Qui hoạch động

Dynamic programming

Nội dung

Tìm hiểu về quy hoạch động (dynamic programming) qua các ví dụ tiêu biểu:

- 0. Fibonacci Numbers
- 1. Longest common subsequence
- 2. Knapsack problem

Sử dụng một phần tài liệu bài giảng CS161 Stanford University

Khởi động: Fibonacci Numbers

• Definition:

```
• F(n) = F(n-1) + F(n-2), with F(1) = F(2) = 1.
```

- The first several are:
 - 1
 - 1
 - 2
 - 3
 - 5
 - 8
 - 13, 21, 34, 55, 89, 144,...

Question:

Given n, what is F(n)?

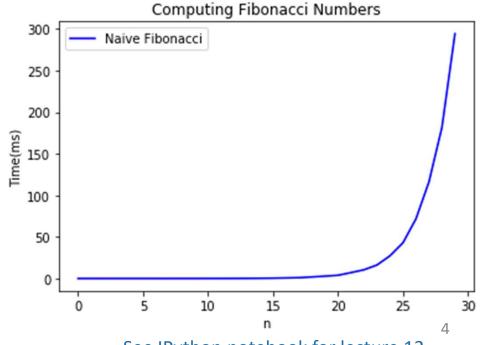
Candidate algorithm

```
    def Fibonacci(n):

            if n == 0, return 0
            if n == 1, return 1
            return Fibonacci(n-1) + Fibonacci(n-2)
```

Running time?

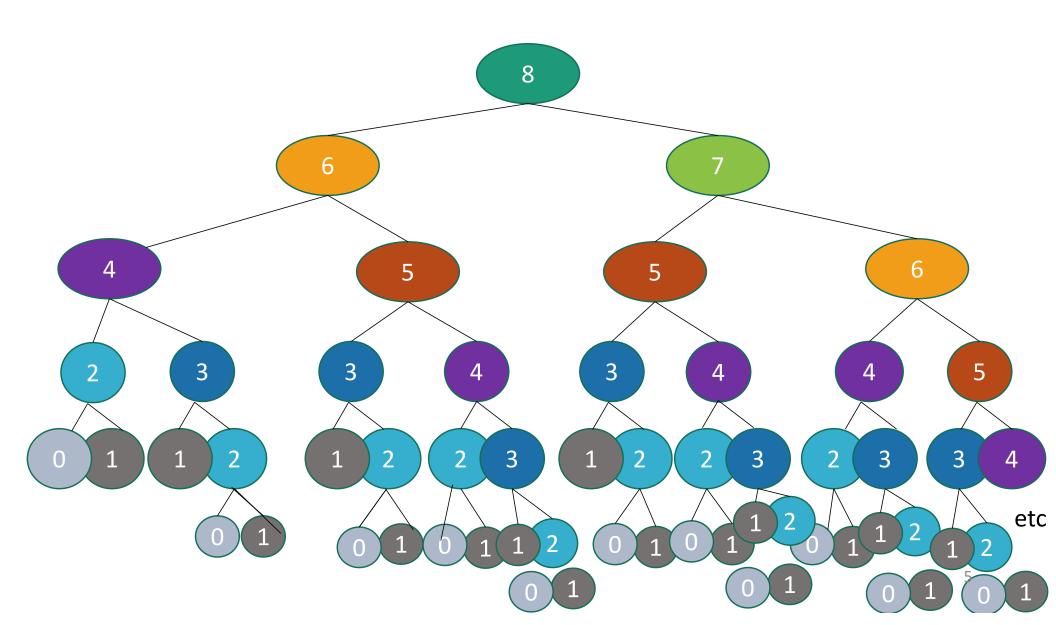
- T(n) = T(n-1) + T(n-2) + O(1)
- $T(n) \ge T(n-1) + T(n-2)$ for $n \ge 2$
- So T(n) grows at least as fast as the Fibonacci numbers themselves...
- You showed in HW1 that this is EXPONENTIALLY QUICKLY!



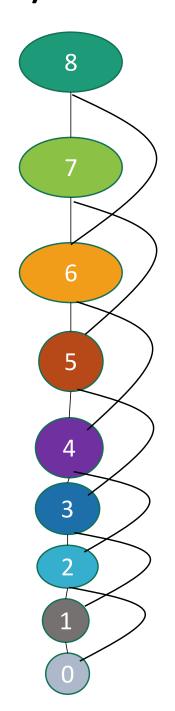
See IPython notebook for lecture 12

What's going on? Consider Fib(8)

That's a lot of repeated computation!



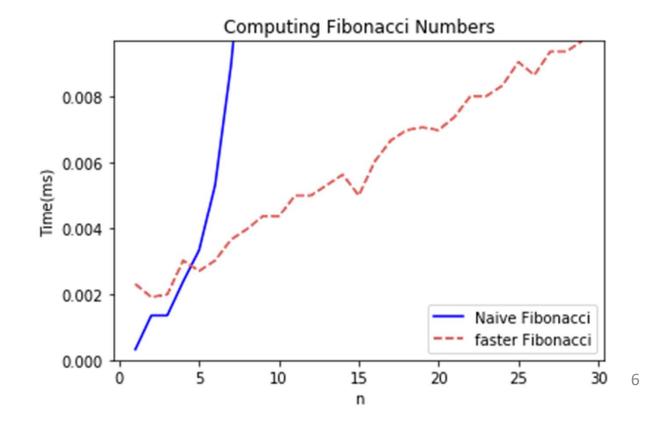
Maybe this would be better:



def fasterFibonacci(n):

- F = [0, 1, None, None, ..., None]
 \\ F has length n + 1
- for i = 2, ..., n:
 - F[i] = F[i-1] + F[i-2]
- return F[n]

Much better running time!



