Methoden der Algorithmik WS1011
Prof. Dr. Michael Kaufmann

Christian Kniep

November 17, 2010

1 Vorlesung 10.11.2010

1.1 Wiederholung

- Anzahl sättigende Pushes : $\leq n * m$
- Anzahl nichtsättigende Pushes : $\leq O(n^2 * m)$
- Anzahl Relabel: $\leq 2n$

1.2 Lemma

Wir haben höchstens $\leq O(n^2 * m)$ nichtsättigende Pushes .

1.3 Beweis

Benutzen Potentialfunktion $\phi = \sum_{v \in A} d(v)$.(A=Menge aller aktiven Knoten)
Da $|A| \leq n, d(v) \leq$ für alle $v\ inV$ und $\phi \leq 2n^2$ anfangs.
Am Ende ist $\phi = 0$.

1. **Relabel**
   ϕ steigt um $\leq \epsilon$, wenn d(v) um ϵ steigt.
   Insgesamt steigt d(v) um $\leq 2n$, also insgesamt steigt ϕ um $\leq 2n^2$.

2. **sättigende Pushes**
   Diese erzeugen über (v,w) vtl. neuen Überfluss an w, w wird somit aktiv, ϕ steigt um $d(w) \leq 2n$. Also steigt ϕ durch sättigende Pushes um $\leq 2n^2 * m$.

3. **nichtsättigende Pushes**
   ϕ wird um d(v) erniedrigt, evtl. jedoch noch um d(w) erhöht, falls w vorher nicht aktiv war.
   Es gilt jedoch $d(v) \geq d(w) + 1 \Rightarrow \phi$ erniedrigt sich um ≥ 1.

Insgesammt:
Erhöhung von ϕ von $\leq 2n^2$ um $2n^2 + 2n^2 * m$ und k Erniedrigungen von ϕ um ≥ 1, führen zusammen zu 0-Potential.
$\Rightarrow k = O(n^2 * m), k =$ # nichtsättigende Pushes

1.4 Satz

Der generische Preflow-Push-Algorithmus läuft in $O(n^2 * m)$.

1.5 Varianten

Wahl der aktiven Knoten.

1. **FIFO**
   Warteschlange aktiver Knoten $O(n^3)$
2. **Highest Label**
   'höchster' Knoten im Netz $O(n^2 \sqrt{m})$
3. **Excess Scaling**
   'höchster' Füllstand zuerst $O(n * m + logc)$, c=grösste Kapazität

1.6 FIFO

1.6.1 Regel

Wende Push/Relabel solange auf denselben Knoten an, bis entweder e(v)=0 oder Relabel-Operation angewendet wurde.
DIe Liste der Knoten wird als FIFO-Queue gehalten.
Wird v relabelt, wird es hinten wieder angefügt.

1.6.2 Phasen

Arbeiten die Liste ab, bis um ersten Mal ein Knoten erscheint, der an dieser Phase schon teilgenommen hat.

1.6.3 Behauptung

Es gibt $\leq 4n^2 + 2n$ Phasen

1.6.4 Beweis

Betrachte jeweils die Änderung der Potentialfunktion.
$\phi = max\{d(v)|vistaktiv\}$

1. **Fall 1**
   Während einer Phase gibt es mindestens eine Relabel-Operation.
   ϕ steigt höchstens so viel wie der d-Wert max. (?). $\Rightarrow$ ϕ steigt in solchen Fällen um $\leq 2n^2$.

2. **Fall 2**
   Alle aktiven Knoten mit max. Distanz werden inaktiv, der maximale Distanzwert sinkt um mindestens 1.

Insgesamt gibt es also $\leq \underbrace{n + 2n^2 + (n + 2n^2)}_{2n+4n^2}$ Pushes.

1.6.5 Satz

Der FIFO-Preflow-Push-Algorithmus läuft in $O(n^3)$

1.7 Highest-Label

1.7.1 Regel

Schicke immer Fluss von einem aktiven Knoten mit höchstem Distanzwert.

