Woche 6 – Übersicht & Tricks

(Disclaimer: Die hier vorzufindenden Notizen haben keinerlei Anspruch auf Korrektheit oder Vollständigkeit und sind nicht Teil des offiziellen Vorlesungsmaterials. Alle Angaben sind ohne Gewähr.)

Allgemeines

- Korrektheitsbeweise mit Invarianten: Die Invariante (INV(i)) muss üblicherweise stark genug sein, dass die Korrektheit des Algorithmus aus der Gültigkeit der Invariant nach der letzten Iteration folgt.
- Programmieraufgaben: "Mein Algo ist doch richtig und trotzdem bekomme ich keine Punkte" - Die Test-Cases die ihr sehen (und mit denen ihr in Expert testen) könnt, sind extrem klein und enthalten normalerweise kaum Edge-Cases. Sie bestätigen euch nur, dass die Idee das Problem "weitestgehend richtig" löst. Allerdings kann euer Algorithmus trotzdem zu langsam sein oder bestimmte Edge-Cases nicht richtig handlen können. Daher: Nur der endgültige Score nach Submission bewertet euren Algorithmus. Also auch an der Prüfung zuerst diesen prüfen, bevor ihr an die nächste Aufgabe geht. Es ist extrem unangenehm, erst kurz vor Schluss der Prüfung die Programmieraufgaben zu submitten und zu realisieren, dass sie keine – oder kaum – Punkte geben. Alles schon vorgekommen.

Dynamic Programming: Allgemeines

Ihr werdet in der Vorlesung 6 verschiedene DP Konzepte kennenlernen (Längste aufsteigende Teilfolge, Längste gemeinsame Teilfolge, Minimale Editierdistanz, Matrixkettenmultiplikation, Subset-Sum, Knapsack). Mit den Ideen hinter diesen Konzepten werdet ihr sehr viele DP-Probleme sehr gut lösen können. Das heisst, wenn ihr ein Problem bekommt, solltet ihr euch zuerst denken (Beispiel):

"Der grösste Turm, den man mit gegebenen Legosteinen, bei denen nur kleinere auf grössere Steine gesetzt werden dürfen, bauen kann? – Kenn ich ein Konzept, dass ein **ähnliches Problem** löst? Klar, das ist ja wie «Längste aufsteigende Teilfolge», nur mit Legosteinen statt Zahlen".

Das funktioniert natürlich nicht immer. Manchmal ist die "Übersetzung" in ein bekanntes Problem nicht so leicht oder mitunter gar nicht wirklich möglich – zum Beispiel weil das Problem eine DP-

Lösung verlangt, die die bekannten Konzepte erweitert oder einfach nichts mit den bekannten Konzepten zu tun hat.

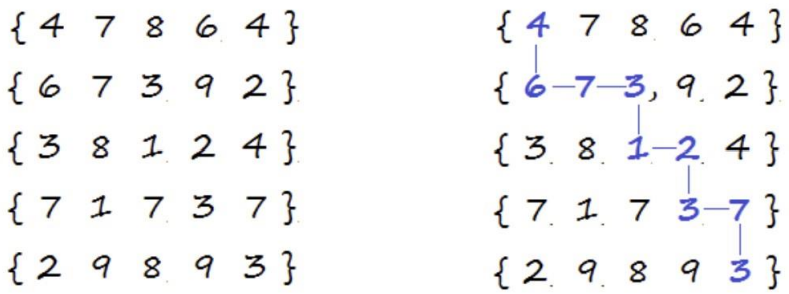
In diesem Fall müsst ihr den Ansatz selbst entwickeln. Das ist aber auch nicht so schwer, wie es am Anfang vorkommt. Ein selbst entwickelter Ansatz besteht immer aus den folgenden Schritten:

1. Überlegen, wie man das Problem in kleinere Teilprobleme zerlegen kann (Divide-and-Conquer Prinzip)
2. Überlegen, wie die Teilprobleme zusammenhängen
3. Überlappende Teilprobleme identifizieren
4. DP-Table konstruieren
5. DP-Table ausfüllen

Hier nun ein Beispiel für ein vergleichsweise einfaches DP-Problem:

Gegeben sei eine Matrix $A \in \mathbb{N}^{m \times n}$, die ein Spielfeld, dass aus $m * n$ Feldern besteht, beschreibt. Jeder einzelne Eintrag der Matrix, $A_{i,j}$ beschreibt die Kosten, um dieses Feld zu traversieren. Wir können von jedem Feld aus nach unten oder nach rechts gehen. Alle Kosten sind positiv. Was ist der geringste Preis, um von links oben nach rechts unten zu gelangen?

