CSE 2320 Notes 7: Dynamic Programming

(CLRS 15.1-15.4)

DYNAMIC PROGRAMMING APPROACH

1. Describe problem input.
2. Determine cost function and base case.
3. Determine general case for cost function. THE HARD PART!!!
4. Appropriate ordering for enumerating subproblems.
   a. Simple bottom-up approach - from small problems towards the entire big problem.
   b. Top-down approach with “memoization” - to attack large problems.
5. Backtrace for solution. *Most of the effort in dynamic programming is ignored at the end.*
   a. Predecessor/back pointers to get to the subproblems whose results are in the solution.
   b. Top-down recomputation of cost function (to reach the same subproblems as 5.a)
      (Providing all solutions is an extra cost feature . . .)

7.A. A SMALL EXAMPLE – Shuttle-to-Airport

![Diagram of Hotel-Airport path with numbers and arrows]
How many different paths (by brute force)?

Observation: To find optimal route, need optimal route to each street corner.

(Could also use Dijkstra’s algorithm, Notes 16, which is more general, but slower.)

1. Describe problem input.

Four arrays of paths, each with $n$ values

- Upper Direct = UD = $ud_1, ud_2, \ldots, ud_n = 9 (2 + 7), 9, 3, 4, 8, 7 (4 + 3)$
- Lower Direct = LD = $ld_1, ld_2, \ldots, ld_n = 12 (4 + 8), 5, 6, 4, 5, 9 (7 + 2)$
- Upper-to-Lower = UL = $ul_1, ul_2, \ldots, ul_n = 2, 3, 1, 3, 4, \infty$
- Lower-to-Upper = LU = $lu_1, lu_2, \ldots, lu_n = 2, 1, 2, 1, \infty$

2. Determine cost function and base case.

$U(i) =$ Cost to reach upper corner $i$
$L(i) =$ Cost to reach lower corner $i$
$U(0) = 0$
$L(0) = 0$

3. Determine general case.

$U(i) = \min \{ U(i - 1) + ud_i, L(i - 1) + ld_i + lu_i \}$
$L(i) = \min \{ L(i - 1) + ld_i, U(i - 1) + ud_i + ul_i \}$

4. Appropriate ordering of subproblems.

$U(i)$ and $L(i)$ cannot be computed without $U(i - 1)$ and $L(i - 1)$

5. Backtrace for solution – either

a. (http://ranger.uta.edu/~weems/NOTES2320/shuttle1.c) explicitly save indication of which of the two cases was used (continue - c, switch - s), or

b. (http://ranger.uta.edu/~weems/NOTES2320/shuttle2.c) recheck during backtrace for which case was used.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>0</td>
<td>9 (c)</td>
<td>17 (s)</td>
<td>20 (c)</td>
<td>24 (c)</td>
<td>31 (s)</td>
<td>38 (c)</td>
</tr>
<tr>
<td>$L$</td>
<td>0</td>
<td>11 (s)</td>
<td>16 (c)</td>
<td>21 (s)</td>
<td>25 (c)</td>
<td>30 (c)</td>
<td>39 (c)</td>
</tr>
</tbody>
</table>
Dynamic programming is:

1. Exhaustive search without brute force.
2. Optimal solution to big problem from optimal solutions to subproblems.

7.B. Weighted Interval Scheduling

Input: A set of $n$ intervals numbered 1 through $n$ with each interval $i$ having start time $s_i$, finish time $f_i$, and positive weight $v_i$.

Output: A set of non-overlapping intervals to maximize the sum of their weights. (Two intervals $i$ and $j$ overlap if either $s_i < s_j < f_i$ or $s_i < f_j < f_i$.)

Brute-force solution: Enumerate the powerset of the input intervals, discard those cases with overlapping intervals, and compute the sum of the weights for each. ( http://ranger.uta.edu/~weems/NOTES2320/wis.power.c )

1. Describe problem input.

Assume the $n$ intervals are in ascending finish time order, i.e. $f_i \leq f_{i+1}$.

