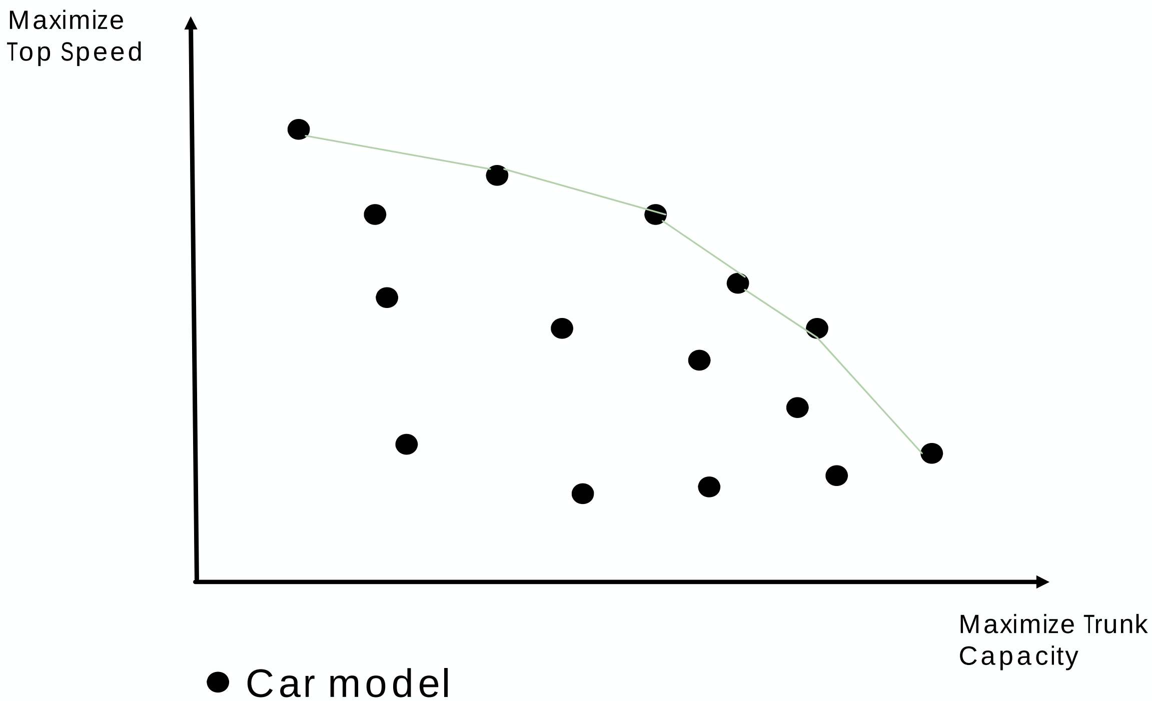


Example – Cont.



The red option is dominated by the circled options because both are better against both objectives. Therefore, the red option is dominated, and you would never choose it.

Example – Cont.