Beispiel:

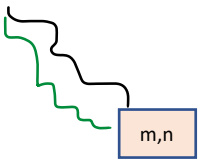


Der Preis des billigsten Weges von $A_{1,1}$ zu $A_{5,5}$ ist hier 36.

Die Lösung:

1. *Überlegen, wie man das Problem in kleinere Teilprobleme zerlegen kann (Divide-and-Conquer Prinzip):*

   Wir sehen: Jede Zelle kann bloss von links oder von oben erreicht werden. Also sehen wir für die Kosten, um A_{mn} zu erreichen: $Cost(m,n) = A_{m,n} \quad + \quad Min\{Cost(m-1,n), \quad Cost(m,n-1)\}$
   Denn wir wollen ja den geringsten Preis zahlen. In Worten bedeutet das: "Um von $A_{1,1}$ nach $A_{m,n}$ zu gelangen, müssen wir entweder die Kosten des Pfads von $A_{1,1}$ zu $A_{m-1,n}$ plus die Kosten von $A_{m,n}$ bezahlen **oder** die Kosten des Pfads von $A_{1,1}$ zu $A_{m,n-1}$ plus die Kosten von $A_{m,n}$ bezahlen. Und da wir uns für den billigsten Pfad interessieren, nehmen wir von den beiden Möglichkeiten die billigere."



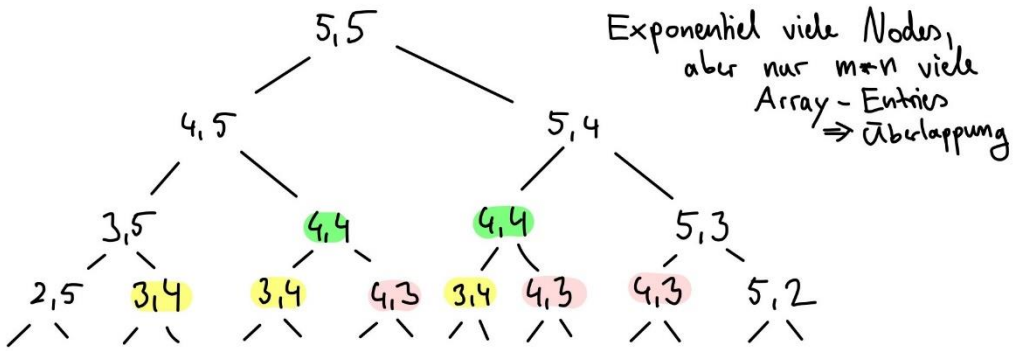
Hier bildlich: Ist der billigste Pfad zu dem oberen Nachbarn von $A_{m,n}$ der schwarze Pfad und der billigste Pfad zu dem linken Nachbarn von $A_{m,n}$ der grüne Pfad, ist der billigste Pfad zu $A_{m,n}$ das Min. Von Grün/Schwarz + $A_{m,n}$.

2. *Überlegen, wie die Teilprobleme zusammenhängen:*

   Das haben wir in diesem Fall schon im ersten Schritt weitestgehend festgestellt. Der billigste Pfad zu einem Feld kann in immer als optimale Wahl von den beiden billigsten Pfäden zu seinen beiden Nachbarn betrachtet werden. Der Preis eines Feldes ist dann der Preis des billigeren Pfad zu einem seiner beiden Nachbarn + der Preis von sich selbst.

3. *Überlappende Teilprobleme identifizieren:*

   Wir können uns als Beispiel einen Berechnungsbaum skizzieren:



   Wir sehen: Da es in dem Berechnungsbaum exponentiell viele Nodes gibt, aber unser Array nur $m * n$ viele Felder hat, muss es sein, dass wir den günstigsten Pfad zu bestimmten Feldern mehrfach berechnen. Das sind überlappende Teilprobleme.

