Conceptual Algorithmic Template for Deterministic Dynamic Programming

Suppose we have $T$ stages and $S$ states. We set up two-level data structures $f$ and $x$ such that

- $f[t][s]$ holds the value of being in state $s$ and time stage $t$ (should be able to accept values of $t$ up to $T+1$)
- $x[t][s]$ holds the best decision to take in state $s$ and time stage $t$

The optimization procedure may then be organized along the following general lines

1. **Set up or read in problem data**
2. **Set up data structures for $x$ and $f$**
   
   For each state $s$, set $f[T+1][s]$ to be the value of being in state $s$ at the end of planning horizon. Depending on the problem and $s$, this value might be
   - Zero
   - Infinity for disallowed ending states ($+\infty$ for minimization, $-\infty$ for maximization)
   - Some “salvage” value (for example, what you could get by selling off excess inventory to a discounter at the end of the time horizon)

3. **Loop over stages $t = T, T-1, \ldots, 1$ (backward!)**
   
   Loop over all possible states $s$
   
   Set the current state value to be $+\infty$ for minimization, $-\infty$ for maximization
   
   Determine which decision moves are possible from stage $t$, state $s$
   
   Loop over all decisions $d$ that are possible from this state
   
   Evaluate the value of each move $d$ using the dynamic programing recursion formula
   
   If the decision $d$ improves on the best seen, record its value and the decision $d$
   
   Store the best decision for stage $t$, state $s$ in $x[t][s]$
   
   Store the corresponding optimal value in $f[t][s]$

4. **When done**, the optimal solution value is in $f[1][i]$, where $i$ is the initial state

To output the optimal sequence of decisions, start by setting $i$ to the initial state, then snake forward through the optimal sequence of states as follows:

5. **Loop over stages $t = 1, 2, \ldots, T$ (forward, this time)**
   
   Output the optimal decision $x[t][i]$
   
   Overwrite $i$ with the optimal state at time $t + 1$, computed from $i$ and $x[t][i]$