20. Dynamic Programming II

Subset Sum Problem, Rucksackproblem, Greedy Algorithmus, Lösungen mit dynamischer Programmierung, FPTAS, Optimaler Suchbaum [Ottman/Widmayer, Kap. 7.2, 7.3, 5.7, Cormen et al, Kap. 15,35.5]

Aufgabe

Hannes und Niklas sollen eine grosse Menge Geschenke unterschiedlichen monetären Wertes bekommen.

Die Eltern wollen die Geschenke vorher so gerecht aufteilen, dass kein Streit eintritt. Frage: wie geht das?

Antwort: wer Kinder hat, der weiss dass diese Aufgabe keine Lösung hat.

536

Realistischere Aufgabe











Teile obige "Gegenstände" so auf zwei Mengen auf, dass beide Mengen den gleichen Wert haben.

Eine Lösung:











Subset Sum Problem

Seien $n \in \mathbb{N}$ Zahlen $a_1, \ldots, a_n \in \mathbb{N}$ gegeben.

Ziel: Entscheide, ob eine Auswahl $I \subseteq \{1, \dots, n\}$ existiert mit

$$\sum_{i \in I} a_i = \sum_{i \in \{1, \dots, n\} \setminus I} a_i.$$

Naiver Algorithmus

Prüfe für jeden Bitvektor $b = (b_1, \dots, b_n) \in \{0, 1\}^n$, ob

$$\sum_{i=1}^{n} b_i a_i \stackrel{?}{=} \sum_{i=1}^{n} (1 - b_i) a_i$$

Schlechtester Fall: n Schritte für jeden der 2^n Bitvektoren b. Anzahl Schritte: $\mathcal{O}(n \cdot 2^n)$.

Algorithmus mit Aufteilung

- Zerlege Eingabe in zwei gleich grosse Teile: $a_1, \ldots, a_{n/2}$ und $a_{n/2+1}, \ldots, a_n$.
- Iteriere über alle Teilmengen der beiden Teile und berechne Teilsummen $S_1^k,\ldots,S_{2^{n/2}}^k$ (k=1,2).
- Sortiere die Teilsummen: $S_1^k \leq S_2^k \leq \cdots \leq S_{2^{n/2}}^k$.
- \blacksquare Prüfe ob es Teilsummen gibt, so dass $S_i^1 + S_j^2 = \frac{1}{2} \sum_{i=1}^n a_i =: h$
 - Beginne mit $i = 1, j = 2^{n/2}$.
 - $\blacksquare \ \, \operatorname{Gilt} S^1_i + S^2_j = h \, \operatorname{dann} \, \operatorname{fertig}$
 - Gilt $S_i^1 + S_j^2 > h$ dann $j \leftarrow j 1$
 - Gilt $S_i^1 + S_j^2 < h$ dann $i \leftarrow i + 1$

540

Beispiel

Menge $\{1,6,2,3,4\}$ mit Wertesumme 16 hat 32 Teilmengen.

Aufteilung in $\{1,6\}$, $\{2,3,4\}$ ergibt folgende 12 Teilmengen mit Wertesummen:

 \Leftrightarrow Eine Lösung: $\{1,3,4\}$

Analyse

- Teilsummegenerierung in jedem Teil: $\mathcal{O}(2^{n/2} \cdot n)$.
- Sortieren jeweils: $\mathcal{O}(2^{n/2}\log(2^{n/2})) = \mathcal{O}(n2^{n/2})$.
- **Z**usammenführen: $\mathcal{O}(2^{n/2})$

Gesamtlaufzeit

$$\mathcal{O}\left(n\cdot 2^{n/2}\right) = \mathcal{O}\left(n\left(\sqrt{2}\right)^n\right).$$

Wesentliche Verbesserung gegenüber ganz naivem Verfahren – aber immer noch exponentiell!

Dynamische Programmierung

Aufgabe: sei $z=\frac{1}{2}\sum_{i=1}^n a_i$. Suche Auswahl $I\subset\{1,\ldots,n\}$, so dass $\sum_{i\in I}a_i=z$.

DP-Tabelle: $[0,\ldots,n] \times [0,\ldots,z]$ -Tabelle T mit Wahrheitseinträgen. T[k,s] gibt an, ob es eine Auswahl $I_k \subset \{1,\ldots,k\}$ gibt, so dass $\sum_{i \in I_k} a_i = s$.