Let $p_i$ be the rightmost preceding interval for interval $i$, i.e. the largest value $j < i$ such that intervals $i$ and $j$ do not overlap. If no such interval $j$ exists, $p_i = 0$. (These values may be computed in $\Theta(n \log n)$ time using binSearchLast() from Notes 1. See http://ranger.uta.edu/~weems/NOTES2320/wis.bs.c )
2. Determine cost function and base case.

\[ M(i) = \text{Cost for optimal non-overlapping subset for the first } i \text{ input intervals.} \]

\[ M(0) = 0 \]

3. Determine general case.

For \( M(i) \), the main issue is: Does the optimal subset include interval \( i \)?

If \( \text{yes} \): optimal subset cannot include any overlapping intervals, so \( M(i) = M(p_i) + v_i \).

If \( \text{no} \): optimal subset is the same as for \( M(i-1) \), so \( M(i) = M(i-1) \).

This observation tells us to compute cost both ways and keep the maximum.

4. Appropriate ordering of subproblems. Simply compute \( M(i) \) in ascending \( i \) order.

5. Backtrace for solution (with recomputation). This is the subset of intervals for \( M(n) \).

\[
\begin{align*}
i &= n; \\
\text{while } & (i > 0) \\
& \quad \text{if } (v[i] + M[p[i]] \geq M[i-1]) \\
& \quad \quad \text{\{ // Interval } i \text{ is in solution} \\
& \quad \quad \quad i = p[i]; \\
& \quad \text{\} else} \\
& \quad i--; \\
\end{align*}
\]

7.C. OPTIMAL MATRIX Multiplication ORDERING (very simplified version of query optimization)

Only one strategy for multiplying two matrices – requires \( mnp \) scalar multiplications (and \( m(n - 1)p \) additions).
There are two strategies for multiplying three matrices:

$$\begin{bmatrix}
1 & 5 \\
A & 5 \\
\end{bmatrix} \begin{bmatrix}
1 & 5 \\
B & C \\
\end{bmatrix} \begin{bmatrix}
5 \\
\end{bmatrix}
$$

$$\text{(A B) C}$$

$$\begin{bmatrix}
1 & [AB] \\
5^s & 5 \\
\end{bmatrix} \begin{bmatrix}
5 & C \\
\end{bmatrix}
$$

10 $s$ total

$$\begin{bmatrix}
1 & 5 \\
ABC & 5^s \\
\end{bmatrix}
$$

$$\begin{bmatrix}
A & (B C) \\
5 & 5 \\
\end{bmatrix}
$$

25 $s$

50 $s$ total

Aside: Ways to parenthesize $n$ matrices? (Catalan numbers)

$$C_0 = 1 \quad \quad C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i} \quad \text{for} \; n \geq 0 \quad \quad C_n = \frac{1}{n+1} \binom{2n}{n}$$


Observation: Final tree cannot be optimal if any subtree is not.

1. Describe problem input.

$n$ matrices $\Rightarrow n + 1$ sizes

$$P_0 \begin{bmatrix}
P_1 \\
M_1 \\
\end{bmatrix} \begin{bmatrix}
P_2 \\
M_2 \\
\end{bmatrix} \cdots \begin{bmatrix}
P_n \\
M_n \\
\end{bmatrix}$$
2. Determine cost function and base case.

\[ C(i, j) = \text{Cost for optimally multiplying } M_i \ldots M_j \]

\[ C(i, i) = 0 \]

3. Determine general case.

Consider a specific case \( C(5, 9) \). The optimal way to multiply \( M_5 \ldots M_9 \) could be any of the following:

\[ C(5, 5) + C(6, 9) + P_{4}P_{5}P_{9} \]
\[ C(5, 6) + C(7, 9) + P_{4}P_{6}P_{9} \]
\[ C(5, 7) + C(8, 9) + P_{4}P_{7}P_{9} \]
\[ C(5, 8) + C(9, 9) + P_{4}P_{8}P_{9} \]

Compute all four and keep the smallest one.