4. *DP-Table konstruieren:*

   Wir orientieren uns ganz einfach an unserer Berechnungsvorschrift: Wir erstellen eine $m * n$ grosse Matrix als DP-Table. Ein Eintrag `DP[i][j]` dieser Matrix speichert Preis des billigsten Pfades von $A_{1,1}$ zum Feld $A_{i,j}$.

5. *DP-Table ausfüllen:*

   Wie ihr (Verweis Zusammenfassung Woche 5) bereits wisst, gibt es zwei Wege, DP-Probleme zu lösen: **Memoization** und **Bottom-Up**. Betrachten wir einfach beide Möglichkeiten:

   - Memoization:
     Ihr schreibt einen rekursiven (iterativ geht es natürlich auch) Algorithmus, $cost(m, n)$ berechnet und dabei jeden neu berechnen Wert in der DP-Matrix speichert – und jedes Mal, bevor er eine neue Lösung berechnet, prüft, ob es diesen Wert nicht schon in der DP-Matrix gibt, er also schon berechnet wurde.

   - Bottom-Up:
     Anstatt mit der Berechnung von $cost(m, n)$ zu beginnen, berechnen wir einfach erst alle anderen Pfade – da wir ja wissen, dass wir diese später eh gebrauchen werden. Dafür müssen wir allerdings zunächst ein paar Base-Cases definieren, auf denen wir aufbauen können: $cost(1,1) = A_{1,1}$. Jetzt können wir durch die DP-Table iterieren und für jeden Entry mithilfe unserer Berechnungsvorschrift

die Kosten bestimmen. Dabei ist es hier offensichtlich am sinnvollsten, Zeile für Zeile zu iterieren, da wir für die berechnung eines Feldes, seinen linken und seinen oberen Nachbarn brauchen. Die finale Lösung finden wir dann in `DP[m][n]`.

ACHTUNG: ICH HABE HIER IMMER EIN 1-INDEXING VERWENDET

Dynamic Programming: Lösungsformat in Aufgaben und Klausur:

In Aufgaben oder an der Prüfung wird von euch erwartet, dass ihr kurz und deutlich mit diesen Punkten euren DP-Algorithmus erklärt:

1. Dimensions of DP-Table
2. Definition of the DP-Table (What is the meaning of an entry)
3. Base Case(s) + Computation of an entry ("Berechnungsvorschrift")
4. Calculation order ("Which order allows the use of computations in previous steps")
5. Extracting the Solution ("Where in the DP-Table can you find the solution")
6. Runtime

Longest Increasing Subsequence (LIS/LAT):

Wir verwenden eine $1 \times n$ – DP-Table. In dem i-ten entry speichern wir das letzte Element einer aufsteigenden Teilfolge der Länge i. Diese DP-Table ist offensichtlich aufsteigend sortiert. Zur Berechnung des Arrays gibt es zwei Möglichkeiten, eine ist besser: Wir iterieren über das gegebene Array und gucken uns in der k-ten Iteration das Element a_k an. Dann suchen wir mit binärer Suche in unserer DP-Table die längste Teilfolge, an die a_k noch angehängt werden kann. Dann ersetzen wir $DP[m+1]$ durch das Element a_k **falls** $a_k < DP[m+1]$. Am Anfang setzen wir DP[0] auf -Inf und DP[k] = Inf für alle $k \leq n$. Die Lösung können wir schlussendlich finden, indem wir durch die DP-Table iterieren und den Index der längsten Teilfolge, für die wir ein Element haben, zurückgeben. Das ganze benötigt in n Iterationen je eine Binäre Suche, läuft also in $O(n * \log n)$.

