

## ADP Aufgabe 2

### 2.A

#### Lernziele

1. Umgang mit verketteten Strukturen vertiefen.
2. Umgang mit abstrakten Datentypen durch Erweiterung lernen.
3. Die Basis für die Implementierung von Listen in Java kennenlernen.
4. Das Laufzeitverhalten eines Algorithmus aus einer doppelt logarithmischen Darstellung hochrechnen sowie den Verdoppelungstest anwenden können.
5. Ursachen für schlechtes Laufzeitverhalten identifizieren und beheben können.

#### 2.A1

Implementieren Sie eine doppelt verkettete Liste als generische Klasse `DoublyLinkedList` mit Wächterknoten für Anfang und Ende. Gehen Sie folgt vor.

1. Leiten Sie von `AbstractList` ab und implementieren Sie **nur die absolut notwendigen** Methoden. Das sind natürlich alle abstrakten Methoden, aber auch alle Methoden, die aus der doppelt verketteten Liste eine modifizierbare Liste machen (Hinzufügen, Löschen, Ändern von Elementen). Studieren Sie die Implementierung von `AbstractList` und identifizieren die **minimale** Menge an Methoden, die Sie überschreiben müssen. **Hinweis:** Iteratoren müssen Sie nicht implementieren. **In A1 geht es nicht um Zeit-Optimierung sondern nur um Funktionalität!**
2. Verwenden Sie eine private innere Klasse `Node` analog der Implementierung von `Stack` etc., und ergänzen Sie die Klasse `Node`, so dass eine doppelte Verkettung entstehen kann.
3. Zu den Wächterknoten (engl. guards): Wächterknoten sind Knoten ohne Inhalt (Inhalt ist `null`), die nur dazu dienen Anfang und Ende einer Liste zu markieren. Zu Beginn besteht die leere Liste nur aus zwei Wächterknoten, einem für `head` und einem für `tail`. Der Nachfolger von `head` ist `tail` und der Vorgänger von `tail` ist `head`. Wächterknoten haben den Vorteil, dass Operationen am Anfang und Ende der Liste nicht gesondert behandelt werden müssen. Das Kopieren und Adaptieren der Implementierung von `LinkedList` der Java Bibliothek ist nicht erlaubt.
4. Implementieren Sie auch den verallgemeinerten Kopier-Konstruktor für Collections.
5. Testen Sie Ihre Implementierung wie folgt:
  - a. Fügen Sie Elemente am Ende, am Anfang und in der Mitte der Liste ein.
  - b. Ändern Sie einen Wert an einem Index.
  - c. Löschen Sie Elemente am Anfang, am Ende und in der Mitte der Liste.
  - d. Löschen Sie die Liste vollständig (mit `clear`).
  - e. Iterieren Sie über die Liste und geben Sie die enthaltenen Elemente aus.
  - f. Erzeugen Sie einen Stream auf der Liste und geben Sie die enthaltenen Elemente aus.
  - g. Wandeln Sie den Inhalt der Liste in ein Array um.
  - h. Wandeln Sie eine Menge und eine `ArrayList` in eine `DoublyLinkedList` um. (Dazu benötigen wir den verallgemeinerten Kopier-Konstruktor!)

#### 2.A2

1. Führen Sie mit der Implementierung von `DoublyLinkedList` einen **Verdoppelungstest mit Gewichtung der Zeiten** durch. In dem Verdoppelungstest soll die Zeit, die für das Iterieren über die gesamte Liste benötigt wird, gemessen werden. Lassen Sie den Test arbeiten bis sich die gewichteten Zeiten ungefähr einpendeln. Ein Beispiel für das Gerüst eines Verdoppelungstests finden Sie im Skript.

2. Tragen Sie die Messungen in einer doppelt logarithmischen Darstellung ein und bestimmen Sie die Steigung (näherungsweise). Die Steigung sei  $b$ .
3. Berechnen Sie die Größe  $a$  der Formel  $T(N) = a * N^b$  und rechnen Sie für den darauffolgenden Verdoppelungsschritt die Laufzeit aus. Vergleichen Sie diese mit der tatsächlich gemessenen Zeit.
4. Führen Sie denselben Test mit der `LinkedList` von Java durch.
5. Identifizieren Sie die Ursache für das schlechte Laufzeitverhalten Ihrer Implementierung (mit Begründung) und verbessern Sie die Implementierung.