Abstractly: Trying to find \( C(i, j) \)

\[
\begin{bmatrix}
P_k \\
P_{i-1}
\end{bmatrix}
\begin{bmatrix}
P_k \\
C(i, k)
\end{bmatrix}
\begin{bmatrix}
P_j \\
P_k
\end{bmatrix}
\begin{bmatrix}
P_{k+1, j}
\end{bmatrix}
\]

\[ C(i,j) = \min_{i \leq k < j} \left\{ C(i,k) + C(k+1,j) + P_{i-1}P_kP_j \right\} \]

4. Appropriate ordering of subproblems.

Since smaller subproblems are needed to solve larger problems, run value for \( j - i \) for \( C(i, j) \) from 0 to \( n-1 \). Suppose \( n = 5 \):

\[
\begin{array}{cccccc}
& 0 & 1 & 2 & 3 & 4 \\
C(1,1) & C(1,2) & C(1,3) & C(1,4) & C(1,5) \\
C(2,2) & C(2,3) & C(2,4) & C(2,5) \\
C(3,3) & C(3,4) & C(3,5) \\
C(4,4) & C(4,5) \\
C(5,5) \\
\end{array}
\]

5. Backtrace for solution – explicitly save the \( k \) value that gave each \( C(i,j) \).
http://ranger.uta.edu/~weems/NOTES2320/optMM.c

// Optimal matrix multiplication order using dynamic programming

#include <stdio.h>

int p[20];
int n;
int c[20][20];
int trace[20][20];

void tree(int left,int right,int indent)
{
    int i;

    if (left==right)
    {
        for (i=0;i<indent;i++)
            printf("   ");
        printf("%d
",left);
        return;
    }
    tree(trace[left][right]+1,right,indent+1);
    for (i=0;i<indent;i++)
        printf("   ");
    printf("%d %d
",left,right);
    tree(left,trace[left][right],indent+1);
}

int main()
{
    int i,j,k;
    int work;

    scanf("%d",&n);

    for (i=0;i<=n;i++)
        scanf("%d",&p[i]);

    for (i=1;i<=n;i++)
        c[i][i]=trace[i][i]=0;

    for (i=1;i<n;i++)
        for (j=1;j<=n-i;j++)
        {
            printf("Compute c[%d][%d]\n",j,j+i);
            c[j][j+i]=999999;
            for (k=j;k<j+i;k++)
            {
                work=c[j][k]+c[k+1][j+i]+p[j-1]*p[k]*p[j+i];
                printf(" k=%d gives cost %3d=c[%d][%d]+c[%d][%d]+p[%d]*p[%d]*p[%d]\n",k,work,j,k,j+i,k,j+i);
                if (c[j][j+i]>work)
                {
                    c[j][j+i]=work;
                    trace[j][j+i]=k;
                }
            }
            printf("  c[%d][%d]=%d,trace[%d][%d]=%d\n",j,j+i,c[j][j+i],j,j+i,trace[j][j+i]);
        }
    return 0;
}
printf("   ");
for (i=1;i<=n;i++)
  printf("   %3d   ",i);
printf("\n");
for (i=1;i<=n;i++)
{
  printf("%2d ",i);
  for (j=1;j<=n;j++)
    if (i>j)
      printf(" ------- ");
    else
      printf(" %3d %3d ",c[i][j],trace[i][j]);
  printf("\n");
}
tree(1,n,0);

It is straightforward to use integration to determine that the k loop body executes about $\frac{2}{3}$ times.
Compute \( c[1][4] \)
- \( c[1][4] = 113, \text{trace}[1][4] = 3 \)

Compute \( c[2][5] \)
- \( c[2][5] = 143, \text{trace}[2][5] = 3 \)

