# Dynamic Programming

*Slides were prepared using the following book:*
"Introduction to Operations Research, 8th Ed." - Chapter 10
F. S. Hillier, G. J. Lieberman
McGraw Hill, 2005

Note: For the solutions to the examples used in the slides, please refer to the book.

# Dynamic Programming

- Mathematical technique
- Systematic procedure for determining optimal combination of decisions
- Requires experience to determine when and how to apply dynamic programming to a problem



# Prototypical Example – The Stagecoach Problem

- A problem specially constructed to illustrate the features and to introduce the terminology of dynamic programming problems

# Prototypical Example – The Stagecoach Problem

- A mythical fortune seeker in Missouri decided to go west to join the gold rush during the mid-19th century.
- The journey required travel through a country where there was serious danger of attack by marauders.
- His starting and end positions were fixed but he could choose through which states he passed on his route.
- Possible routes are shown in the figure on the next slide, where states are represented as nodes.
- Direction of travel is always from left to right.
- 4 coach runs are needed to go from A to J.



# Prototypical Example – The Stagecoach Problem

- The fortune seeker was concerned about his safety.
- Coach passengers were offered life insurance policies.
- The cost of the policy depends on the safety of the traveled route.
- So, the safest route would be the cheapest.
- The fortune seeker wants to determine a route so that the total cost of the insurance policy would be minimized.

# Prototypical Example – The Stagecoach Problem

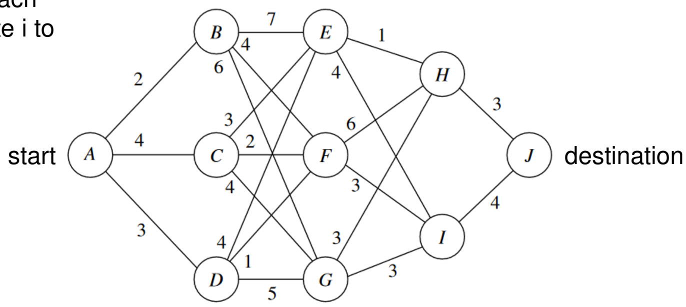
| | B | C | D |
|---|---|---|---|
| A | 2 | 4 | 3 |

| | E | F | G |
|---|---|---|---|
| B | 7 | 4 | 6 |
| C | 3 | 2 | 4 |
| D | 4 | 1 | 5 |

| | H | I |
|---|---|---|
| E | 1 | 4 |
| F | 6 | 3 |
| G | 3 | 3 |

| | J |
|---|---|
| H | 3 |
| I | 4 |

The cost for the standard policy on the stage coach run from state i to state j: $c_{ij}$






# Solving the Problem

- Choosing cheapest policy at each stage gives

  A -> B -> F -> I -> J with cost 13
- However, A -> D -> F is cheaper overall than A -> B -> F
- Dynamic programming
  - starts with a smaller part of the original problem and finds the optimal solution for this smaller problem
  - then gradually enlarges the problem, finding the current optimal solution from the preceding one, until the original problem is solved in its entirety

# Solving the Problem: Stagecoach

- Start with smaller problem: Fortune seeker (FS) is at last stage
- At each iteration, the problem is enlarged by increasing by 1 the number of stages left to complete journey
- For the enlarged problem, the optimal solution for where to go next from each possible state is found using results obtained in preceding iteration.



# Formulation

- Stages $x_n$ (n=1,2,3,4)
- $f_n(s,x_n)$: total cost of best overall policy for the remaining stages, given that the FS is in stage s and selects $x_n$ as the immediate destination.
- $x_n^*$ is any value of $x_n$ (not necessarily unique) that minimizes $f_n(s,x_n)$, $f_n^*(s)$ is the corresponding minimum value of $f_n(s,x_n)$, i.e. $f_n(s,x_n^*)$
- $f_n(s,x_n)$ = immediate cost at stage n + minimum future cost for stages n+1 onward, i.e.
  - $f_n(s,x_n) = c_{sxn} + f_{n+1}^*(x_n)$
- Since the ultimate destination (state J) is reached at the end of stage 4, $f_5^*(J)= 0$.
- Dynamic Programming finds the route successively by finding $f_4^*(s)$, $f_3^*(s)$, $f_2^*(s)$ for each possible state s, then uses $f_2^*(s)$ to solve for $f_1^*(A)$.