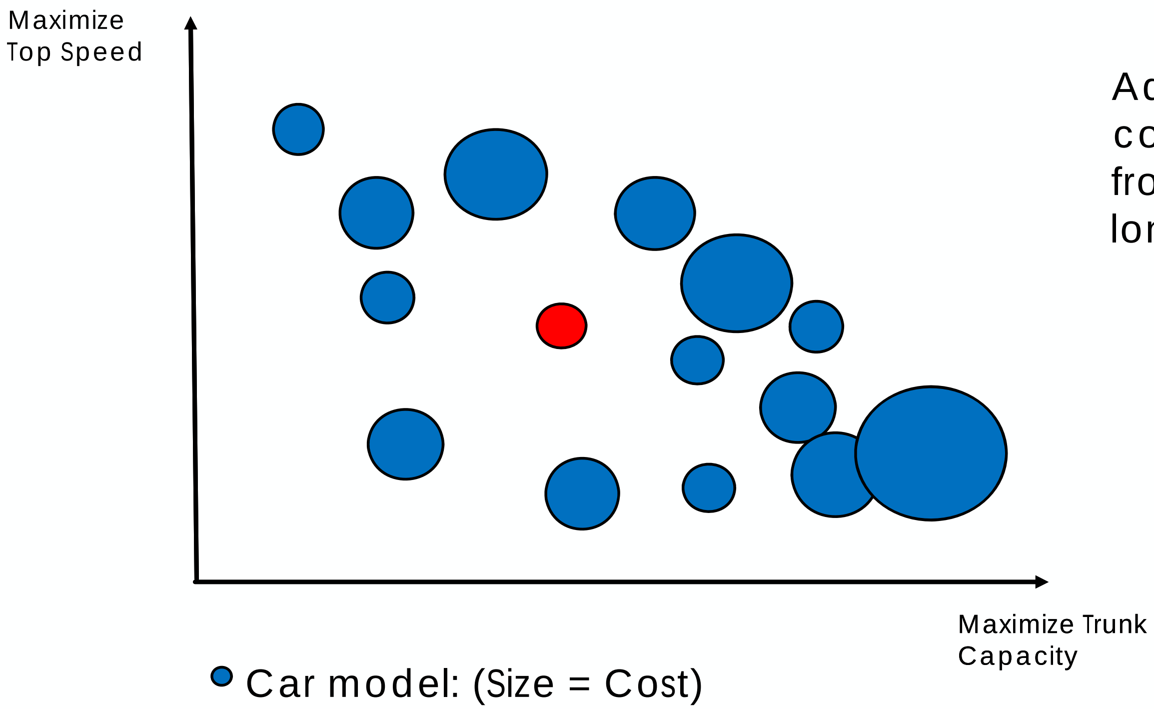


The set of non-dominated options Constitute the **Pareto Optimal (efficient) Frontier.** To choose an option that is not on The frontier is inefficient.

the **Pareto frontier** or **Pareto set** is the set of parameterizations (allocations) that are all Pareto efficient.

Pareto efficiency or **Pareto optimality** is a state of allocation of resources from which it is impossible to reallocate so as to make any one individual or preference criterion better off without making at least one individual or preference criterion worse off.

Example – Cont.



Adding a third objective, minimize cost, changes the Pareto optimal frontier. Not that the red option is no longer dominated.

Value Functions

How do you choose a solution from the Pareto Optimal frontier, particularly when there are multiple objectives?

Given a set of alternatives $a=[a_1, a_2, a_3, \ldots a_n]$

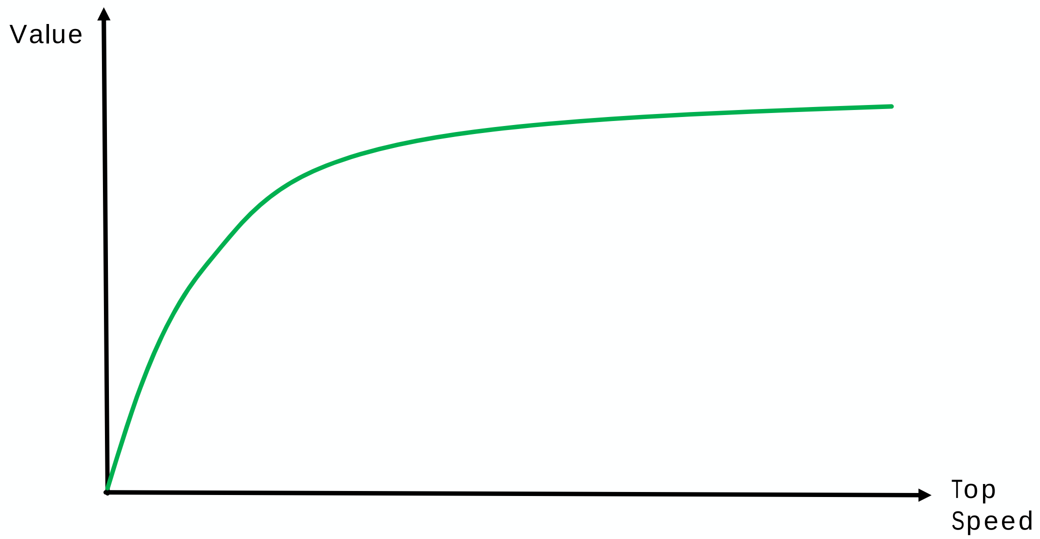
With attribute values $x_i = [x_{1i}, x_{2i}, x_{3i}, \ldots x_{mi}]$ for alternative a_i