This was an example of...



What is *dynamic programming*?

- Là một chiến lược thiết kế thuật toán (an algorithm design paradigm)
 - Các bạn đã học chiến lược chia để trị (divide-andconquer).
- Thường dùng để giải các bài toán tối ưu
 - Ví dụ, các dạng tìm đường đi ngắn nhất
 - Tính số Fibonacci không phải là bài toán tối ưu, nhưng là một ví dụ đơn giản để thuyết minh DP

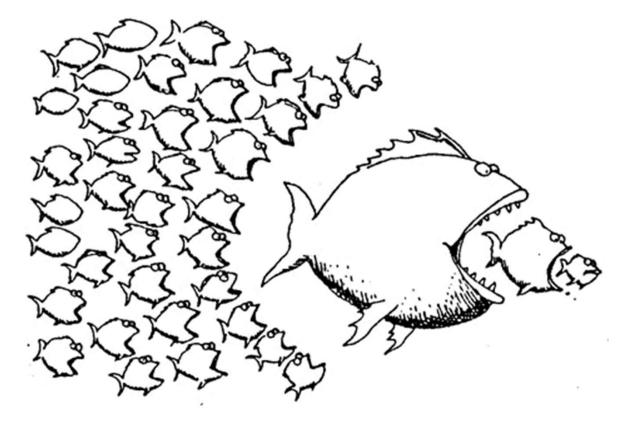
Elements of dynamic programming

- Optimal substructure.
 - Optimal solutions to sub-problems can be used to find the optimal solution of the original problem.
- Overlapping subproblems.
 - The subproblems show up again and again
- Using these properties, we can design a dynamic programming algorithm:
 - Keep a table of solutions to the smaller problems.
 - Use the solutions in the table to solve bigger problems.
 - At the end we can use information we collected along the way to find the solution to the whole thing.

Two ways to think about and/or implement DP algorithms

Top down

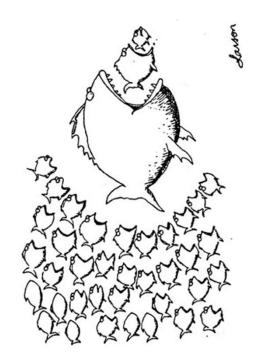
Bottom up





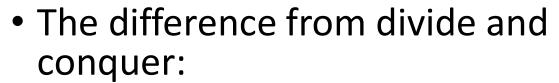
Bottom up approach what we just saw.

- For Fibonacci:
- Solve the small problems first
 - fill in F[0],F[1]
- Then bigger problems
 - fill in F[2]
- •
- Then bigger problems
 - fill in F[n-1]
- Then finally solve the real problem.
 - fill in F[n]



Top down approach

- Think of it like a recursive algorithm.
- To solve the big problem:
 - Recurse to solve smaller problems
 - Those recurse to solve smaller problems
 - etc..



- Keep track of what small problems you've already solved to prevent re-solving the same problem twice.
- Aka, "memo-ization"





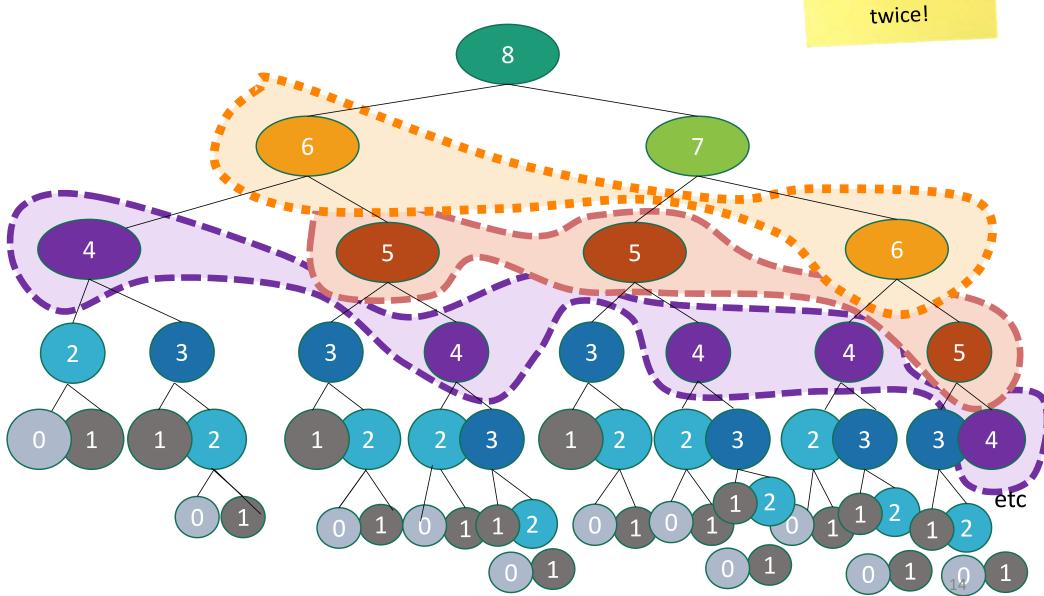
Example of top-down Fibonacci

```
• define a global list F = [0,1,None, None, ..., None]
• def Fibonacci(n):
    • if F[n] != None:
         • return F[n]
    • else:
         • F[n] = Fibonacci(n-1) + Fibonacci(n-2)
    • return F[n]
                                              Computing Fibonacci Numbers
                                0.008
                           Time(ms)
9000
    Memo-ization:
  Keeps track (in F) of
   the stuff you've
     already done.
                                0.002
                                                             Naive Fibonacci
                                                             faster Fibonacci, bottom-up
                                                            faster Fibonacci, top-down
                                0.000
                                                  10
                                                                20
```