Longest Common Subsequence (LCS/LGT):

Gegeben sind Wörter/Sequencen/Arrays A, B. Wir verwenden eine $m \times n$ – DP-Table. Das (i,j)-te Entry speichert die LCS für die Teilwörter $A[:i], B[:,j]$. Die Berechnungsvorschrift ist offensichtlich: Die LCS von zwei Wörtern $A[:i], B[:,j]$ ist $\max\{1 + LCS(A[:i-1], B[:,j-1])$ (falls $a_i = b_j$), $LCS(A[:i], B[:,j-1]),\ LCS(A[:i-1], B[:,j])\}$. Denn für zwei Wörter der Längen i,j ist die Teilfolge entweder die längste Teilfolge in den beiden möglichen Wortverkürzungen – oder die Folge, die an (i,j) endet. Da wir wieder links/oben/diagonal-oben stehende Einträge für das Berechnen des Entry $DP[i][j]$ brauchen, ist zeilenweises iterieren hier die beste Wahl. Die Lösung kann letzten Endes aus $DP[m][n]$ ausgelesen werden. Da wir über das gesamte Array iterieren, ist dieser Algorithmus in $O(m * n)$

Minimum Edit Distance (MED):

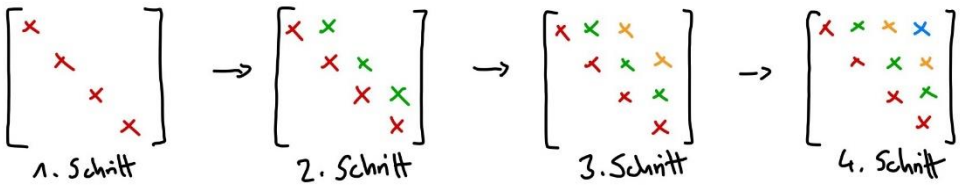
Gegeben sind sind Wörter/Sequencen/Arrays A, B. Wir verwenden eine $m \times n$ – DP-Table. Das (i,j)-te Entry speichert die MED für die Teilwörter $A[:i], B[:,j]$. Die Berechnungsvorschrift ist offensichtlich: Die MED von zwei Wörtern $A[:i], B[:,j]$ ist $\min\{MED(A[:i-1], B[:,j-1])$ (falls $a_i = b_j$), $MED(A[:i], B[:,j-1]) + 1,\ MED(A[:i-1], B[:,j]) + 1\}$. Denn für zwei Wörter der Längen i,j ist die MED entweder die MED der beiden Wortkürzungen plus das Einfügen je eines weiteren Buchstabens/Zeichens oder die MED für beide Wortkürzungen ohne editing, falls die Buchstaben/Zeichen gleich sind. Da wir wieder links/oben/diagonal-oben stehende Einträge für das Berechnen des Entry $DP[i][j]$ brauchen, ist zeilenweises iterieren hier die beste Wahl. Die Lösung

kann letzten Endes aus $DP[m][n]$ ausgelesen werden. Da wir über das gesamte Array iterieren, ist dieser Algorithmus in $O(m * n)$

Matrix Chain Multiplication (Konzept):

Gegeben ist eine Sequenz aus Objekten, $A_1, \dots, A_n$, die miteinander verrechnet werden, wobei bspw. die Reihenfolge der Berechnung entscheidend ist, beispielsweise Matrizen. Wir verwenden eine $n \times n$ – DP-Table. Das (i, j)-te Entry speichert die Kosten / optimale Lösung der Sub-berechnung von A_i bis A_j. Wir berechnen es folgendermassen: wir können die Sequenz von i nach j immer in je zwei Teil-Sequenzen aufteilen. Die optimale Lösung ist dann die Kosten/optimale Lösung der Teilsequenzen + die Kosten/optimale Lösung der Verrechnung der beiden Teilsequenzen. Das berechnen wir für jeden Trenner und dann nehmen wir davon den optimalen (normalerweise den maximalen) Wert. Also formal: $DP[i][j] = \min_{i \le p < j} \{M[i][p] + M[p+1][j] + Verrechnungskosten\}$.

Wir berechnen also nur die rechts-obere Dreiecksmatrix in DP. Dafür berechnen wir erst die Hauptdiagonale (Base Cases) und arbeiten und uns dann anschliessend in jeder "grossen" Iteration je einen Schritt in der Matrix nach rechts.