1. Es gibt $O(n^3)$ nichtsättigende Pushes .
   Sei $h = max\{d(v)|vaktiv\}$
   Zuerst werden aktive Knoten v mit $d(v) = h$ betrachtet,
   dann $h - 1, h - 2, usw.$ Bei Relabelings fängt das ganze neu an.
   Es gibt jedoch nur $\leq 2n^2$ Relabelings.
   Gibt es n nichtsättigende Pushes hintereinander, dann sind wir fertig. Alle Knoten sind dann inaktiv.

Wie findet man aktive Knoten mit höchstem DIstanzlabel?
Halten Liste $(k) = v|vaktivundd(v) = k$.
Merken max. Index der den höchsten d-Wert angibt. Betrachte Listen (max), Listen $(max - 1)$, usw...
Relabel erhöht max.

1.8 Satz

Der Highest-Label-Preflow-Push macht $O(n^2 \sqrt{m}$ nichtsättigende Pushes und hat somit $O(n^3)$ Gesamtlaufzeit.

2 Vorlesung 17.11.2010

2.1 Überblick randomisierte Algorithmen

2.1.1 Definition

Ein Algorithmusder Entscheidungen zufällig trifft heisst Randomisierter Algorithmus

2.1.2 Vorteile

- Die Randomisierter Algorithmussind oft schneller als die deterministischen.
- sind (oft) viel einfacher
- keine oder wenige Worst-Case Eingaben

Es gibt zwei Typen von Randomisierter Algorithmus:

1. **L**as-Vegas-Algorithmen Liefern immer das korrekte Ergebnis!
   Durch zufällig getroffene Entscheidung variiert die Laufzeit.
   Ziel der Analyse: Laufzeit des LV Algorithmus ist mit hoher Wahrscheinlichkeit (m.h.W.) $\leq X$.
2. **M**onte-Carlo Algorithmus Knnen falsches Ergebnis liefern!
   Ziel der Analyse: Wahrscheinlichkeit für falsches Ergebnis $\leq X$ mit hoher Wahrscheinlichkeit (m.h.W.) .

Für Entscheidungsprobleme gibt es 2 Arten von Monte-Carlo Algorithmen.

- **e**inseitiger Fehler: Algorithmus kann sich nur in eine Richtung irren.
  Z.B. ist Ergebnis von Algo. 'Ja' ist dies immer korrekt (Fehlerwahrscheinlichkeit ist 0).
  ist Ergebnis 'Nein', dann ist die Fehlerwahrscheinlichkeit ≥ 0.
- **b**eidseitiger Fehler Algorithmus kann sich in beide Richtungen irren.
  Fehlerwahrscheinlichkeit in beiden Fällen $\geq o$.

2.1.3 Wahrscheinlichkeitstheorie

Sei Ω Ereignisraum. Ein Ereignis ist ELement aus Ω.
Wahrscheinlichkeit ist Abbilsung von Ereignissen auf reelle Zahlen, so dass die Summe über alle Ereignisse = 1.
$prob : \Omega \rightarrow \mathbb{R}_0^+ \sum_{w \in \Omega} prob(w) = 1$
Ist Ω endlich und es gilt $prob(w) = \frac{1}{|\Omega|}$ so heisst prob Gleichverteilung.