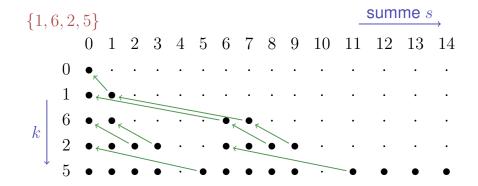
Initialisierung: T[0,0] = true. T[0,s] = false für s > 1.

Berechnung:

$$T[k,s] \leftarrow \begin{cases} T[k-1,s] & \text{falls } s < a_k \\ T[k-1,s] \lor T[k-1,s-a_k] & \text{falls } s \ge a_k \end{cases}$$

für aufsteigende k und innerhalb k dann s.

Beispiel



Auslesen der Lösung: wenn T[k,s] = T[k-1,s] dann a_k nicht benutzt und bei T[k-1,s] weiterfahren, andernfalls a_k benutzt und bei $T[k-1,s-a_k]$ weiterfahren.

544

Rätselhaftes

Der Algorithmus benötigt $\mathcal{O}(n \cdot z)$ Elementaroperationen.

Was ist denn jetzt los? Hat der Algorithmus plötzlich polynomielle Laufzeit?

Aufgelöst

Der Algorithmus hat nicht unbedingt eine polynomielle Laufzeit. z ist eine Zahl und keine Anzahl!

Eingabelänge des Algorithmus \cong Anzahl Bits zur *vernünftigen* Repräsentation der Daten. Bei der Zahl z wäre das $\zeta = \log z$.

Also: Algorithmus benötigt $\mathcal{O}(n \cdot 2^{\zeta})$ Elementaroperationen und hat exponentielle Laufzeit in ζ .

Sollte z allerdings polynomiell sein in n, dann hat der Algorithmus polynomielle Laufzeit in n. Das nennt man *pseudopolynomiell*.

NP

Man weiss, dass der Subset-Sum Algorithmus zur Klasse der *NP*-vollständigen Probleme gehört (und somit *NP-schwer* ist).

P: Menge aller in Polynomialzeit lösbarer Probleme.

NP: Menge aller Nichtdeterministisch in Polynomialzeit lösbarer Probleme.

Implikationen:

- NP enthält P.
- Probleme in Polynomialzeit *verifizierbar*.
- Unter der (noch?) unbewiesenen²⁷ Annahme, dass NP ≠ P, gibt es für das Problem *keinen Algorithmus mit polynomieller Laufzeit*.

Das Rucksackproblem

Wir packen unseren Koffer und nehmen mit ...

Zahnbürste

Zahnbürste

Zahnbürste

Hantelset

Luftballon

Kaffemaschine

- Kaffemaschine
- Taschenmesser
- Taschenmesser

- Oh jeh zu schwer.
- Ausweis

Ausweis

Hantelset

- Oh jeh zu schwer.
- Oh jeh zu schwer.

Wollen möglichst viel mitnehmen. Manche Dinge sind uns aber wichtiger als andere.

548

Rucksackproblem (engl. Knapsack problem)

Gegeben:

- Menge von $n \in \mathbb{N}$ Gegenständen $\{1, \ldots, n\}$.
- Jeder Gegenstand i hat Nutzwert $v_i \in \mathbb{N}$ und Gewicht $w_i \in \mathbb{N}$.
- Maximalgewicht $W \in \mathbb{N}$.
- Bezeichnen die Eingabe mit $E = (v_i, w_i)_{i=1,...,n}$.

Gesucht:

eine Auswahl $I\subseteq\{1,\ldots,n\}$ die $\sum_{i\in I}v_i$ maximiert unter $\sum_{i\in I}w_i\leq W$.

Gierige (engl. greedy) Heuristik

Sortiere die Gegenstände absteigend nach Nutzen pro Gewicht v_i/w_i : Permutation p mit $v_{p_i}/w_{p_i} \ge v_{p_{i+1}}/w_{p_{i+1}}$

Füge Gegenstände in dieser Reihenfolge hinzu $(I \leftarrow I \cup \{p_i\})$, sofern das zulässige Gesamtgewicht dadurch nicht überschritten wird.

Das ist schnell: $\Theta(n \log n)$ für Sortieren und $\Theta(n)$ für die Auswahl. Aber ist es auch gut?