Compute \( c[3][6] \)
- \( c[3][6] = 185, \text{trace}[3][6] = 3 \)

Compute \( c[4][7] \)
- \( c[4][7] = 95, \text{trace}[4][7] = 3 \)

Compute \( c[1][5] \)
- \( c[1][5] = 118, \text{trace}[1][5] = 4 \)

Compute \( c[2][6] \)
- \( c[2][6] = 228, \text{trace}[2][6] = 4 \)

Compute \( c[3][7] \)
- \( c[3][7] = 152, \text{trace}[3][7] = 4 \)

Compute \( c[1][6] \)
- \( c[1][6] = 168, \text{trace}[1][6] = 5 \)

Compute \( c[2][7] \)
- \( c[2][7] = 209, \text{trace}[2][7] = 4 \)

Compute \( c[1][7] \)
- \( c[1][7] = 196, \text{trace}[1][7] = 4 \)
(Aside) Like optimal matrix multiplication, the order-preserving Huffman code problem mentioned in Notes 06 requires a solution with the leaves in alphabetic order. The cost function is based on minimizing the expected bits/symbol under this restriction:

$$C(\text{left}, \text{right}) = \sum_{i=\text{left}}^{\text{right}} p_i + \min_{\text{left} \leq k < \text{right}} \{C(\text{left}, k) + C(k+1, \text{right})\}$$

$$C(i,i) = 0$$

7.D. LONGEST COMMON SUBSEQUENCE (not substring, http://ranger.uta.edu/~weems/NOTES2320/LCS.c )

Has important applications in genetics research.

1. Describe problem input.

Two sequences:

$$X = x_1 x_2 \ldots x_m$$
$$Y = y_1 y_2 \ldots y_n$$

2. Determine cost function and base case.

$$C(i,j) = \text{length of LCS for } x_1 x_2 \ldots x_i \text{ and } y_1 y_2 \ldots y_j$$
$$C(i,j) = 0 \text{ if } i = 0 \text{ or } j = 0$$
3. Determine general case.

Suppose \( C(i, j) \) has

\[
\begin{array}{cccccc}
  x_1 & x_2 & \ldots & x_{i-1}A & y_1 & y_2 & \ldots & y_{j-1}A \\
\end{array}
\]

Since \( x_i = y_j \), \( C(i, j) = C(i-1, j-1) + 1 \)

Now suppose \( x_i \neq y_j \):

\[
\begin{array}{cccccc}
  x_1 & x_2 & \ldots & x_{i-1}A & y_1 & y_2 & \ldots & y_{j-1}B \\
\end{array}
\]

But ‘\( B \)’ may appear in \( x_1 x_2 \ldots x_{i-1} \) or ‘\( A \)’ may appear in \( y_1 y_2 \ldots y_{j-1} \):

\[
C(i, j) = \max\{C(i, j-1), C(i-1, j)\} \text{ if } x_i \neq y_j
\]

4. Appropriate ordering of subproblems.

Before computing \( C(i, j) \), must have \( C(i-1, j-1) \), \( C(i, j-1) \), and \( C(i-1, j) \) available.
Use \((m+1) \times (n+1)\) matrix to store \( C \) values.

5. Backtrace for solution – either explicitly save indication of which of the three cases was used or recheck \( C \) values.

Takes \( \Theta(mn) \) time. (Aside: Can be done using \( \Theta(m + n) \) space.)
Example:

```
ababab
aabbbaa
LCS is abaa, length==4
```

```
0 0 0 0 0 0 0
a 0 1 1 1 1 1 1
b 0 1 1 2 2 2
a 0 1 2 2 2 3 3
b 0 1 2 3 3 3 3
a 0 1 2 3 3 4 4
b 0 1 2 3 4 4 4
```

7.E. LONGEST INCREASING SUBSEQUENCE

Monotone: For an input sequence \( Y = y_1 \ldots y_n \), find a longest subsequence in increasing (\( \leq \)) order.