## Characteristics of Dynamic Programming Problems

- One way to recognize that a situation can be formulated as a dynamic programming problem is to notice that its basic structure is analogous to the stage coach problem.



## Characteristics of Dynamic Programming Problems

- The problem can be divided into stages, with a policy decision required at each stage
  - Requires making a *sequence of interrelated decisions*
- Each stage has a number of states associated with the beginning of that stage, i.e., various possible conditions the system may be in at that stage of the problem.
- The effect of the policy decision at each stage is to transform the current state to a state associated with the beginning of the next stage.
- The solution procedure is designed to find an optimal policy for the overall problem, i.e., an optimal policy decision at each stage for each of the possible states.

# Characteristics of Dynamic Programming Problems

Given the current state, an *optimal policy for the remaining stages* is *independent* of the policy decisions made in *previous stages*. The optimal immediate decision depends only on the current state and not on how you got there. This is called the "principle of optimality" for dynamic programming.

- The solution procedure begins by finding an *optimal policy for the last stage*.
- A recursive relationship, that identifies the optimal policy for stage n, given the optimal policy for stage n+1, is available.
  - e.g., for the stagecoach problem, $f_n^*(s)=\min_{x_n} (c_{sx_n} + f_{n+1}^*(x_n))$
- When we use this recursive procedure, the solution procedure starts at the end and moves *backward* stage by stage, each time finding the optimal policy for that stage; until it finds the optimal policy starting at the *initial* stage.

13

# Characteristics of Dynamic Programming Problems

| $s_n$ \ $x_n$ | $f_n(s_n, x_n)$ | $f_n^*(s_n)$ | $x_n^*$ |
|---|---|---|---|
| | | | |

Obtain such a table for each stage (n = N, N-1, …, 1).
When this table is finally obtained for the initial stage (n=1), the problem is solved.

14

# Example 2: Distributing Medical Teams to Countries

- World Health Council has five available medical teams to allocate among three countries.
- The council needs to decide how many teams (if any) to allocate to each of these three countries to maximize the total effectiveness of the five teams.
- Teams must be allocated as a whole, i.e., the number allocated to each country must be an integer; no partial team assignments are allowed.
- Performance measure is:

  (increased life expectancy in years * country's population)

- Which allocation maximizes this measure?

15

# Example 2: Distributing Medical Teams to Countries

| Teams | Country1 | Country2 | Country3 |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 45 | 20 | 50 |
| 2 | 70 | 45 | 70 |
| 3 | 90 | 75 | 80 |
| 4 | 105 | 110 | 100 |
| 5 | 120 | 150 | 130 |

In terms of thousands of additional person-years of life

16

# Questions to Ask to Determine the States

- What changes from one stage to next?
- Given that the decisions have been made at the previous stages, how can the status of the situation at the current stage be described?
- What information about the current state is necessary to determine the optimal policy from then on?



# Formulation

- State of the system $x_n$ is the number of available medical teams
- At stage 1 (country 1), where all three countries remain under consideration for allocations, $s_1$=5.
- However, at stage 2 or 3 (country 2 or 3), $s_n$ is just 5 minus the number of teams allocated at preceding stages, so that the sequence of states is

$$s_1=5 , \qquad s_2 = 5 - x_1 , \qquad s_3 = s_2 - x_2$$

- When we are solving at stage 2 or 3, we shall not yet have solved for the allocations at the preceding stages.
- So, we consider every possible state we could be in at stage 2 or 3, namely, $s_n$=0, 1, 2, 3, 4, or 5.