34

²⁷Die bedeutenste ungelöste Frage der theoretischen Informatik!

Gegenbeispiel zur greedy strategy

$$v_1 = 1$$
 $w_1 = 1$ $v_1/w_1 = 1$ $v_2 = W - 1$ $w_2 = W$ $v_2/w_2 = \frac{W-1}{W}$

Greedy Algorithmus wählt $\{v_1\}$ mit Nutzwert 1. Beste Auswahl: $\{v_2\}$ mit Nutzwert W-1 und Gewicht W.

Greedy kann also beliebig schlecht sein.

Dynamic Programming

Unterteile das Maximalgewicht.

Dreidimensionale Tabelle m[i,w,v] ("machbar") aus Wahrheitswerten.

m[i, w, v] = true genau dann wenn

- Auswahl der ersten i Teile existiert ($0 \le i \le n$)
- deren Gesamtgewicht höchstens w ($0 \le w \le W$) und
- Nutzen mindestens v ($0 \le v \le \sum_{i=1}^n v_i$) ist.

552

Berechnung der DP Tabelle

Initial

- \blacksquare $m[i, w, 0] \leftarrow$ true für alle $i \geq 0$ und alle $w \geq 0$.
- $\blacksquare m[0, w, v] \leftarrow$ false für alle $w \ge 0$ und alle v > 0.

Berechnung

$$m[i,w,v] \leftarrow \begin{cases} m[i-1,w,v] \lor m[i-1,w-w_i,v-v_i] & \text{falls } w \ge w_i \text{ und } v \ge v_i \\ m[i-1,w,v] & \text{sonst.} \end{cases}$$

aufsteigend nach i und für festes i aufsteigend nach w und für festes i und w aufsteigend nach v.

Lösung: Grösstes v, so dass m[i, w, v] = true für ein i und w.

Beobachtung

Nach der Definition des Problems gilt offensichtlich, dass

- für m[i, w, v] = true gilt: $m[i', w, v] = \text{true } \forall i' \geq i$, $m[i, w', v] = \text{true } \forall w' \geq w$, $m[i, w, v'] = \text{true } \forall v' \leq w$.
- für m[i, w, v] = false gilt: m[i', w, v] = false $\forall i' \leq i$, m[i, w', v] = false $\forall w' \leq w$, m[i, w, v'] = false $\forall v' \geq w$.

Das ist ein starker Hinweis darauf, dass wir keine 3d-Tabelle benötigen.

DP Tabelle mit 2 Dimensionen

Tabelleneintrag t[i,w] enthält statt Wahrheitswerten das jeweils grösste v, das erreichbar ist²⁸ mit

- den Gegenständen $1, \ldots, i \ (0 \le i \le n)$
- bei höchstem zulässigen Gewicht w ($0 \le w \le W$).

Berechnung

Initial

 \bullet $t[0,w] \leftarrow 0$ für alle $w \geq 0$.

Berechnung

$$t[i,w] \leftarrow \begin{cases} t[i-1,w] & \text{falls } w < w_i \\ \max\{t[i-1,w],t[i-1,w-w_i]+v_i\} & \text{sonst.} \end{cases}$$

aufsteigend nach i und für festes i aufsteigend nach w.

Lösung steht in t[n, w]

Beispiel

$$E = \{(2,3), (4,5), (1,1)\} \qquad \underbrace{w} \qquad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7$$

$$\emptyset \qquad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0 \quad 0 \quad 0$$

$$(2,3) \qquad 0, \quad 0, \quad 3, \quad 3, \quad 3 \quad 3 \quad 3$$

$$i \qquad (4,5) \qquad 0, \quad 0 \quad 3, \quad 3 \quad 5, \quad 5 \quad 8, \quad 8$$

$$(1,1) \qquad 0 \quad 1 \quad 3 \quad 4 \quad 5 \quad 6 \quad 8 \quad 9$$

Auslesen der Lösung: wenn t[i,w]=t[i-1,w] dann Gegenstand i nicht benutzt und bei t[i-1,w] weiterfahren, andernfalls benutzt und bei $t[i-1,s-w_i]$ weiterfahren.