Extra Material:

Bonusaufgabe HS20:

A	1	3	2	1
3	2	1	1	B

Figure 1: Runner problem for a cost array of size 2×5.

Imagine, a runner wants to run from A to B in Fig. 1. There are two lanes available. One is represented by the first row and the other by the second row. On some sections, the first lane is faster than the second lane, and vice versa. The runner can change lanes at any time, but this costs 1 minute every time. In this exercise, you are supposed to provide a dynamic programming algorithm that computes the optimal track.

Formally, the problem is defined in terms of a cost array $c \in \mathbb{N}^{2 \times n}$. In Fig. 1 $n = 5$. Now, the runner starts at position $(1, 1)$ and wants to run to $(2, n)$. Running along a lane from field (i, j) to the field $(i, j + 1)$ requires $c_{i,j+1}$ minutes. Changing lanes from field $(1, j)$ to $(2, j)$ requires $1 + c_{2,j}$ minutes, and from field $(2, j)$ to $(1, j)$ requires $1 + c_{1,j}$ minutes.

Provide an algorithm using dynamic programming that computes the optimal track from A to B. Your algorithm should compute the optimal sequence $(1, 1), (i_1, j_1), \dots, (i_k, j_k), (2, n)$ and its cost (= the time required by the runner to run the sequence).

Dimension der DP-Table: 2 x n, somit linear

Definition der DP-Table: $DP[i,j]$ enthält die geringsten Kosten, um das Feld (i,j) zu erreichen, ohne rückwärts zu laufen.

Berechnung eines Eintrages: Als Base Case haben wir $DP[1,1] = 0$ und $DP[2,1] = 1 + c_{2,1}$

Die Einträge mit $j > 1$ berechnen wir wie folgt:

$$DP[1,j] = \min(DP[1,\mathrm{j}-1] + c_{1,j}, DP[2,\mathrm{j}-1] + c_{2,j} + c_{1,j} + 1)$$

und

$$DP[2,j] = \min(DP[2,\mathrm{j}-1] + c_{2,j}, DP[1,\mathrm{j}-1] + c_{1,j} + c_{2,j} + 1)$$

Berechnungsreihenfolge: Wir berechnen die Werte von DP Spalte für Spalte von links nach rechts.

Lesen des Ergebnisses: Das Ergebnis befindet sich zum Schluss in $DP[2,n]$.

Den optimalen Pfad finden wir durch Backtracking. Wir starten beim Feld (2,n) und schauen, ob wir zu diesem Feld mit oder ohne wechseln der line gekommen sind: Falls wir ohne wechseln dorthin gekommen sind, gilt DP[2,n] = DP[2,n − 1] + $c_{2,n}$ und wir fügen (2,n-1) zu dem optimalen Pfad hinzu. Falls wir die line gewechselt haben, gilt DP[2,n] = DP[1,n − 1] + $c_{1,n}$ + $c_{2,n}$ +1 und wir fügen (1, n) und (1,n-1) zu, optimalen Pfad hinzu und fahren fort, indem wir wieder genau gleich checken, wie wir zu (1,n-1) gelangt ist. Dies führen wir durch, bis entweder (1,1) oder (2,1) erreicht sind. Falls wir (2,1) erreicht haben, fügen wir (1,1) hinzu und sind fertig.

Laufzeit: Jeder Eintrag von *DP* kann in $\Theta(1)$ berechnet werden. Somit liegt die Laufzeit für die Berechnung von DP $\Theta(n)$. Die Lösung auszulesen, erfolgt in $\Theta(1)$, das Backtracking in $\Theta(n)$ da wir exakt n-1 mal checken müssen, von welchem Feld wir kamen. Somit liegt die Gesamtlaufzeit in $\Theta(n)$.

Bonusaufgabe HS21:

Consider the recurrence

$$F_1 = 1$$
$$F_n = \left(\min_{1 \le i < n} F_i^2 + F_{n-i}^2 \right) \bmod 3n \quad \text{for } n \ge 2,$$

where $a \bmod b$ is the remainder of dividing a by b.
Compute $F(n)$ bottom-up using dynamic programming and state the running time of your algorithm.