Strict: For an input sequence \( Y = y_1 \ldots y_n \), find a longest subsequence in strictly increasing (\( < \)) order.

Both versions may be solved inefficiently by reduction to LCS:

Monotone: \( \Theta(n^2) \) worst-case time by taking LCS of sequence and its elements sorted in ascending order.

```
1122346778
6178213472
LCS is 11347, length==5
```

```
6 1 7 8 2 1 3 4 7 2
0 0 0 0 0 0 0 0 0 0
1 0 0 1 1 1 1 1 1 1
1 0 0 1 1 1 2 2 2 2
2 0 0 1 1 1 2 2 2 2 3
2 0 0 1 1 1 2 2 2 2 3
3 0 0 1 1 1 2 2 3 3 3 3
4 0 0 1 1 1 2 3 3 4 4
6 0 1 1 1 1 2 3 4 4 4
7 0 1 1 2 2 2 3 4 5 5
7 0 1 1 2 2 2 3 4 5 5
8 0 1 1 2 3 3 3 4 5 5
```
Strict: $\Theta(mn)$ worst-case time, where $m$ is the number of unique integers occurring in input.

LCS is 12347, length==5

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>1</th>
<th>7</th>
<th>8</th>
<th>2</th>
<th>1</th>
<th>3</th>
<th>4</th>
<th>7</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Both versions may be solved in $\Theta(n \log n)$ worst-case time, using an appropriate DP cost function and $n$ binary searches.

Monotone (http://ranger.uta.edu/~weems/NOTES2320/LIS.c):

1. Describe problem input. $Y = y_1y_2 \ldots y_n$

2. Determine cost function and base case.

   $C(i) = \text{Length of longest increasing subsequence ending with } y_i$.

   $C(0) = 0$

3. Determine general case for cost function.

   $C(i) = 1 + \max_{j<i \text{ and } y_j \leq y_i} \left\{ C(j) \right\}$ (The $j$ that gives $C(i)$ may be saved for backtrace.)

4. Appropriate ordering of subproblems - iterate over the prefix length, saving $C$ and $j$ for each $i$.

\[
\begin{array}{cccccccccccc}
  i & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
  y_i & 60 & 10 & 70 & 80 & 20 & 10 & 30 & 40 & 70 & 20 \\
  C & 1 & 1 & 2 & 3 & 2 & 2 & 3 & 4 & 5 & 3 \\
  j & 0 & 0 & 2 & 3 & 2 & 2 & 6 & 7 & 8 & 6 \\
\end{array}
\]

5. Backtrace for solution.

Find the rightmost occurrence of the maximum $C$ value. The corresponding $y$ will be minimized.

Appears to take $\Theta(n^2)$, but binSearchLast() from Notes 1 may be used to find each $C$ and $j$ pair in $\Theta(\log n)$ time to give $\Theta(n \log n)$ overall:
// Initialize table for binary search for DP
bsTabC[0]=(-999999);    // Must be smaller than all input values.
bsTabI[0]=0;            // Index of predecessor (0=grounded)
for (i=1;i<=n;i++)
    bsTabC[i]=999999;   // Must be larger than all input values.

C[0]=0;     // DP base case
j[0]=0;

for (i=1;i<=n;i++)
{
    // Find IS that y[i] could be appended to.
    // See CSE 2320 Notes 01 for binSearchLast()
    k=binSearchLast(bsTabC,n+1,y[i]);
    C[i]=k+1;         // Save length of LIS for y[i]
    j[i]=bsTabI[k];   // Predecessor of y[i]
    bsTabC[k+1]=y[i]; // Decrease value for this length IS
    bsTabI[k+1]=i;
}

\[
\begin{array}{cccccccccccc}
  i &  1 &  2 &  3 &  4 &  5 &  6 &  7 &  8 &  9 & 10 \\
y_i & 60 & 10 &  70 &  80 &  20 &  10 &  30 &  40 &  70 & 20 \\
\end{array}
\]