Analyse

556

Die beiden Algorithmen für das Rucksackproblem haben eine Laufzeit in $\Theta(n\cdot W\cdot \sum_{i=1}^n v_i)$ (3d-Tabelle) und $\Theta(n\cdot W)$ (2d-Tabelle) und sind beide damit pseudopolynomiell, liefern aber das bestmögliche Resultat.

Der greedy Algorithmus ist sehr schnell, liefert aber unter Umständen beliebig schlechte Resultate.

Im folgenden beschäftigen wir uns mit einer Lösung dazwischen.

²⁸So etwas ähnliches hätten wir beim Subset Sum Problem auch machen können, um die dünnbesetzte Tabelle etwas zu verkleinern

Approximation

Sei ein $\varepsilon\in(0,1)$ gegeben. Sei $I_{\rm opt}$ eine bestmögliche Auswahl. Suchen eine gültige Auswahl I mit

$$\sum_{i \in I} v_i \ge (1 - \varepsilon) \sum_{i \in I_{\mathsf{opt}}} v_i.$$

Summe der Gewichte darf ${\cal W}$ natürlich in keinem Fall überschreiten.

Andere Formulierung des Algorithmus

Bisher: Gewichtsschranke $w \to \text{maximaler Nutzen } v$ Umkehrung: Nutzen $v \to \text{minimales Gewicht } w$

- \Rightarrow Alternative Tabelle: g[i,v] gibt das minimale Gewicht an, welches
- lacktriangle eine Auswahl der ersten i Gegenstände ($0 \le i \le n$) hat, die
- einen Nutzen von genau v aufweist $(0 \le v \le \sum_{i=1}^n v_i)$.

560

Berechnung

Initial

- $g[0,0] \leftarrow 0$
- $g[0,v] \leftarrow \infty$ (Nutzen v kann mit 0 Gegenständen nie erreicht werden.).

Berechnung

$$g[i,v] \leftarrow \begin{cases} g[i-1,v] & \text{falls } v < v_i \\ \min\{g[i-1,v], g[i-1,v-v_i] + w_i\} & \text{sonst.} \end{cases}$$

aufsteigend nach i und für festes i aufsteigend nach v.

Lösung ist der grösste Index v mit $g[n, v] \leq w$.

Beispiel

$$E = \{(2,3), (4,5), (1,1)\} \qquad \qquad \underbrace{v} \\ 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9$$

$$\emptyset \quad 0 \leftarrow \infty \quad \infty$$

$$(2,3) \quad 0 \leftarrow \infty \quad \infty \quad 2 \leftarrow \infty \quad \infty \quad \infty \quad \infty \quad \infty$$

$$\downarrow (4,5) \quad 0_{\kappa} \quad \infty \quad \infty \quad 2_{\kappa} \quad \infty \quad 4_{\kappa} \quad \infty \quad \infty \quad 6_{\kappa} \quad \infty$$

$$(1,1) \quad 0 \quad 1 \quad \infty \quad 2 \quad 3 \quad 4 \quad 5 \quad \infty \quad 6 \quad 7$$

Auslesen der Lösung: wenn g[i,v]=g[i-1,v] dann Gegenstand i nicht benutzt und bei g[i-1,v] weiterfahren, andernfalls benutzt und bei $g[i-1,b-v_i]$ weiterfahren.

Der Approximationstrick

Pseudopolynomielle Laufzeit wird polynomiell, wenn vorkommenden Werte in Polynom der Eingabelänge beschränkt werden können.

Sei K>0 *geeignet* gewählt. Ersetze die Nutzwerte v_i durch "gerundete Werte" $\tilde{v_i}=\lfloor v_i/K \rfloor$ und erhalte eine neue Eingabe $E'=(w_i,\tilde{v_i})_{i=1...n}$.

Wenden nun den Algorithmus auf Eingabe E^\prime mit derselben Gewichtsschranke W an.

Idee

Beispiel K=5

Eingabe Nutzwerte

$$\begin{array}{c} 1,2,3,4,5,6,7,8,9,10,\ldots,98,99,100 \\ \rightarrow \\ 0,0,0,0,1,1,1,1,1,2,\ldots,19,19,20 \end{array}$$

Offensichtlich weniger unterschiedliche Nutzwerte

564

Eigenschaften des neuen Algorithmus

- Auswahl von Gegenständen aus E' ist genauso gültig wie die aus E. Gewicht unverändert!
- Laufzeit des Algorithmus ist beschränkt durch $\mathcal{O}(n^2 \cdot v_{\max}/K)$ $(v_{\max} := \max\{v_i | 1 \le i \le n\})$

Wie gut ist die Approximation?