15

Memo-ization visualization

Collapse
repeated nodes
and don't do the
same work
twice!



C++ Implementation

```
int fibo(int f[], int n) {
    if (f[n] >=0) return f[n];
    return (f[n] = fibo(f, n-1) + f[n-2]);
}
```

```
int main() {
    int n;
    cin >> n;
    int a[n+1];
    a[0] = 0, a[1] = 1;
    for (int i=2; i<=n; i++) a[i] = -1;
    cout << fibo(a, n) << endl;
}</pre>
```

Nội dung

Tìm hiểu về quy hoạch động (dynamic programming) qua các ví dụ tiêu biểu:

- 0. Fibonacci Numbers
- 1. Longest common subsequence
- 2. Knapsack problem

Longest Common Subsequence

How similar are these two species?



AGCCCTAAGGGCTACCTAGCTT

DNA:



DNA:
GACAGCCTACAAGCGTTAGCTTG

Longest Common Subsequence

How similar are these two species?



DNA:

DNA:
GACAGCCTACAAGCGTTAGCTTG

Pretty similar, their DNA has a long common subsequence:

AGCCTAAGCTTAGCTT

Longest Common Subsequence

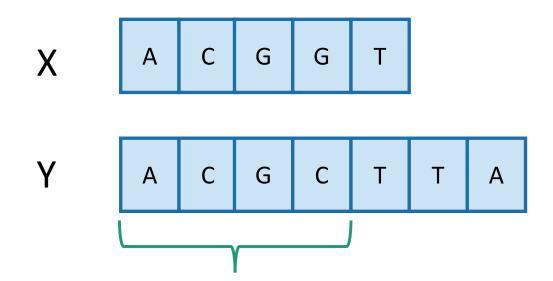
- Subsequence:
 - BDFH is a subsequence of ABCDEFGH
- If X and Y are sequences, a **common subsequence** is a sequence which is a subsequence of both.
 - BDFH is a common subsequence of ABCDEFGH and of ABDFGHI
- A longest common subsequence...
 - ...is a common subsequence that is longest.
 - The longest common subsequence of ABCDEFGH and ABDFGHI is ABDFGH.

Recipe for applying Dynamic Programming

- Step 1: Identify optimal substructure.
- Step 2: Find a recursive formulation for the length of the longest common subsequence.
- Step 3: Use dynamic programming to find the length of the longest common subsequence.
- Step 4: If needed, keep track of some additional info so that the algorithm from Step 3 can find the actual LCS.
- Step 5: If needed, code this up like a reasonable person.

Step 1: Optimal substructure

Prefixes:



Notation: denote this prefix **ACGC** by Y₄

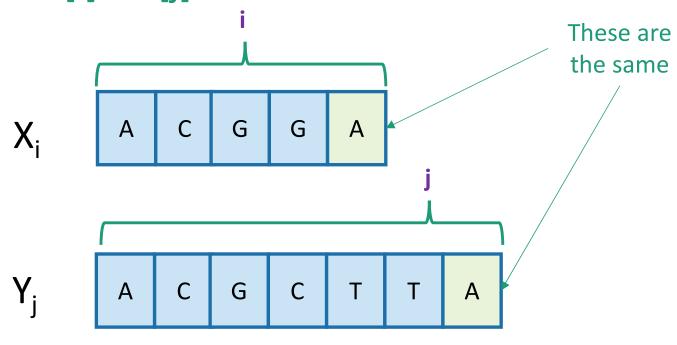
- Our sub-problems will be finding LCS's of prefixes to X and Y.
- Let C[i,j] = length_of_LCS(X_i, Y_i)

Examples: C[2,3] = 2C[4,4] = 3

Two cases

Case 1: X[i] = Y[j]

- Our sub-problems will be finding LCS's of prefixes to X and Y.
- Let C[i,j] = length_of_LCS(X_i, Y_i)

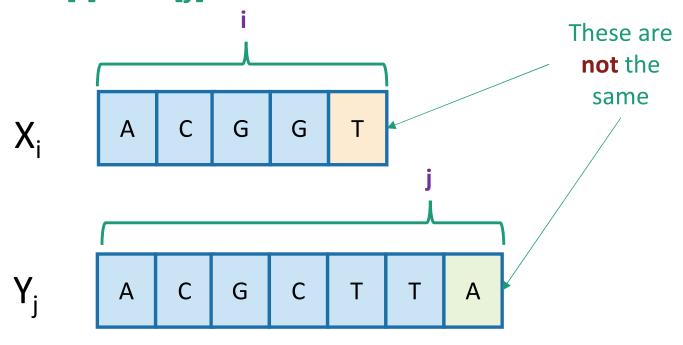


- Then C[i,j] = 1 + C[i-1,j-1].
 - because $LCS(X_i,Y_j) = LCS(X_{i-1},Y_{j-1})$ followed by