## 2.B Zusammenhangskomponenten

### Lernziele

1. Varianten der Algorithmen für die Lösung des Union-Find-Problems anwenden.
2. Laufzeitverhalten unterschiedlicher Implementierungen analysieren.
3. Abhängigkeit des Laufzeitverhaltens von der Eingabe verstehen.

### Aufgabenstellung

Sie bekommen die folgende Eingabe (Abfolge von Paaren) für ein Union-Find-Problem mit anfänglich  $N = 10$  Komponenten.

```
0 1
2 1
1 3
6 9
4 7
5 8
1 5
1 8
9 1
8 2
8 3
8 4
```

1. Zeichnen Sie für den Quick-Union-Algorithmus nach jedem `union`-Schritt den Inhalt des `id[]` Arrays. Zeichnen Sie auch den entsprechenden Wald von Bäumen.
2. Wie ist die Wachstumsordnung des Quick-Union-Algorithmus, wenn am Ende alle Komponenten in einer Zusammenhangskomponente liegen sollen?
3. Zeichnen Sie für den Weighted-Quick-Union-Algorithmus nach jedem `union` Schritt den Inhalt des `id[]` und des `sz[]` Arrays.
4. Wie ist die Wachstumsordnung des Weighted-Quick-Union-Algorithmus, wenn am Ende alle Komponenten in einer Zusammenhangskomponente liegen sollen?
5. Implementieren Sie Pfadkompression für den Quick-Union-Algorithmus und vergleichen Sie die Laufzeiten von Quick-Union und Quick-Union mit Pfadkompression mit Hilfe eines Verdoppelungstests.
6. Welches Kostenmodell wird bei der Bestimmung der Wachstumsordnung der Union-Find Algorithmen verwendet?
7. Geben Sie für alle Union-Find Algorithmen Beispiele für eine beste und eine schlechteste Eingabe.

## 2.C Speicherbedarf

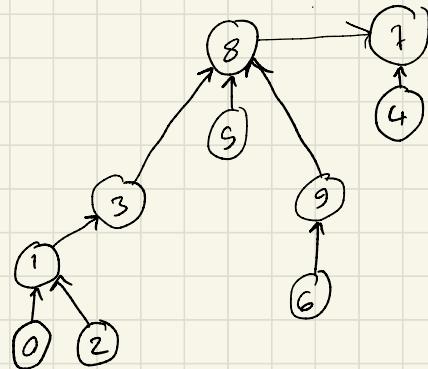
Berechnen Sie für `Integer`, `Date`, `Counter`, `int[]`, `double[]`, `double[][]`, `String`, `Node` und `Stack` (beide Implementierungen) den Speicherbedarf auf einem fiktiven 48 Bit Rechner. Gehen Sie davon aus, dass die Referenzen 6 Byte und der Objekt-Overhead 12 Byte belegen, dass auf ein Vielfaches von 6 Byte aufgefüllt wird und die Klasse `Node` als statische innere Klasse implementiert ist.

## 2.D Insertion- und Selection-Sort

1. Warum ist die Wachstumsordnung für Selection-Sort immer quadratisch? Wieviel Tauschoperationen benötigt Selection-Sort im besten Fall?
2. Im Skript finden Sie die Behauptung: **Für ein Array mit N Elementen unterscheidet sich die Laufzeit von Insertion- und Selection-Sort nur um einen kleinen Faktor.** Finden Sie Argumente, die die Behauptung stützen und berücksichtigen Sie dabei die Aussagen der Sätze zur Anzahl der Vergleichs- und Tauschoperationen für beide Algorithmen.
3. Gegeben die Eingabe V I R E N B E F A L L. Fertigen Sie für Selection- oder Insertion-Sort ein Handprotokoll der ersten 8 Sortierdurchläufe an. Markieren Sie das Element, dass aktuell betrachtet wird, die Bereiche die sich ändern und die Bereiche, die nicht betrachtet werden. Geben Sie auch die Werte der typischen Variablen des Sortierverfahrens an.

1. Zeichnen Sie für den Quick-Union-Algorithmus nach jedem `union`-Schritt den Inhalt des `id[]` Arrays. Zeichnen Sie auch den entsprechenden Wald von Bäumen.