Choose the alternative a_i such that it maximizes the value function $V(x_i)$

In standard notation: $\max V(x_i)$

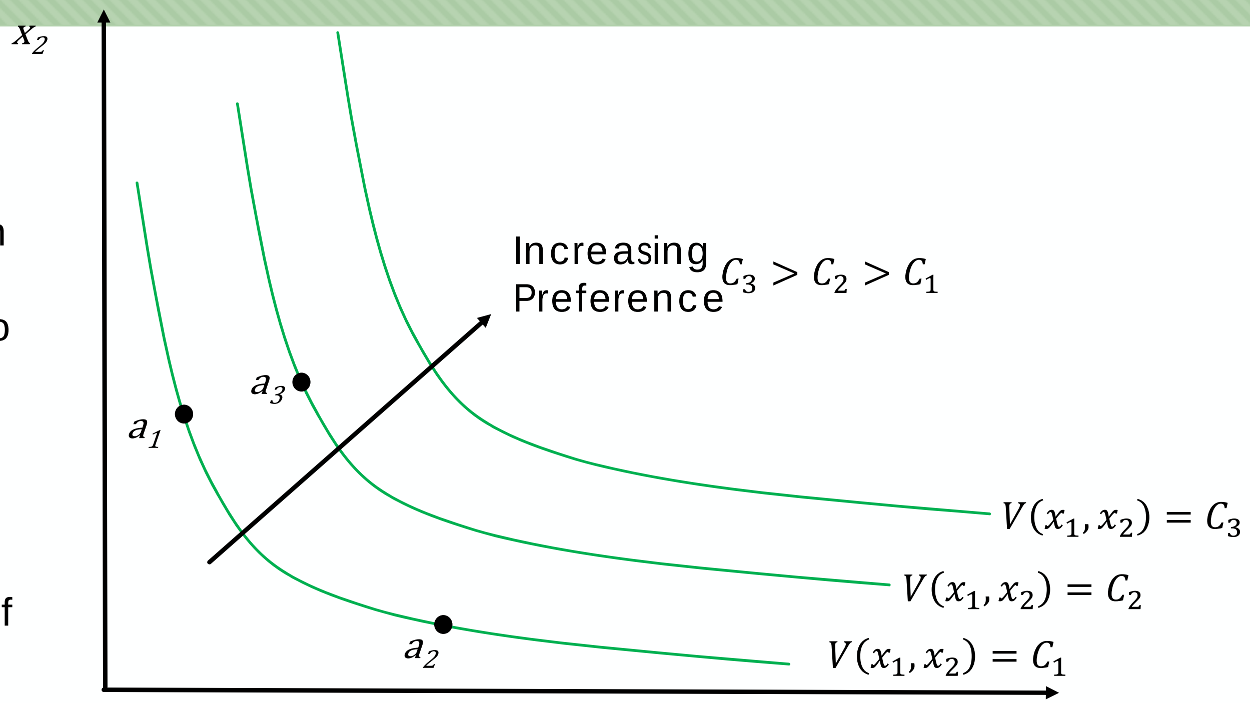
Value Functions

The idea behind value functions is to mathematically encode your preferences as a function. This is easiest to visualize in one dimension.