Two cases

Case 2: X[i] != Y[j]

- Our sub-problems will be finding LCS's of prefixes to X and Y.
- Let C[i,j] = length_of_LCS(X_i, Y_i)



- Then C[i,j] = max{ C[i-1,j], C[i,j-1] }.
 - either $LCS(X_i, Y_i) = LCS(X_{i-1}, Y_i)$ and \top is not involved,
 - or $LCS(X_i,Y_i) = LCS(X_i,Y_{i-1})$ and A is not involved,
 - (maybe both are not involved, that's covered by the "or"),

Recursive formulation of the optimal solution

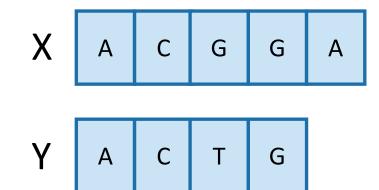
• $C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1],C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$

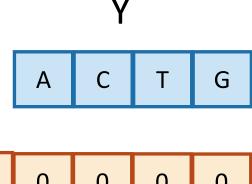
LCS DP

- LCS(X, Y):
 - C[i,0] = C[0,j] = 0 for all i = 0,...,m, j=0,...n.
 - **For** i = 1,...,m and j = 1,...,n:
 - **If** X[i] = Y[j]:
 - C[i,j] = C[i-1,j-1] + 1
 - Else:
 - C[i,j] = max{ C[i,j-1], C[i-1,j] }
 - Return C[m,n]

Running time: O(nm)

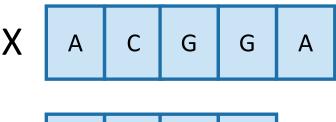
$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1],C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$





0	0	0	0	0
0				
0				
0				
0				
0				

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$





Υ					
Α	С	Т	G		

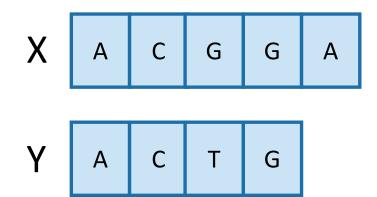
0	0	0	0	0
0	1	1	1	1
0	1	2	2	2
0	1	2	2	3
0	1	2	2	3
0	1	2	2	3

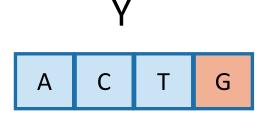
So the LCM of X and Y has length 3.

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$

Recipe for applying Dynamic Programming

- Step 1: Identify optimal substructure.
- Step 2: Find a recursive formulation for the length of the longest common subsequence.
- Step 3: Use dynamic programming to find the length of the longest common subsequence.
- Step 4: If needed, keep track of some additional info so that the algorithm from Step 3 can find the actual LCS.
- Step 5: If needed, code this up like a reasonable person.

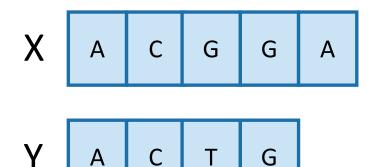


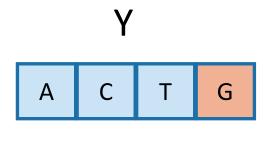


 Once we've filled this in, we can work backwards.

0	0	0	0	0
0	1	1	1	1
0	1	2	2	2
0	1	2	2	3
0	1	2	2	3
0	1	2	2	3

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$





• Once we've filled this in, we can work backwards.

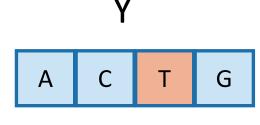
0	0	0	0	0
0	1	1	1	1
0	1	2	2	2
0	1	2	2	3
0	1	2	2	3
0	1	2	2	3

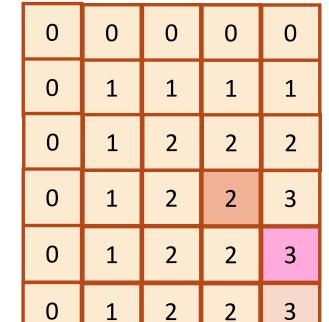
That 3 must have come from the 3 above it.

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$



Y A C T G





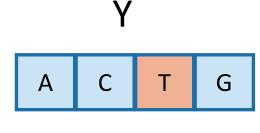
- Once we've filled this in, we can work backwards.
- A diagonal jump means that we found an element of the LCS!

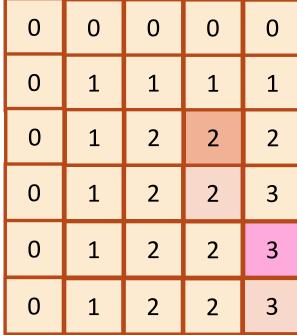
This 3 came from that 2 – we found a match!