**C**

\[
\begin{array}{cccccccc}
  j &  1 &  2 &  3 &  4 &  5 \\
\end{array}
\]

2. Determine cost function and base case.

\[
C(i) = \text{Length of longest strictly increasing subsequence ending with } y_i.
\]

\[
C(0) = 0
\]

3. Determine general case for cost function.

\[
C(i) = 1 + \max_{j<i \text{ and } y_j<y_i} \{C(j)\} \quad (\text{The } j \text{ that gives } C(i) \text{ must be saved to allow backtrace.})
\]

Finally, any \(y_i\) that is found by \text{binSearchLast()} will be ignored.

**Strict (http://ranger.uta.edu/~weems/NOTES2320/LSIS.c):** Similar to monotone with the following exceptions:

2. Determine cost function and base case.

\[
C(i) = \text{Length of longest strictly increasing subsequence ending with } y_i.
\]

\[
C(0) = 0
\]

3. Determine general case for cost function.

\[
C(i) = 1 + \max_{j<i \text{ and } y_j<y_i} \{C(j)\} \quad (\text{The } j \text{ that gives } C(i) \text{ must be saved to allow backtrace.})
\]

Finally, any \(y_i\) that is found by \text{binSearchLast()} will be ignored.
for (i=1; i<=n; i++)
{
    // Find SIS that y[i] could be appended to.
    // See CSE 2320 Notes 01 for binSearchLast()
    k=binSearchLast(bsTabC,n+1,y[i]);
    // y[i] only matters if it is not already in table.
    if (bsTabC[k]<y[i]) {
        C[i]=k+1;         // Save length of LIS for y[i]
        j[i]=bsTabI[k];   // Predecessor of y[i]
        bsTabC[k+1]=y[i]; // Decrease value for this length IS
        bsTabI[k+1]=i;
    }
    else
    {
        C[i]=(-1);        // Mark as ignored
        j[i]=(-1);
    }
}

<table>
<thead>
<tr>
<th>i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>y[i]</td>
<td>60</td>
<td>10</td>
<td>70</td>
<td>80</td>
<td>20</td>
<td>10</td>
<td>30</td>
<td>40</td>
<td>70</td>
<td>20</td>
</tr>
</tbody>
</table>

C

j

1

2

3

4

5

Suppose a value $k$ appears for $p$ entries in the cost function table ($C$) for an instance of the longest monotonically increasing subsequence problem. Going left-to-right across the corresponding input sequence values ($y_i$), which statement is true?

(Stated formally: For $i_1 < i_2 < \cdots < i_p$, suppose $C_{i_1} = C_{i_2} = \cdots = C_{i_p} = k$. Which statement is true regarding $y_{i_1}, y_{i_2}, \ldots, y_{i_p}$?)

_____  A. They are monotonically decreasing  
  B. They are strictly increasing  
  C. They are monotonically increasing  
  D. They are strictly decreasing
7.F. SUBSET SUM (http://ranger.uta.edu/~weems/NOTES2320/subsetSum.c)

Given a “set” of $n$ positive integer values, find a subset whose sum adds to a value $m$.

Optimization?

Enumerating subsets (combinations) would take exponential time.

1. Describe problem input. Array $S = S_1, S_2, ..., S_n$ and $m$.

2. Determine cost function and base case.

   $C(i) = \text{Smallest index } j \text{ such that there is some combination of } S_1, S_2, ..., S_j , \text{ that includes } S_j \text{ and sums to } i.$

   $C(0) = 0 \text{ (Will assume that } S_0 = 0)$

3. Determine general case for cost function.

   $C(i) = \min_{j \text{ s.t. } C(i-S_j) \text{ is defined} \text{ and } C(i-S_j) < j} \{j\}$

4. Appropriate ordering of subproblems:

   a. Iterate over $i$ looking backwards (like the cost function) to previous “finalized” solutions.