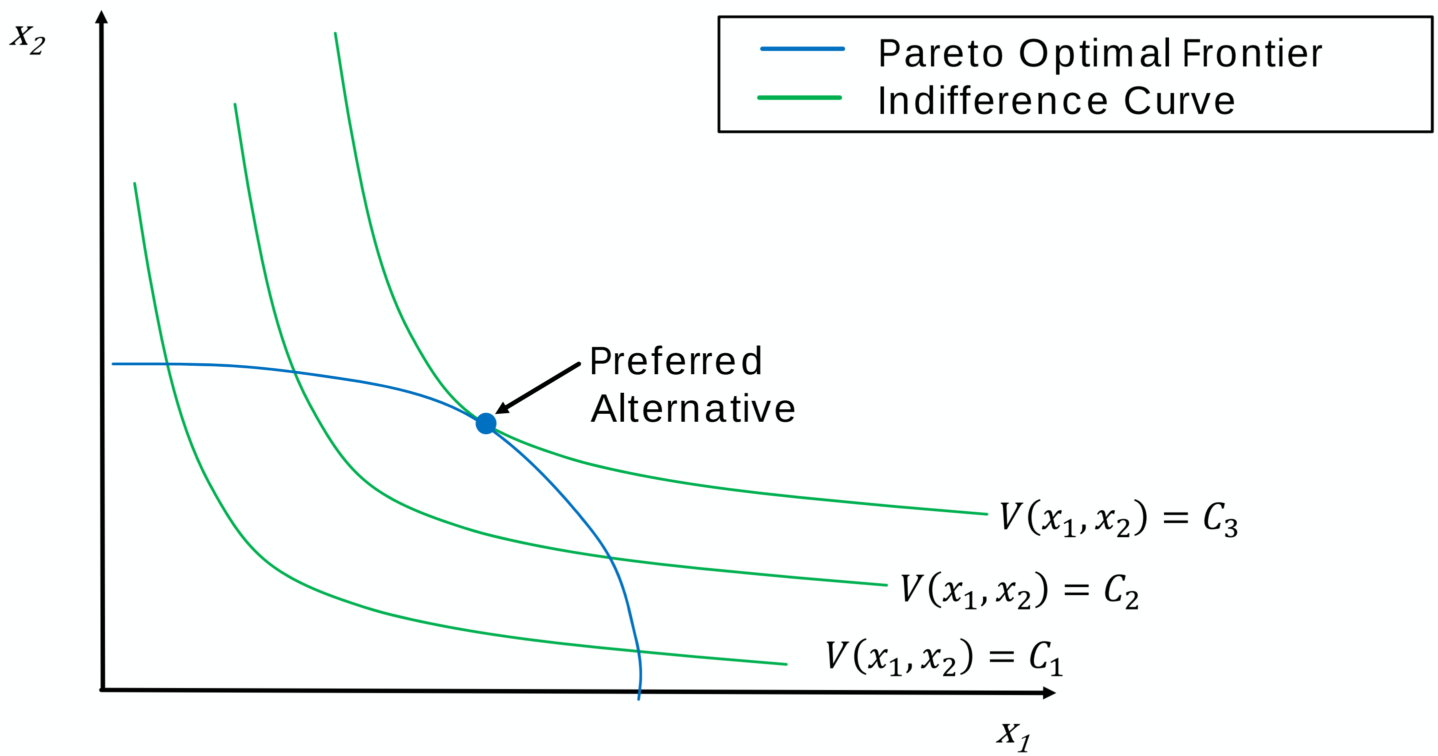
Indifference Curves

In economics, an **indifference curve** connects points on a graph representing different quantities of two goods, points between which a consumer is indifferent. That is, the consumer has no preference for one combination or bundle of goods over a different combination on the same **curve**.



a_3 is preferred to a_1
a_3 is preferred to a_2
Indifferent between a_1 and a_2

Relation to the Pareto Frontier



Reducing the Complexity of the tradeoffs

- Real world decisions often involve many objectives and many options
- Reconsider the car example: What if we included all of the objectives and all of the alternatives?
  - Price, fuel efficiency, handling, seating, trunk capacity, total ownership cost, appearance, resale value, reliability, accessories, etc.
  - Honda, Toyota, Ford, GM, Chrysler, BMW, Mercedes, Volkswagen, Fiat, Ferrari, Tesla, Aston Martin, Saab, Volvo, Mitsubishi, Audi, Lexus, Acura, Infinity, Nissan, etc.
  - …and that’s not even including the models and trim packages!
- We would like to avoid assessing value functions for all of these objectives and alternatives
  - Eliminate infeasible alternatives, eliminate dominated alternatives, use even swaps to eliminate objectives

Eliminate infeasible alternatives

- If an alternative does not meet your needs or violates a constraint, eliminate it
- For example, if you can’t afford a Ferrari, why would you waste effort assessing it?
- This seems obvious, but I have seen engineers leave infeasible alternatives in their trade studies and give them low scores
- Example: A trade study to select a new type of server assessed the value of an alternative that would not fit in the available rack space
  - We couldn’t change the available space so why consider it? Just document the rationale for eliminating the alternative and move on

Consequence Table

Example: Selecting New Office Space

	Parkway	Lombard	Baranov	Montana	Pierpoint
Commute in Minutes	45	25	20	25	30
Customer Access (%)	50	80	70	85	75
Office Services	A	B	C	A	C
Office Size (sq ft)	800	700	500	950	700
Monthly Cost	1850	1700	1500	1900	1750

Source: Hammond, J. S., R. L. Keeney, and H. Raiffa, Even Swaps: A Rational Method for Making Trade-offs, *Harvard Business Review*, March-April 1998, pp. 137-149.

Notation

- In the notation we are using, a consequence table is a matrix formed by combining the attribute vectors for each alternative
- Consequence $= [\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_i \quad \cdots \quad \boldsymbol{x}_n]$
- Consequence $= \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1i} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2i} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{ki} & \cdots & x_{kn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mi} & \cdots & x_{mn} \end{bmatrix}$
- Where there the set of alternatives are in indexed $i = 1$ to n and the attributes are indexed $k = 1$ to m

Before Lecture 3:

- Read Chapters 5, 7 and 11.
- Homework 2 is due next week.
- Make sure you have the Decision Tool software installed and I encourage you to read the tutorial and get yourselves familiar with it.

STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®