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$



Y A C T G

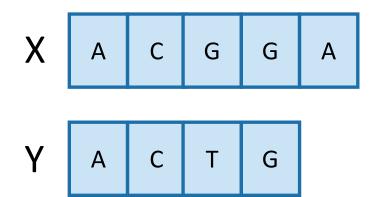


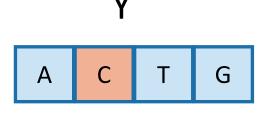


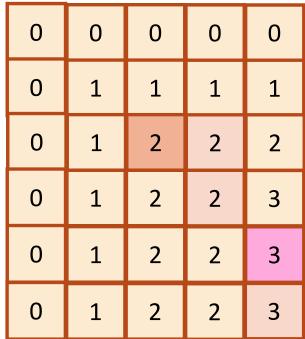
- Once we've filled this in, we can work backwards.
- A diagonal jump means that we found an element of the LCS!

That 2 may as well have come from this other 2.

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$



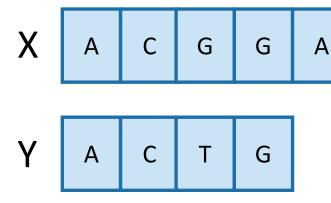


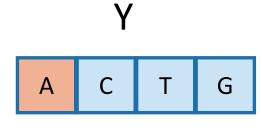


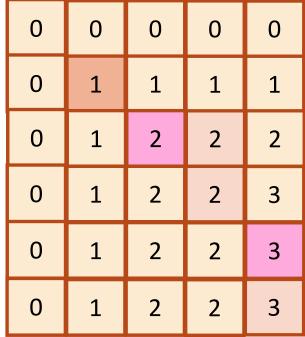
- Once we've filled this in, we can work backwards.
- A diagonal jump means that we found an element of the LCS!

G

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$



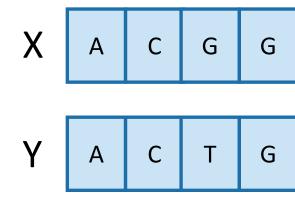




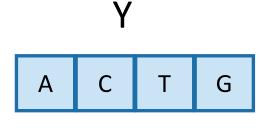
- Once we've filled this in, we can work backwards.
- A diagonal jump means that we found an element of the LCS!

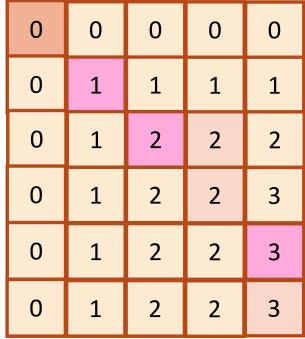
C G

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$

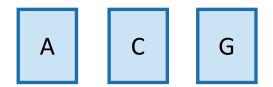


Α





- Once we've filled this in, we can work backwards.
- A diagonal jump means that we found an element of the LCS!



This is the LCS!

$$C[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ C[i-1,j-1] + 1 & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max\{C[i,j-1], C[i-1,j]\} & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$

Nội dung

Tìm hiểu về quy hoạch động (dynamic programming) qua các ví dụ tiêu biểu:

- 0. Fibonacci Numbers
- 1. Longest common subsequence
- 2. Knapsack problem

Example 2: Knapsack Problem

We have n items with weights and values:

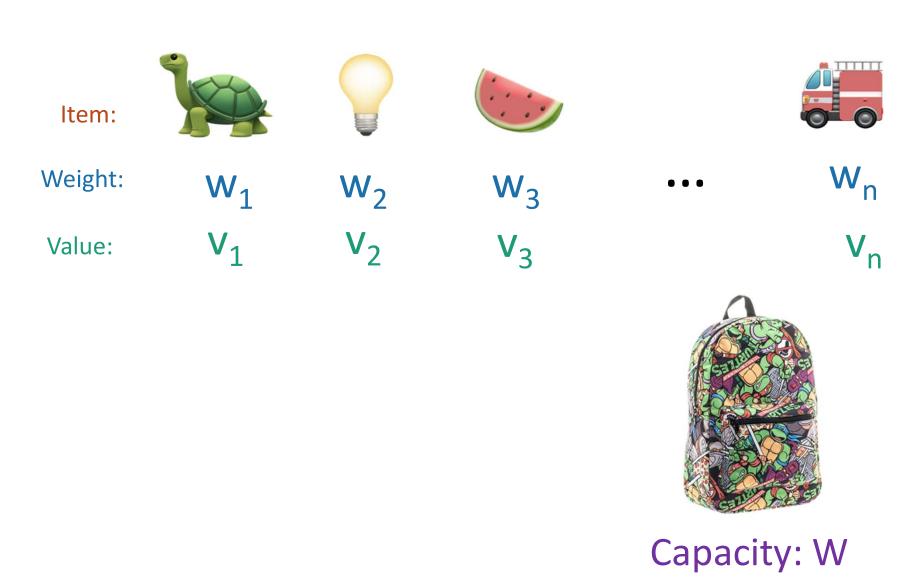


- And we have a knapsack:
 - it can only carry so much weight:



Capacity: 10

Some notation





Capacity: 10













Weight:

Item:

6

2

4

3

13

11

35

Value:

20

8

14

Unbounded Knapsack:

- Suppose I have infinite copies of all of the items.
- What's the most valuable way to fill the knapsack?