   b. (Aside, Dijkstra’s algorithm-like) Iterate over finalized $C(i)$ to compute $i + S_j$ for each $j > C(i)$ and attempt update forward. After updates, $C(i + 1)$ has final value.

   c. (Aside) Maintain ordered list of finalized solutions from using $S_1, S_2, ..., S_{j-1}$ and generate new ordered list that also uses $S_j$ to reach some new values.
5. Backtrace for solution - if $C(m)$ is defined, then backtrace using $C$ values to subtract out each value in subset. (Indices will appear in strictly decreasing order during backtrace.)

```
// Initialize table for DP
C[0]=0;  // DP base case
// For each potential sum, determine the smallest index such that its input value is in a subset to achieve that sum.
for (potentialSum=1; potentialSum<=m; potentialSum ++)
{
    for (j=1;j<=n;j++)
    {
        leftover=potentialSum-S[j];  // To be achieved with other values
        if (leftover>=0 &&  // Possible to have a solution
            C[leftover]<j)  // Indices are included in ascending order.
            break;
    }
    C[potentialSum]=j;
}
```

if (C[m]==n+1)
    printf("No solution\n");
else
{
    printf("Solution\n");
    printf(" i   S\n");
    printf("-------\n");
    for (i=m;i>0;i--=S[C[i]])
        printf("%3d %3d\n",C[i],S[C[i]]);
}

Example: $m=12, n=4$

```
i  0  1  2  3  4
S_i 0  3  6  7  9  [The $S_i$ values do not require ordering.]
i  0  1  2  3  4  5  6  7  8  9  10  11  12
C_i
```

Time is $\Theta(mn)$. Space is $\Theta(m)$. [What happens if $m$ and each $S_i$ are multiplied by the same constant?]
Another subset sum example ([http://ranger.uta.edu/~weems/NOTES2320/subsetSum.c](http://ranger.uta.edu/~weems/NOTES2320/subsetSum.c))

<table>
<thead>
<tr>
<th>i</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>i</td>
<td>C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
</tr>
</tbody>
</table>

Solution

<table>
<thead>
<tr>
<th>i</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

7.G. 0/1 (INDIVISIBLE, UNBOUNDED) KNAPSACK - OPTIMAL SOLUTION ([http://ranger.uta.edu/~weems/NOTES2320/knapsackTypeRS.c](http://ranger.uta.edu/~weems/NOTES2320/knapsackTypeRS.c))

CLRS 15.1 calls this the *rod-cutting problem*.

*n* item *types*, each with an integer size and value (CLRS - type = rod, size = length, value = price).

Unlike conventional version (Notes 6), *unlimited* supply of each type.

*m*, the integer capacity of the knapsack (length of the longer rod to be cut)

**Goal:** Select a combination from the unlimited supply of items that

1) maximizes the sum of the values, and

2) the sum of the sizes does not exceed *m*.

1. **Describe problem input.** Array *size* of *n* weights, array *val* of *n* values, and *m*.

2. **Determine cost function and base case.**
\( maxKnown(i) = \) Maximum sum of values achievable by some combination of items whose weights sum to no more than \( i \).

\( maxKnown(0) = 0 \)

3. Determine general case for cost function.

\[
maxKnown(i) = \max_{k \text{ s.t. } maxKnown(i-size_k) \text{ is defined}} \{ maxKnown(i-size_k) + val_k \}
\]

4. Appropriate ordering of subproblems - since goal is to compute \( maxKnown(m) \), extra cases could be computed. Use array of \( maxKnown(i) \) values along with unknown indicator to implement memoization (top-down).

```c
// From Sedgewick
int knap(int M, int level)
{
    int i, space, max, maxi = 0, t;
    for (i=0; i<level; i++)
        printf(".");
    if (maxKnown[M] != unknown)
    {
        printf("Reusing knap(%d)=%d\n", M, maxKnown[M]);
        return maxKnown[M];
    }
    printf("Start knap(%d)\n", M);
    for (i = 0, max = 0; i < N; i++)
        if ((space = M - items[i].size) >= 0)
            if ((t = knap(space, level+1) + items[i].val) > max)
                { max = t; maxi = i; }
    maxKnown[M] = max; itemKnown[M] = items[maxi];
    for (i=0; i<level; i++)
        printf(".");
    printf("Finish knap(%d)\n", M);
    return max;
}
```