<u>P</u>	<u>q</u>	0	1	2	3	4	5	6	7	8	9
0	1	1	1	2	3	4	5	6	7	8	9
2	7	7	7	1	3	4	5	6	7	8	9
1	3	1	3	1	3	4	5	6	7	8	9
6	9	1	3	1	3	4	5	9	7	8	9
4	7	1	3	1	3	7	5	9	7	8	9
5	8	1	3	1	3	7	8	9	7	8	9
1	5	1	3	1	8	7	6	9	7	8	9
1	8	1	3	1	8	7	6	9	7	8	9
7	1	1	3	1	8	7	8	8	7	8	8
8	2	1	3	1	8	7	8	8	7	8	8
8	3	1	3	1	8	7	6	8	7	8	8
8	4	7	3	1	8	7	8	8	7	8	8



2. Wie ist die Wachstumsordnung des Quick-Union-Algorithmus, wenn am Ende alle Komponenten in einer Zusammenhangskomponente liegen sollen?

Best Case:  $N + 3 \cdot (N-1)$       Worst Case:  $N + N^2$

3. Zeichnen Sie für den Weighted-Quick-Union-Algorithmus nach jedem `union` Schritt den Inhalt des `id[]` und des `sz[]` Arrays.

$p$	$q$	$0$	$1$	$2$	$3$	$4$	$5$	$6$	$7$	$8$	$9$
$0$	$1$	$0$	$1$	$2$	$3$	$4$	$5$	$6$	$7$	$8$	$9$
$0$	$1$	$0$	$0$	$2$	$3$	$4$	$5$	$6$	$7$	$8$	$9$
$2$	$1$	$0$	$2$	$2$	$3$	$4$	$5$	$6$	$7$	$8$	$9$
$1$	$3$	$0$	$2$	$2$	$2$	$4$	$5$	$6$	$7$	$8$	$9$
$6$	$9$	$6$	$2$	$2$	$2$	$4$	$5$	$6$	$7$	$8$	$6$
$4$	$7$	$0$	$2$	$2$	$2$	$4$	$5$	$6$	$4$	$8$	$6$
$5$	$8$	$0$	$2$	$2$	$2$	$4$	$5$	$6$	$4$	$5$	$6$
$1$	$9$	$0$	$2$	$2$	$2$	$4$	$2$	$6$	$4$	$5$	$6$
$1$	$8$	$0$	$2$	$2$	$2$	$4$	$2$	$6$	$4$	$5$	$6$
$9$	$1$	$0$	$2$	$2$	$2$	$4$	$2$	$2$	$4$	$5$	$6$
$8$	$2$	$0$	$2$	$2$	$2$	$4$	$2$	$2$	$4$	$2$	$6$
$8$	$3$	$0$	$2$	$2$	$2$	$4$	$2$	$2$	$4$	$2$	$6$
$8$	$4$	$0$	$2$	$4$	$2$	$4$	$2$	$2$	$4$	$2$	$6$

$pRoot == qRoot$

$p$	$q$	$0$	$1$	$2$	$3$	$4$	$5$	$6$	$7$	$8$	$9$
$0$	$1$	$7$	$7$	$7$	$7$	$7$	$7$	$7$	$1$	$7$	$7$
$0$	$1$	$2$	$1$	$1$	$1$	$1$	$1$	$1$	$1$	$1$	$1$
$2$	$1$	$2$	$1$	$2$	$1$	$1$	$1$	$1$	$1$	$1$	$1$
$1$	$3$	$2$	$2$	$2$	$1$	$1$	$1$	$1$	$1$	$1$	$1$
$6$	$9$	$2$	$2$	$2$	$1$	$1$	$1$	$2$	$1$	$1$	$1$
$4$	$7$	$2$	$2$	$2$	$1$	$2$	$1$	$2$	$1$	$7$	$1$
$5$	$8$	$2$	$2$	$2$	$1$	$2$	$2$	$2$	$1$	$1$	$1$
$1$	$5$	$2$	$2$	$3$	$1$	$2$	$2$	$2$	$1$	$1$	$1$
$1$	$8$	$2$	$2$	$3$	$1$	$2$	$2$	$2$	$1$	$1$	$1$
$7$	$1$	$2$	$2$	$4$	$1$	$2$	$2$	$2$	$1$	$1$	$1$
$8$	$2$	$2$	$2$	$5$	$1$	$2$	$2$	$2$	$1$	$1$	$1$
$8$	$3$	$2$	$2$	$6$	$1$	$2$	$2$	$1$	$1$	$1$	$1$
$8$	$4$	$2$	$2$	$6$	$1$	$3$	$2$	$1$	$1$	$1$	$1$

4. Wie ist die Wachstumsordnung des Weighted-Quick-Union-Algorithmus, wenn am Ende alle Komponenten in einer Zusammenhangskomponente liegen sollen?