Total weight: 10 Total value: 42



• 0/1 Knapsack:

- Suppose I have only one copy of each item.
- What's the most valuable way to fill the knapsack?







Total weight: 9 Total value: 35

Recipe for applying Dynamic Programming

• Step 1: Identify optimal substructure.



- Step 2: Find a recursive formulation for the value of the optimal solution.
- Step 3: Use dynamic programming to find the value of the optimal solution.
- Step 4: If needed, keep track of some additional info so that the algorithm from Step 3 can find the actual solution.
- Step 5: If needed, code this up like a reasonable person.

Optimal substructure: try 2

Sub-problems:

• 0/1 Knapsack with fewer items.

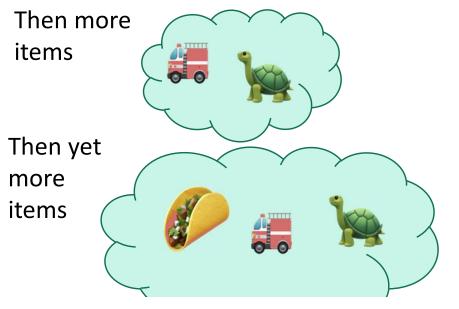
First solve the problem with few items







We'll still increase the size of the knapsacks.



(We'll keep a two-dimensional table).

Our sub-problems:

Indexed by x and j

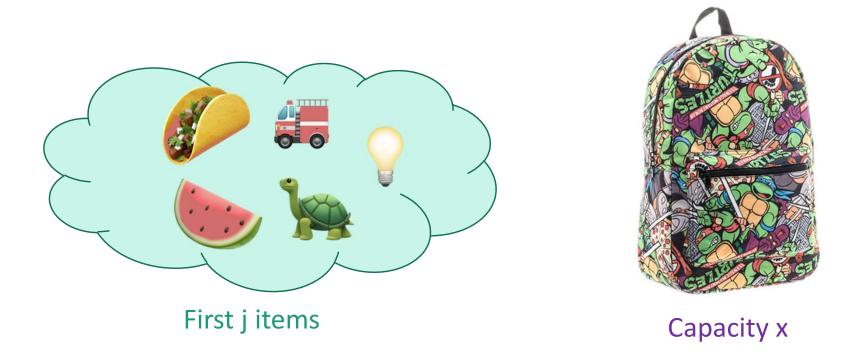


Capacity x

K[x,j] = optimal solution for a knapsack of size x using only the first j items.

Relationship between sub-problems

Want to write K[x,j] in terms of smaller sub-problems.



K[x,j] = optimal solution for a knapsack of size x using only the first j items.



- Case 1: Optimal solution for j items does not use item j.
- Case 2: Optimal solution for j items does use item j.



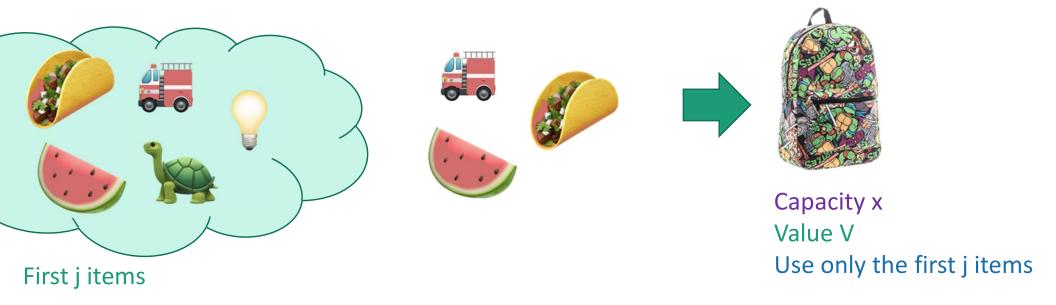


Capacity x

K[x,j] = optimal solution for a knapsack of size x using only the first j items.



Case 1: Optimal solution for j items does not use item j.



What lower-indexed problem should we solve to solve this problem?

1 min think; (wait) 1 min share

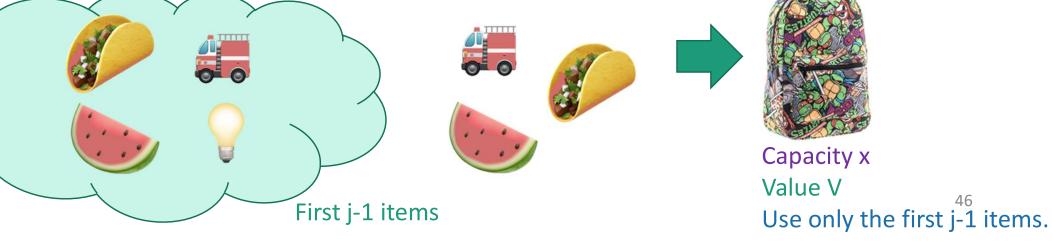




• Case 1: Optimal solution for j items does not use item j.



Then this is an optimal solution for j-1 items:





item j

• Case 2: Optimal solution for j items uses item j.