// Since knap() uses memoization, a bottom-up loop is not needed.
printf("Maximum for %d is %d\n", m, knap(m, 0));

5. Backtrace for solution - backtrace using \( maxKnown \) and itemKnown.

Example: \( m=46 \)

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>11</td>
<td>13</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>val</td>
<td>10</td>
<td>14</td>
<td>16</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>20</th>
<th>22</th>
<th>24</th>
<th>27</th>
<th>29</th>
<th>33</th>
<th>35</th>
<th>46</th>
</tr>
</thead>
<tbody>
<tr>
<td>maxKnown</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>20</td>
<td>24</td>
<td>28</td>
<td>28</td>
<td>34</td>
<td>34</td>
<td>48</td>
</tr>
<tr>
<td>size</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>17</td>
<td>19</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>val</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>10</td>
<td>14</td>
</tr>
</tbody>
</table>

Time is \( \Theta(nm) \). Space is \( \Theta(m) \).
Start knap(46)
..Start knap(35)
...Start knap(24)
....Start knap(13)
......Start knap(2)
......Finish knap(2)
......Start knap(0)
......Finish knap(0)
......Finish knap(13)
....Start knap(11)
......Reusing knap(0)=0
....Finish knap(11)
...Start knap(7)
...Finish knap(7)
...Start knap(5)
...Finish knap(5)
..Finish knap(24)
..Start knap(22)
...Reusing knap(11)=10
...Start knap(9)
...Finish knap(9)
...Reusing knap(5)=0
...Start knap(3)
...Finish knap(3)
..Finish knap(22)
..Start knap(18)
...Reusing knap(7)=0
...Reusing knap(5)=0
...Start knap(1)
...Finish knap(1)
..Finish knap(18)
..Start knap(16)
....Reusing knap(5)=0
....Reusing knap(3)=0
..Finish knap(16)
.Finish knap(35)
..Start knap(33)
..Reusing knap(22)=20
..Start knap(20)
...Reusing knap(9)=0
...Reusing knap(7)=0
...Reusing knap(3)=0
...Reusing knap(1)=0
..Finish knap(20)
..Reusing knap(16)=14
..Start knap(14)
...Reusing knap(3)=0
...Reusing knap(1)=0
..Finish knap(14)
.Finish knap(33)
..Start knap(29)
...Reusing knap(18)=16
...Reusing knap(16)=14
...Start knap(12)
...Reusing knap(1)=0
...Finish knap(12)
..Start knap(10)
...Finish knap(10)
..Finish knap(10)
..Finish knap(29)
..Start knap(27)
...Reusing knap(16)=14
...Reusing knap(14)=14
...Reusing knap(10)=0
..Start knap(8)
..Finish knap(8)
..Finish knap(27)
Finish knap(46)
Maximum for 46 is 48

<table>
<thead>
<tr>
<th>i</th>
<th>size</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

i max size val

<table>
<thead>
<tr>
<th>i</th>
<th>max size</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>18</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>13</td>
</tr>
<tr>
<td>29</td>
<td>28</td>
<td>13</td>
</tr>
<tr>
<td>33</td>
<td>34</td>
<td>13</td>
</tr>
<tr>
<td>35</td>
<td>34</td>
<td>11</td>
</tr>
<tr>
<td>46</td>
<td>48</td>
<td>13</td>
</tr>
</tbody>
</table>

Solution has value 48:

<table>
<thead>
<tr>
<th>i</th>
<th>size</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

Unused capacity 1