$\log_2 N$

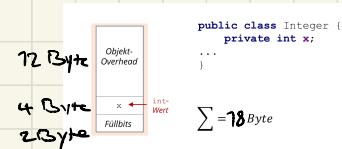
## 2.C Speicherbedarf

Berechnen Sie für Integer, Date, Counter, int[], double[], double[][][], String, Node und Stack (beide Implementierungen) den Speicherbedarf auf einem fiktiven 48 Bit Rechner. Gehen Sie davon aus, dass die Referenzen 6 Byte und der Objekt-Overhead 12 Byte belegen, dass auf ein Vielfaches von 6 Byte aufgefüllt wird und die Klasse Node als statische innere Klasse implementiert ist.

Integer: 12 Byte Object-Overhead

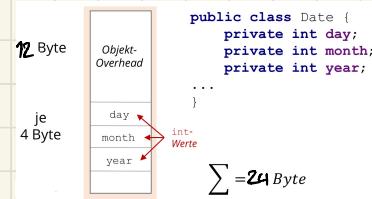
4 Byte Instanzvariablen

2 Byte Füllbits



Date: 12 Byte Object-Overhead

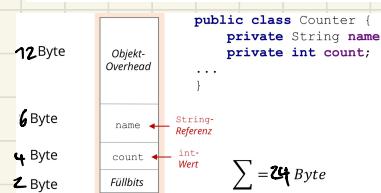
$4 \cdot 3 = 12$  Byte Instanzvariablen



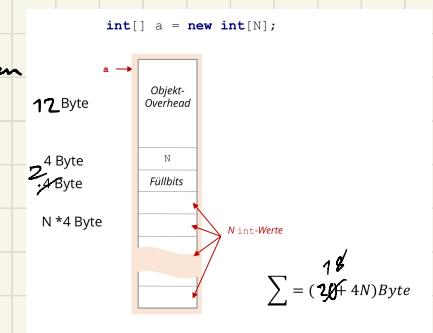
Counter: 12 Byte Object-Overhead

10 Byte Instanzvariablen

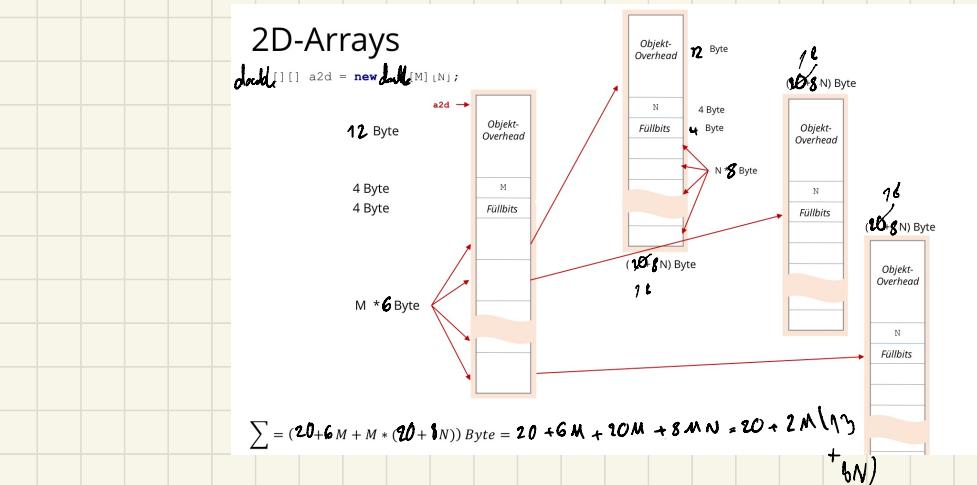
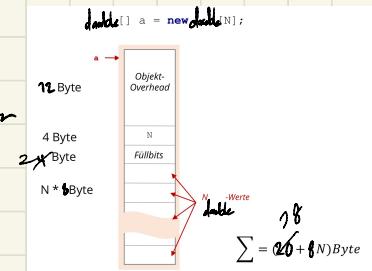
2 Byte Füllbits