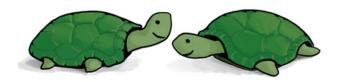




Capacity x
Value V
Use only the first j items

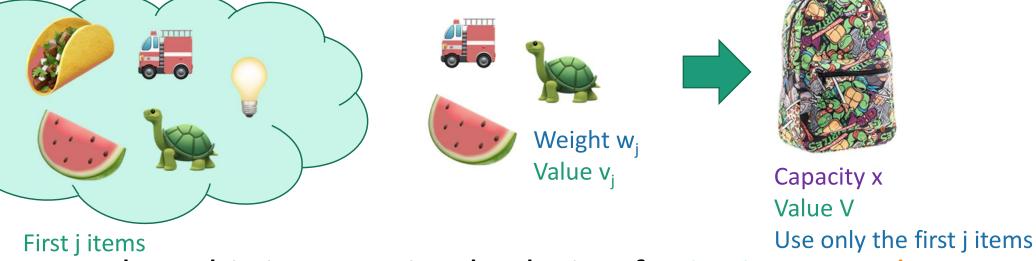
What lower-indexed problem should we solve to solve this problem?

1 min think; (wait) 1 min share





Case 2: Optimal solution for j items uses item j.



Then this is an optimal solution for j-1 items and a



Recursive relationship

- Let K[x,j] be the optimal value for:
 - capacity x,
 - with j items.

$$K[x,j] = max\{ K[x, j-1], K[x - w_{j,} j-1] + v_{j} \}$$
Case 1

Case 2

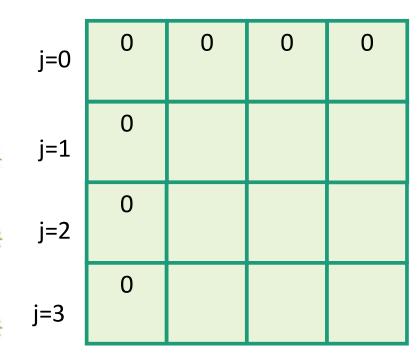
• (And K[x,0] = 0 and K[0,j] = 0).

Bottom-up DP algorithm

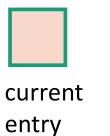
return K[W,n]

```
Zero-One-Knapsack(W, n, w, v):
   • K[x,0] = 0 for all x = 0,...,W
   • K[0,i] = 0 for all i = 0,...,n
   • for x = 1,...,W:
       • for j = 1,...,n:
                                Case 1
           • K[x,i] = K[x, i-1]
           • if w_i \leq x:
                                                 Case 2
               • K[x,j] = max\{ K[x,j], K[x-w_i, j-1] + v_i \}
```

$$x=0$$
 $x=1$ $x=2$ $x=3$



- Zero-One-Knapsack(W, n, w, v):
 - K[x,0] = 0 for all x = 0,...,W
 - K[0,i] = 0 for all i = 0,...,n
 - for x = 1,...,W:
 - for j = 1,...,n:
 - K[x,j] = K[x, j-1]
 - if $w_i \le x$:
 - $K[x,j] = max\{ K[x,j],$ $K[x - w_{i}, j-1] + v_{i}$
 - return K[W,n]





relevant previous entry



Weight:













6



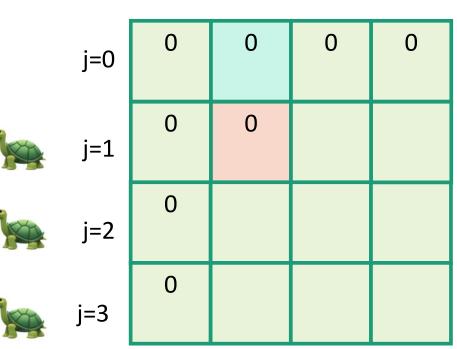
Capacity: 3

Value:

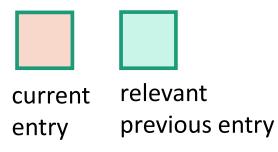
1

1

x=0	x=1	x=2	x=3



- Zero-One-Knapsack(W, n, w, v):
 - K[x,0] = 0 for all x = 0,...,W
 - K[0,i] = 0 for all i = 0,...,n
 - for x = 1,...,W:
 - **for** j = 1,...,n:
 - K[x,j] = K[x, j-1]
 - if $w_i \le x$:
 - $K[x,j] = max\{ K[x,j],$ $K[x - w_{i}, j-1] + v_{i}$
 - return K[W,n]





Item:

Weight:











1 Value:

1

4

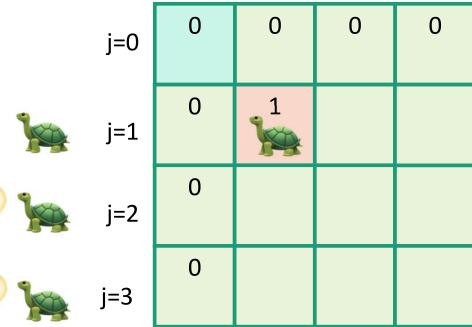
2

6

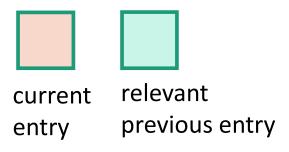
3

Capacity: 3





- Zero-One-Knapsack(W, n, w, v):
 - K[x,0] = 0 for all x = 0,...,W
 - K[0,i] = 0 for all i = 0,...,n
 - for x = 1,...,W:
 - for j = 1,...,n:
 - K[x,j] = K[x, j-1]
 - if $w_i \le x$:
 - $K[x,j] = max\{ K[x,j],$ $K[x - w_{i}, j-1] + v_{i}$
 - return K[W,n]