Dimension der DP-Table: Wir sehen, um einen wert F_k zu berechnen, brauchen wir nur die Werte $F_1, F_2, \dots, F_{k-1}$. Somit können wir das mit einem $1 \times n$ - Table lösen.

Definition der DP-Table: $DP[i]$ enthält F_i für $1 \le i \le n$

Berechnung eines Eintrages: Als Base Case haben wir $DP[1] = F_1 = 1$.
Die Einträge mit $2 \le i \le n$ berechnen wir wie folgt:

$$DP[i] = \left(\min_{1 \le j \le \left\lfloor \frac{i}{2} \right\rfloor} DP[j]^2 + DP[i-j]^2 \right) mod\ 3i$$

Berechnungsreihenfolge: Wir berechnen die Werte von DP vom kleinsten Index bis zum Grössten.

Lesen des Ergebnisses: Das Ergebnis befindet sich zum Schluss in $DP[n]$.

Laufzeit: Jeder Eintrag $DP[i]$ kann in $\Theta(i)$ berechnet werden, weil wir $\left\lfloor \frac{i}{2} \right\rfloor$ Werte berechnen müssen und das Minimum davon nehmen und noch von diesem Wert $mod\ 3i$ berechnen müssen. Somit ist die Gesamtlaufzeit: $\sum_{i=1}^{n} \Theta(i) = \Theta(n^2)$

Klausur FS18: (Programmieraufgabe T1)

ALGO-Turm

Gegeben sei eine Kiste mit $n \leq 5000$ *ALGO-Steinen*, welche von 1 bis n durchnummeriert sind (siehe untenstehendes Beispiel). Der i-te ALGO-Stein ist ein Quader mit Grundfläche $\ell_i \times b_i$ und Höhe h_i, mit $\ell_i, b_i \in \{1, 2, \ldots, 10000\}$ und $h_i \in \{0, 1, ..., 1000\}$.

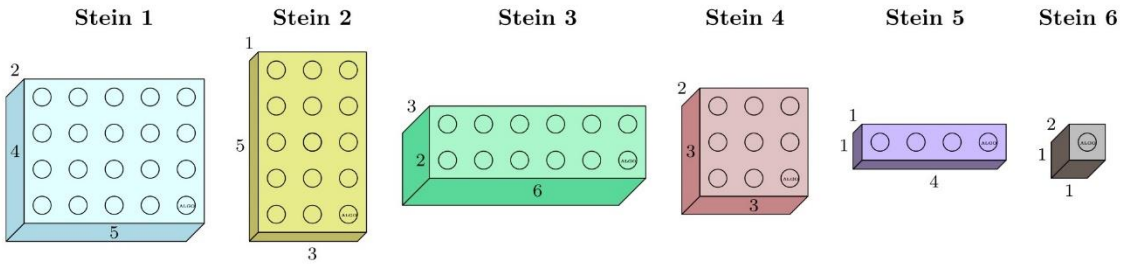
Aus diesen Steinen soll nun ein *möglichst hoher* Turm gebaut werden. Dazu dürfen die vorhandenen Steine aufeinandergestapelt werden, solange dabei *keine Überhänge* entstehen. Formaler: Der i-te Stein darf genau dann auf den j-ten Stein aufgesetzt werden, wenn mindestens eine der folgenden zwei Bedingungen gilt:

- $\ell_j \geq \ell_i$ und $b_j \geq b_i$; oder
- $\ell_j \geq b_i$ und $b_j \geq \ell_i$.

Intuitiv gesprochen entspricht die zweite Bedingung einer Drehung des i-ten Steins um 90° Grad.

Ein *ALGO-Turm* ist eine Liste $\langle t_1, t_2, ..., t_k \rangle$ von paarweise unterschiedlichen Indizes, sodass für alle $1 \leq i < k$ der Stein t_{i+1} auf den Stein t_i aufgesetzt werden darf. Die Höhe H eines solchen Turms ergibt sich als Summe der Höhen aller verwendeten Steine, das heisst $H := \sum_{i=1}^{k} h_{t_i}$.