`int []:` 12 Byte Object-Overhead  
 + 4 Byte int Wert (Länge des Arrays)  
 $N * 4$  Byte für die Instanzvariablen

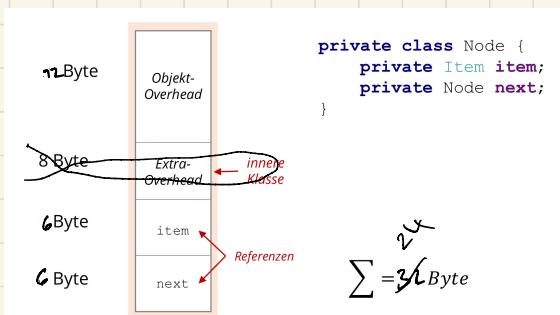
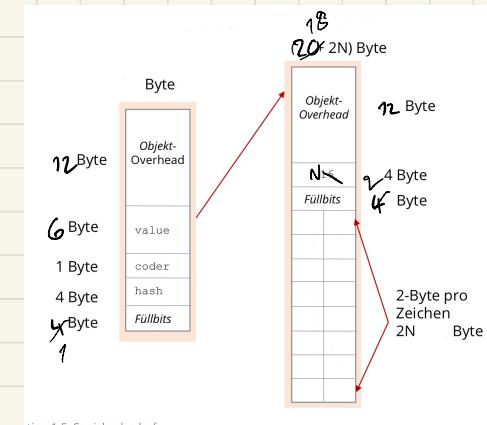


`double []:` 12 Byte Object-Overhead  
 + 4 Byte Länge des Arrays  
 $N * 8$  Byte für die Instanzvariablen



```
public final class String ... {
    private final byte[] value;
    private final byte coder;
    private int hash;
    ...
}
```

$$81 + (20 + 2N)$$



## Stack mit VKL (Speicherbedarf für N-Elemente)

- Stack mit int-Werten:
  - 12 Byte Objektoverhead
  - 4 Byte für die Länge
  - 6 Byte für die Referenz auf first
  - 2 Byte Füllbits
- $N * (24 \text{ Bytes pro Node})$ 
  - 12 Byte Objektoverhead
  - 6 Byte für die Referenz auf next
  - 4 Byte für den int Wert
  - 2 Byte Füllbits
- $\Sigma = (24 + 24N) \text{ Byte} \approx 24N \text{ Byte}$

- Stack mit Integer-Werten:
  - 24 Byte wie für den Stack mit int-Werten
  - $N * (52 \text{ Bytes pro Node})$ 
    - 12 Byte Objektoverhead
    - 6 Byte für die Referenz auf next
    - 6 Byte für die Referenz auf den Integer Wert
    - 18 Byte für den Integer Wert
  - $\Sigma = (24 + 92N) \text{ Byte} \approx 92N \text{ Byte}$

# Stack mit Resizing-Array (Speicherbedarf für N int-Werte)

- Stack mit int-Werten:
  - 12 Byte Objektoverhead
  - 4 Byte für die Länge
  - 6 Byte für die Referenz auf das int[] Array
  - 2 Byte Füllbits
  - Speicher für das int[] Array
    - 1. Array wird im nächsten Schritt vergrößert ( $N = a.length() - 1$ ):
      - i. 12 Byte Objektoverhead (Array)
      - ii. 4 Byte für die Länge des Arrays
      - iii. 2 Byte für ein Vielfaches von 6
      - iv. 4N Byte für die N-int Werte
      - v.  $\Sigma = 24 + 16 + 4N \text{ Byte} \approx 4N \text{ Byte}$
    - 2. Array Länge ist 4N und wird im nächsten Schritt um die ½ verkleinert.
      - i. 12 Byte Objektoverhead (Array)
      - ii. 4 Byte für die Länge des Arrays
      - iii. 2 Byte für ein Vielfaches von 6
      - iv. 4\*(4N) Byte für ein Array mit 4\*N int-Werten
      - v.  $\Sigma = (24 + 16 + 16N) \text{ Byte} \approx 2N \text{ Byte}$