1









1 Value:

Item:

Weight:

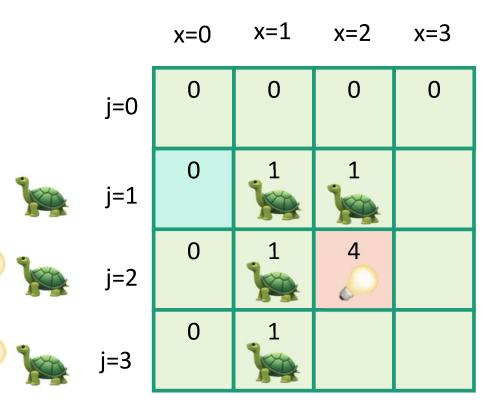
4

2

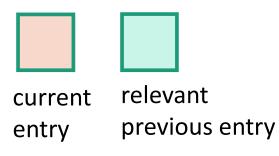
6

3

Capacity: 3



- Zero-One-Knapsack(W, n, w, v):
 - K[x,0] = 0 for all x = 0,...,W
 - K[0,i] = 0 for all i = 0,...,n
 - for x = 1,...,W:
 - **for** j = 1,...,n:
 - K[x,j] = K[x, j-1]
 - if $w_i \le x$:
 - $K[x,j] = max\{K[x,j],$ $K[x - w_{i}, j-1] + v_{i}$
 - return K[W,n]















Value:

Weight:

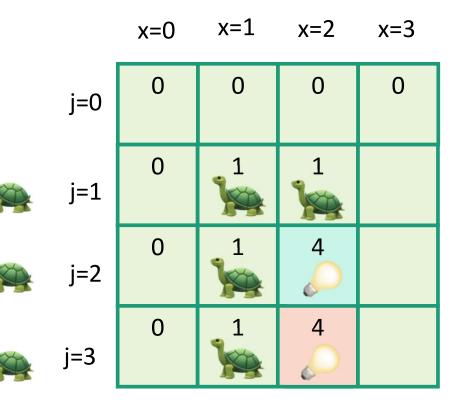
1

4

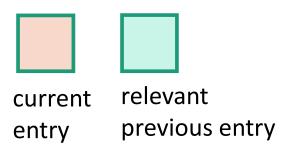
6

3

Capacity: 3



- Zero-One-Knapsack(W, n, w, v):
 - K[x,0] = 0 for all x = 0,...,W
 - K[0,i] = 0 for all i = 0,...,n
 - for x = 1,...,W:
 - **for** j = 1,...,n:
 - K[x,j] = K[x, j-1]
 - if $w_i \le x$:
 - $K[x,j] = max\{K[x,j],$ $K[x - w_{i}, j-1] + v_{i}$
 - return K[W,n]







Item:

Weight:



2



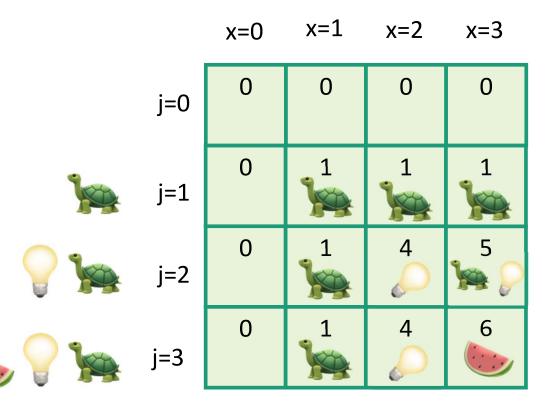




4 1 Value:

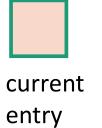
6

Capacity: 3



- Zero-One-Knapsack(W, n, w, v):
 - K[x,0] = 0 for all x = 0,...,W
 - K[0,i] = 0 for all i = 0,...,n
 - for x = 1,...,W:
 - **for** j = 1,...,n:
 - K[x,j] = K[x, j-1]
 - if $w_i \le x$:
 - $K[x,j] = max\{K[x,j],$ $K[x-w_j, j-1] + v_j\}$
 - return K[W,n]

So the optimal solution is to put one watermelon in your knapsack!





relevant previous entry



Weight: Value:



-

1













6



Capacity: 3

What have we learned?

Dynamic programming:

- Paradigm in algorithm design.
- Uses optimal substructure
- Uses overlapping subproblems
- Can be implemented **bottom-up** or **top-down**.
- It's a fancy name for a pretty common-sense idea:

Don't duplicate work if you don't have to!

Why "dynamic programming"?

- Programming refers to finding the optimal "program."
 - as in, a shortest route is a *plan* aka a *program*.
- Dynamic refers to the fact that it's multi-stage.
- But also it's just a fancy-sounding name.



Why "dynamic programming"?

- Richard Bellman invented the name in the 1950's.
- At the time, he was working for the RAND Corporation, which was basically working for the Air Force, and government projects needed flashy names to get funded.
- From Bellman's autobiography:
 - "It's impossible to use the word, dynamic, in the pejorative sense...I thought dynamic programming was a good name. It was something not even a Congressman could object to."