no. of items	items	Bag Capacity (kg)				
available	available	0	1	2	3	4
0		0	0	0	0	0
1	A	0	0	5(A)	5(A)	5(A)
2	A,B	0	0	5(A)	6(B)	6(B)
3	A,B,C	0	0	5(A)	6(B)	7(C)
4	A,B,C,D	0	2(D)	5(A)	7(A,D)	8(B,D)

- We work through the table from top left corner to bottom right, going firstly from left to right across each row.
- The numbers in the body of the table are the *euro* values we can get, followed in brackets by the item choices.
- The number at the bottom right corner is the answer (8 euro: choose B,D)

To formalize this: let

PROF(i,j) ≡ Best profit we can obtain from choosing amongst the first i items, with bag capacity j.

With this definition:

- PROF(0,j) = 0 (there are no items to put in our bag).
- PROF(i,0) = 0 (our bag has capacity zero).

When we introduce a new item i (i.e. we go to a new row in the table above), we can either choose the new item or not:

Don't choose the new item i Then PROF(i,j) is just PROF(i-1,j) (i.e. our previous choice for best profit, not including new item i).

Choose the new item i From choosing the item i we get value v(i). But now the current bag capacity j is reduced to j - w(i). With this remaining bag capacity, the best profit we can obtain is $PROF(i-1, j-w(i))$. So our total best profit here is $v(i) + PROF(i-1, j-w(i))$.

Which of these two choices we make depends on which is best, i.e. we should choose the maximum! So we have

$$PROF(i,j) = MAX(PROF(i-1,j), v(i) + PROF(i-1, j-w(i)))$$

Note:

1. The Dynamic Programming Strategy here gives us all the intermediate results (if we wanted them!)
2. The way we order the items is immaterial to the final result, but affects the intermediate results.
3. What if our bag capacity (or item weights) were not integers? Then we need to make more "fine steps", in steps of e.g. 0.1, or 0.01 in the bag size.
4. The total complexity is just the size of the table (i.e. (n)(K)) so the algorithm is **O**(Kn). For example, for 20 items, bag capacity 50, we need to calculate (20)(50) numbers, i.e. 1,000 numbers. A Brute Force solution here requires $2^n = 2^{20} \approx 1,000,000$ calculations.

Addendum

The complexity quoted above for the Dynamic Programming solution to the Knapsack Problem *depends on how the solution is implemented*:

Iteration using a (2-dimensional) array We simply fill out the table above, from left to right and then from top to bottom: i.e.

- Calculate (write down!) PROF[0][0], PROF[0][1], PROF[0][2], ... (all these numbers are 0).
- Use the main equation to calculate PROF[1][0], PROF[1][1], PROF[1][2], ...
- Use the main equation to calculate PROF[2][0], PROF[2][1], PROF[2][2], ...
- etc.

- Use the main equation to calculate PROF[n][0], PROF[n][1], PROF[n][2], ... PROF[n][K]

Recursion We can just call PROF(i,j) as a *function*(of two variables i and j):

Base Case If either i or j is zero, return 0.

Recursive Step Use directly the main equation. (We must be careful here to use an `IF` statement to check that `w[i] < j`, otherwise we may end up with a negative index and hence infinite recursion.)

1. The initial call to PROF(n,K) leads to 2 new calls (specifically to PROF(n-1,K) and PROF(n-1, K-w[n]).
2. These then lead to 2 new calls each, a total of 4 calls to PROF(n-2,⁎).

3. This leads (in worst case) to 8 new calls to PROF(n-3,⁎).

In worst case this is $\mathbf{O}(2^n)$. In actual operation it will not be 2^n because in many cases (when w[i]>j: i.e. the item does not fit in the bag) the call to PROF(i-1, j-w[i]) will not be made. In the best case, w[i] is always greater than j (i.e. none of the items fit in to the bag!) and the problem is linear, $\mathbf{O}(n)$ with only n calls to the function.

So to summarize:

Small n, small K Doesn't matter which method you use!!

Small n, large K Use Recursion.

Large n, small K Use Iteration (2-dim array).

Large n, large K Use Iteration (2-dim array).

Nota Bene

When we say above, e.g. “small n, large K”, the comparison is not directly between n and K, rather between 2^n and K. For example, if $n = 20$ and $K = 1,000$, this might look like “small n, large K”, but when we compare 2^n to K,
$2^{10} = 1,048,576 \gg 1,000$, so here we should still use iteration:
K needs to be of the order of 2^n for recursion to win out.

Dynamic Programming Example 3: The Shortest Path Problem

We want to find the shortest path between any 2 cities on a (road) network. The distances of each “direct road” connection between neighboring cities are given.

Notation: we will write the distance between each of the n cities as an $n \times n$ matrix called A where $A(i, j)$ is the distance from city i to city j:

1. $A(i,j) = A(j,i)$, i.e. the matrix is symmetric (the distance from city i to city j is the same as the distance from city j to city i).
2. $A(i,i) = \infty \ \ \forall i$ (i.e. we don't want the traveller to go from a city to itself).
3. $A(i,j) = \infty$ if there is no direct road from i to j (i.e. if i and j are not neighboring cities).