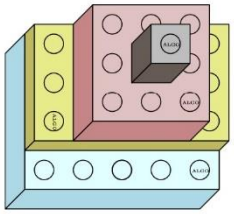
Ihre Aufgabe ist es nun, die Höhe eines höchstmöglichen ALGO-Turms zu berechnen, welcher mit den vorhandenen ALGO-Steinen gebaut werden kann. Dabei dürfen Sie annehmen, dass alle Grundflächen paarweise unterschiedlich sind, und dass die Steine nach absteigender Grundfläche sortiert sind. In anderen Worten gilt für $1 \leq i < n$, dass $\ell_i \cdot b_i > \ell_{i+1} \cdot b_{i+1}$.



Die oben abgebildeten $n = 6$ ALGO-Steine haben die folgenden Dimensionen:

Stein 1 $\ell_1 = 5, b_1 = 4, h_1 = 2;$
Stein 2 $\ell_2 = 3, b_2 = 5, h_2 = 1;$
Stein 3 $\ell_3 = 6, b_3 = 2, h_3 = 3;$
Stein 4 $\ell_4 = 3, b_4 = 3, h_4 = 2;$
Stein 5 $\ell_5 = 4, b_5 = 1, h_5 = 1;$
Stein 6 $\ell_6 = 1, b_6 = 1, h_6 = 2;$

Der höchste ALGO-Turm, welcher mit diesen Steinen gebaut werden kann, ist demzufolge $\langle 1, 2, 4, 6 \rangle$, siehe auch die untenstehende Abbildung. Er hat Höhe $h_1 + h_2 + h_4 + h_6 = 2 + 1 + 2 + 2 = 7$.



Hierbei musste Stein 2 (in gelb) *gedreht* werden, um ihn auf Stein 1 (in blau) aufsetzen zu können.

```java
import java.io.*;
import java.util.LinkedList;
import java.util.Scanner;
import java.lang.Math;
import java.lang.Integer;
import java.lang.String;

class Main {
    public static void main(String[] args) {
        // Uncomment the following two lines if you want to read from a file
        In.open("public/custom.in");
        Out.compareTo("public/custom.out");

        // Read the number of test cases
        int ntestcases = In.readInt();
        for(int t=0; t<ntestcases; t++)
        {
          int n = In.readInt();

          int[] l = new int[n+1];
          int[] b = new int[n+1];
          int[] h = new int[n+1];

          for(int i=1; i<=n; i++)
            l[i] = In.readInt();

          for(int i=1; i<=n; i++)
            b[i] = In.readInt();

          for(int i=1; i<=n; i++)
            h[i] = In.readInt();

          Out.println(solve(n, l, b, h));
        }

        // Uncomment the following line if you want to read from a file
        // In.close();
    }



    //n bezeichnet die Anzahl der ALGO Steine.
    //l[i] enthält die Länge des i-ten ALGO Steins, für i=1,...,n
    //b[i] enthält die Breite des i-ten ALGO Steins, für i=1,...,n
    //h[i] enthält die Höhe des i-ten ALGO Steins, für i=1,...,n
    static int solve(int n, int[] l, int[] b, int[] h)
    {
          //TODO: Geben Sie die Höhe des höchsten ALGO Turms, der aus den vorhandenen Steinen
          //gebaut werden kann, zurück.
          int[] LIS = new int[n+1];
          for(int i = 0; i<LIS.length; i++){
            LIS[i] = h[i];
          }
          int iP = 2, jP = 1, max = 0;
          while(iP<n+1){
            jP = 1;
            while(jP<iP){
              if(fits(jP, iP, b, l, h)){
                LIS[iP] = Math.max(LIS[iP], LIS[jP] + h[iP]);
              }
              jP++;
            }
            iP++;
          }
          for(int i = 0; i<LIS.length; i++){
            max = Math.max(max, LIS[i]);
          }
          return max;
    }
    static boolean fits(int j, int i, int[]b, int[]l, int[]h){
      return((l[j] >= l[i] && b[j] >= b[i]) || (l[j] >= b[i] && b[j] >= l[i]));
    }


}
```

Quellen:

Skript (Wie immer)

https://www.techiedelight.com/find-minimum-cost-reach-last-cell-matrix-first-cell/