# Stack mit Resizing-Array (Speicherbedarf für N-Integer Werte)

- Stack mit Integer-Werten:
  - 24 Byte, wie für Stack mit N int-Werten
  - Speicher für das Integer[] Array
    - 1. Array wird im nächsten Schritt vergrößert ( $N = a.length() - 1$ ):
      - i. 12 Byte Objektoverhead (Array)
      - ii. 4 Byte für die Länge des Arrays
      - iii. 2 Byte für ein Vielfaches von 6
      - iv. 6N Byte für die N Integer-Referenzen
      - v. 18 N Byte für die N-Integer Werte
      - vi.  $\Sigma = 24 + 18 + 24N \text{ Byte} \approx 6N \text{ Byte}$
    - 2. Array Länge ist 4N und um ½ verkleinert wird.
      - i. 12 Byte Objektoverhead (Array)
      - ii. 4 Byte für die Länge des Arrays
      - iii. 2 Byte für ein Vielfaches von 6
      - iv.  $4 * 6N = 24N \text{ Byte für die } 4N \text{ Integer-Referenzen des Arrays}$
      - v.  $18N \text{ Byte für die } N \text{ Integer-Werte}$
      - vi.  $\Sigma = (24 + 16 + 24N + 18N) \text{ Byte} \approx 2N \text{ Byte}$

## 2.D Insertion- und Selection-Sort

- Warum ist die Wachstumsordnung für Selection-Sort immer quadratisch? Wieviel Tauschoperationen benötigt Selection-Sort im besten Fall?

Selection-Sort ist ein iteratives Sortierverfahren, d.h. in jeder Iteration wird das Minimum gefunden und gegen das aktuelle Element getauscht und diese Tauschoperation findet immer (ohne Einschränkung) statt (tauschen sogar gegen sich selbst). In jeder Iteration werden alle verbleibenden Elemente durchlauft, da es in jeder Iteration immer weniger Elemente gibt, handelt es sich um eine arithmetische Reihe und die Summe der arithmetische Reihe ist quadratisch.

Es hat keinen Einfluss, ob der Array bereits sortiert ist oder nicht, Selection-Sort benötigt trotzdem alle Elemente durchzulaufen.

```
for (int j = i + 1; j < N; j++) {  
    if (less(a[j], a[min])) min = j;  
}  
  
exch(a, i, min);
```

findet immer statt

- Im Skript finden Sie die Behauptung: **Für ein Array mit N Elementen unterscheidet sich die Laufzeit von Insertion- und Selection-Sort nur um einen kleinen Faktor.** Finden Sie Argumente, die die Behauptung stützen und berücksichtigen Sie dabei die Aussagen der Sätze zur Anzahl der Vergleichs- und Tauschoperationen für beide Algorithmen.

ShellSort

$$\frac{N^2}{4} + \frac{N^2}{9} = \frac{2N^2}{9} = \underset{\textcolor{red}{\sim}}{\cancel{N^2}} \cdot \frac{2}{3}$$

$\Rightarrow$  L2V  
Insertion  
↳ Eingabe-abhängig

$$\cancel{\frac{N^2}{2}} + \cancel{N} \Rightarrow \begin{array}{l} \text{L2V} \\ \text{Selection} \end{array}$$

3. Gegeben die Eingabe V I R E N B E F A L L. Fertigen Sie für Selection- oder Insertion-Sort ein Handprotokoll der ersten 8 Sortierdurchläufe an. Markieren Sie das Element, dass aktuell betrachtet wird, die Bereiche die sich ändern und die Bereiche, die nicht betrachtet werden. Geben Sie auch die Werte der typischen Variablen des Sortierverfahrens an.

<u>i</u>	<u>j</u>	0	1	2	3	4	5	6	7	8	9	10
		V	I	R	E	N	B	E	F	A	L	L
1	0	I	V	R	E	N	B	E	F	A	L	L
2	1	I	R	V	E	N	B	E	F	A	L	L
3	0	E	I	R	V	N	B	E	F	A	L	L
4	2	E	I	N	R	V	B	E	F	A	L	L
5	0	B	E	I	N	R	V	E	F	A	L	L
6	2	B	E	E	I	N	R	V	F	A	L	L
7	3	B	E	E	F	I	N	R	V	A	L	L
8	0	A	B	E	E	F	I	N	R	V	L	L
9	6	A	B	E	E	F	I	L	N	R	V	L
10	7	A	B	F	E	F	I	L	L	N	R	V