Example

Consider the road network defined by the 4×4 matrix

$$\begin{pmatrix} \infty & \infty & 2 & 12 \\ \infty & \infty & \infty & 3 \\ 2 & \infty & \infty & 5 \\ 12 & 3 & 5 & \infty \end{pmatrix}.$$

The Dynamic Programming solution to this problem is commonly known as Floyd's Algorithm. We denote by

$$D_{ij}^{k} \equiv \text{The length of the shortest path from } i \text{ to } j \text{ where only intermediate vertices with label less than or equal to } k \text{ are allowed.}$$

Note that

$$D_{ij}^{0} = A_{ij}.$$

(since D_{ij}^{0} is the length of the shortest path from i to j with no intermediate vertices). Consider D_{ij}^{1}: This is the shortest path from i to j where we allow (possibly) as an intermediate vertex city 1. There are two possibilities:

We don't travel via city 1 Then the shortest path is given by D_{ij}^{0} (i.e. the previous shortest path with no intermediate vertex).

We travel via city 1 The the shortest path is made up of the addition of [shortest path from i to 1] + [shortest path from 1 to j], i.e.

$$D^0_{i1} + D^0_{1j}$$

(the superscript is zero here since both of these paths themselves do not involve any intermediate vertex).

Which of these two options we pick depends on which is shorter, i.e.

$$D^1_{ij} = \min(D^0_{ij}, D^0_{i1} + D^0_{1j})$$

We can use this equation to calculate the new $n \times n$ matrix D^1 from the original matrix $D^0 = A$.
In the general case, when we introduce a new city k, we have

$$D^k_{ij} = \min(D^{k-1}_{ij}, D^{k-1}_{ik} + D^{k-1}_{kj})$$

Using this equation repeatedly we obtain:

$$D^1 = \begin{pmatrix} \infty & \infty & 2 & 12 \\ \infty & \infty & \infty & 3 \\ 2 & \infty & 4 & 5 \\ 12 & 3 & 5 & 24 \end{pmatrix} \qquad D^2 = \begin{pmatrix} \infty & \infty & 2 & 12 \\ \infty & \infty & \infty & 3 \\ 2 & \infty & 4 & 5 \\ 12 & 3 & 5 & 6 \end{pmatrix}$$

$$D^3 = \begin{pmatrix} 4 & \infty & 2 & 7 \\ \infty & \infty & \infty & 3 \\ 2 & \infty & 4 & 5 \\ 7 & 3 & 5 & 6 \end{pmatrix} \qquad D^4 = \begin{pmatrix} 4 & 10 & 2 & 7 \\ 16 & 6 & 8 & 3 \\ 2 & 8 & 4 & 5 \\ 7 & 3 & 5 & 6 \end{pmatrix}$$

where at each stage we have highlighted in blue the numbers that have changed from the previous matrix. Note that

1. The diagonal entries in D^4 do not have much "real world" meaning: They denote the shortest distance from a city to itself, which is of course zero.
2. Since the graph is connected, there are no infinities in the final matrix.

Complexity Analysis

Recursion We can write the equation

$$D_{ij}^{k} = \min(D_{ij}^{k-1}, D_{ik}^{k-1} + D_{kj}^{k-1})$$

in the more functional/programming form

$$D(k,i,j) = \min(D(k-1,i,j), D(k-1,i,k)+D(k-1,k,j))$$

Used recursively, this breaks a problem of size n in to 3 new problems each of size $n-1$, then 9 problems of size $n-2$, 27 problems of size $n-3$ etc., so the complexity (no. of function calls we must make) is $\mathbf{O}(3^n)$.

Iteration using a table This is a 3-dimensional problem: We build up a "table of tables". We need to calculate n matrices, each of size $n \times n$, so we need to calculate n^3 numbers. Our complexity is $\mathbf{O}(n^3)$.

This comparison is “not fair”: In recursion above, we are only calculating the one pair shortest path, while in iteration we are calculating the all pairs shortest path. To make a fair comparison, to calculate the all-pairs shortest path using recursion requires $\mathbf{O}(n^2 3^n)$ function calls.

Some improvements on the Floyd Algorithm...

In the example above, each 4×4 matrix required 16 numbers to be calculated. We can reduce this as follows:

1. Since each matrix is guaranteed symmetric, we only need to calculate the “upper triangular” (and diagonal) part, and then copy to the lower triangular part. This means each time round we need only calculate 10 new numbers (4 diagonal, 6 off-diagonal) instead of 16.

2. In the transition from D_{ij}^{k-1} to D_{ij}^{k} we can be sure that none of the entries in row k or column k will change, i.e.

$$D_{ij}^{k} = D_{ij}^{k-1} \qquad \text{if } i = k \text{ or } j = k$$

(The reason follows: If $i = k$ or $j = k$, then in D_{ij}^{k} we are not really introducing a new *intermediate* city k since k is a source or destination city.) This saves the calculation of 4 numbers each time round.

Using these two observations, we reduce the number of new numbers we must calculate at each iteration from 16 to 6. (In the general case we can calculate the reduction is from n^2 to $n(n-1)/2$.)