Es gilt

$$v_i - K \le K \cdot \left\lfloor \frac{v_i}{K} \right\rfloor = K \cdot \tilde{v_i} \le v_i$$

Sei I_{opt}^{\prime} eine optimale Lösung von E^{\prime} . Damit

$$\begin{split} \left(\sum_{i \in I_{\mathsf{opt}}} v_i\right) - n \cdot K &\overset{|I_{\mathsf{opt}}| \leq n}{\leq} \sum_{i \in I_{\mathsf{opt}}} (v_i - K) \leq \sum_{i \in I_{\mathsf{opt}}} (K \cdot \tilde{v}_i) = K \sum_{i \in I_{\mathsf{opt}}} \tilde{v}_i \\ & \leq K \sum_{I_{\mathsf{opt}}' \mathsf{optimal}} K \sum_{i \in I_{\mathsf{opt}}'} \tilde{v}_i = \sum_{i \in I_{\mathsf{opt}}'} K \cdot \tilde{v}_i \leq \sum_{i \in I_{\mathsf{opt}}'} v_i. \end{split}$$

Wahl von K

Forderung:

$$\sum_{i \in I'} v_i \ge (1 - \varepsilon) \sum_{i \in I_{\mathsf{opt}}} v_i.$$

Ungleichung von oben:

$$\sum_{i \in I_{\mathsf{opt}}'} v_i \ge \left(\sum_{i \in I_{\mathsf{opt}}} v_i\right) - n \cdot K$$

Also:
$$K = \varepsilon \frac{\sum_{i \in I_{\mathsf{opt}}} v_i}{n}$$
.

Wahl von K

Wähle $K=arepsilon rac{\sum_{i\in I_{\mathrm{opt}}} v_i}{n}$. Die optimale Summe ist aber unbekannt, daher wählen wir $K'=arepsilon rac{v_{\mathrm{max}}}{n}$. 29

Es gilt $v_{\max} \leq \sum_{i \in I_{\text{opt}}} v_i$ und somit $K' \leq K$ und die Approximation ist sogar etwas besser.

Die Laufzeit des Algorithmus ist beschränkt durch

$$\mathcal{O}(n^2 \cdot v_{\text{max}}/K') = \mathcal{O}(n^2 \cdot v_{\text{max}}/(\varepsilon \cdot v_{\text{max}}/n)) = \mathcal{O}(n^3/\varepsilon).$$

FPTAS

Solche Familie von Algorithmen nennt man *Approximationsschema*: die Wahl von ε steuert Laufzeit und Approximationsgüte.

Die Laufzeit $\mathcal{O}(n^3/\varepsilon)$ ist ein Polynom in n und in $\frac{1}{\varepsilon}$. Daher nennt man das Verfahren auch ein voll polynomielles Approximationsschema *FPTAS - Fully Polynomial Time Approximation Scheme*

Optimale binäre Suchbäume

Gegeben: Suchwahrscheinlichkeiten p_i zu jedem Schlüssel k_i $(i=1,\ldots,n)$ und q_i zu jedem Intervall d_i $(i=0,\ldots,n)$ zwischen Suchschlüsseln eines binären Suchbaumes. $\sum_{i=1}^n p_i + \sum_{i=0}^n q_i = 1$.

Gesucht: Optimaler Suchbaum T mit Schlüsseltiefen $\operatorname{depth}(\cdot)$, welcher die erwarteten Suchkosten

$$C(T) = \sum_{i=1}^{n} p_i \cdot (\operatorname{depth}(k_i) + 1) + \sum_{i=0}^{n} q_i \cdot (\operatorname{depth}(d_i) + 1)$$
$$= 1 + \sum_{i=1}^{n} p_i \cdot \operatorname{depth}(k_i) + \sum_{i=0}^{n} q_i \cdot \operatorname{depth}(d_i)$$

minimiert.

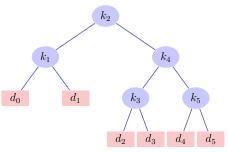
²⁹Wir können annehmen, dass vorgängig alle Gegenstände i mit $w_i > W$ entfernt wurden.

Beispiel

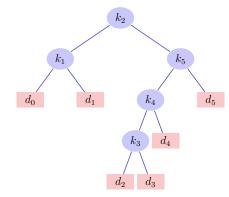
Beispiel

Erwartete Häufigkeiten

i	0	1	2	3	4	5
p_i		0.15	0.10	0.05	0.10	0.20
q_i	0.05	0.10	0.05	0.05	0.05	0.10



Suchbaum mit erwarteten Kosten 2.8

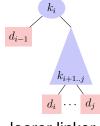


Suchbaum mit erwarteten Kosten 2.75

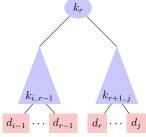
572

Struktur eines optimalen Suchbaumes

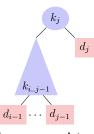
- Teilsuchbaum mit Schlüsseln k_i, \ldots, k_i und Intervallschlüsseln d_{i-1},\ldots,d_j muss für das entsprechende Teilproblem optimal sein. 30
- Betrachten aller Teilsuchbäume mit Wurzel k_r , $i \le r \le j$ und optimalen Teilbäumen k_i, \ldots, k_{r-1} und k_{r+1}, \ldots, k_i



leerer linker Teilsuchbaum



Teilsuchbäume links und rechts nichtleer



573

leerer rechter Teilsuchbaum

Teilsuchbäume

³⁰Das übliche Argument: wäre er nicht optimal, könnte er durch eine bessere Lösung ersetzt werden, welche die Gesamtlösung verbessert.

Erwartete Suchkosten

Sei $\operatorname{depth}_T(k)$ die Tiefe des Knotens im Teilbaum T. Sei k_r die Wurzel eines Teilbaumes T_r und T_{L_r} und T_{R_r} der linke und rechte Teilbaum von T_r . Dann

$$depth_T(k_i) = depth_{T_{L_r}}(k_i) + 1, (i < r)$$

$$depth_T(k_i) = depth_{T_{R_r}}(k_i) + 1, (i > r)$$

Erwartete Suchkosten

Seien e[i, j] die Kosten eines optimalen Suchbaumes mit Knoten k_i, \ldots, k_j .

Basisfall: e[i, i-1], erwartete Suchkosten d_{i-1}

Sei
$$w(i,j) = \sum_{l=i}^{j} p_l + \sum_{l=i-1}^{j} q_l$$
.

Wenn k_r die Wurzel eines optimalen Teilbaumes mit Schlüsseln k_i, \ldots, k_j , dann

$$\begin{split} e[i,j] &= p_r + (e[i,r-1] + w(i,r-1)) + (e[r+1,j] + w(r+1,j)) \\ \text{mit } w(i,j) &= w(i,r-1) + p_r + w(r+1,j) \\ e[i,j] &= e[i,r-1] + e[r+1,j] + w(i,j). \end{split}$$

Dynamic Programming

$e[i,j] = \begin{cases} q_{i-1} & \text{falls } j=i-1, \\ \min_{i \leq r \leq j} \{e[i,r-1] + e[r+1,j] + w[i,j]\} & \text{falls } i \leq j \end{cases}$

Berechnung

Tabellen $e[1\dots n+1,0\dots n], w[1\dots n+1,0\dots m], r[1\dots n,1\dots n]$ Initial

 $lacksquare e[i,i-1] \leftarrow q_{i-1}, w[i,i-1] \leftarrow q_{i-1} \text{ für alle } 1 \leq i \leq n+1.$

Berechnung

$$w[i,j] = w[i,j-1] + p_j + q_j$$

$$e[i,j] = \min_{i \le r \le j} \{e[i,r-1] + e[r+1,j] + w[i,j]\}$$

$$r[i,j] = \arg\min_{i \le r \le j} \{e[i,r-1] + e[r+1,j] + w[i,j]\}$$

für Intervalle [i,j] mit ansteigenden Längen $l=1,\ldots,n$, jeweils für $i=1,\ldots,n-l+1$. Resultat steht in e[1,n], Rekonstruktion via r. Laufzeit $\Theta(n^3)$.

57

Beispiel

		1				
p_i		0.15	0.10	0.05	0.10	0.20 0.10
q_i	0.05	0.10	0.05	0.05	0.05	0.10