2.1.4 Beispiel

$\Omega = 1, 2, 3, 4, 5, 6$ und $prob(w) = \frac{1}{6}$ für alle $w \in \Omega$. $\Omega_{gerade} = 2, 4, 6$ und $prob(w) \in \Omega_{gerade} = 3\frac{1}{6} = \frac{1}{2}$

2.1.5 Markovsche Ungleichung

Für bestimmtes Ereignis X und dessen Erwartungswert E(X) gilt:
$prob(X \geq k * E(X)) \leq \frac{1}{k}$

2.2 Random Walk

BILD

2.2.1 Frage

$t_n(i) = E(AnzahlSchrittebisAbgrunderreicht)$ für $n = Feld, i = PositionderPerson$

2.2.2 Analyse

I
$t(0) = 0$
$t(1) = 1 + \frac{1}{2}t(0) + \frac{1}{2}t(2)$
$t(2) = 1 + \frac{1}{2}t(1) + \frac{1}{2}t(3)$
.
.
$t(n-1) = 1 + \frac{1}{2}t(n-2) + \frac{1}{2}t(n)$ $t(n) = 1 + t(n-1)$

II
$t(n-1) \leq 1 + \frac{1}{2}t(n-2) + \frac{1}{2}[1 + t(n-1)] = \frac{3}{2} + \frac{1}{2}t(n-2) + \frac{1}{2}t(n-1)$
Gleichung $*2*t(n-1)$:
$t(n-1) \leq 3 + t(n-2)$
$t(n-2) \leq 1 + \frac{1}{2}t(n-3) + \frac{1}{2}[3 + t(n-2)] = \frac{5}{2} + \frac{1}{2}t(n-3) + \frac{1}{2}t(n-2)$
$\leq 5 + t(n-3)$
Allgemein:
$t(i) \leq 2(n-1) + 1 + t(i-1)$
$t(2) \leq 2(n-2) + 1 + t(1) = 2n - 3 + t(1)$
$t(1) \leq 2(n-1) + 1 + \underbrace{t(0)}_{0} = 2n - 1$

III
$t(1) \leq 2n - 1$
$t(2) \leq 2n - 3 + 2n - 1 = 4n - 4$
$t(3) \leq 2n - 5 + 4n - 4 = 6n - 9$ Vermutung: $t(i) \leq 2in - i^2$

Vollstq̈ndige Induktion über i
IA: gezeigt für $i = 0, 1, 2, 3$
IS: $t(i+1) \leq 2(n-(i+1)) + 1 + t(i)$

$\leq 2n - 2i - 1 + 2ni - i^2$
$\leq 2n(i + 1) - (i + 1)^2$ □

$\Rightarrow$ Also gilt: $t_n(n) \leq 2n^2 - n^2 = n^2$
Klar ist, dass $t_n(i - 1) \leq t_n(i) \forall i$, daher ist $t_1(i) \leq n^2 \forall i \leq n$
$\Rightarrow$ Random Walk terminiert in $O(n^2)$ Schritten mit hoher Wahrscheinlichkeit (m.h.W.)

Einfache Anwendung von Random Walk: 2-SAT.
Gegeben: Boolscher Term in KNF, wobei alle Klauseln ≤ 2 Literale haben.
Beispiel: $(a_1 \vee a_2) \wedge (\bar{a}_3 \vee a_4) \wedge (a_5) \wedge ...$ Gesucht: Variablenbelegung aller a_i mit true/false, so dass alle Klauseln erfüllt sind. A1: $a_1 = t, a_2 = f, a_3 = t$
A2: $a_1 = t, a_2 = t, a_3 = t$
Hamming Abstand von zwei Variablenbelegungen ist die Anzahl der Variablen an denen sich die Belegung unterscheidet.

Frage Ist F erfüllbar?
2-SAT von Papadimitriou (1991)

1. Rate Variablenbelegung $a_1 ... a_n =: A$
2. Wiederhole ($2n^2$ Mal)
   if (A erfüllt alle Klauseln) return true
   while zufällig eine Klausel C, die von A nicht erfüllt wird
   und wähle einen der beidenen Literale aus C und negiere den Wert.
   $\Rightarrow$ dies ist neues A

3. return false; (mit hoher Wahrscheinlichkeit (m.h.W.) nicht erfüllbar)

Analyse
(richtige Belegung e ist 'Abgrund', $\bar{e}$ ist ganz links an der Wand und hat grössten Hamming Abstand.)

Angenommen es flipt erfüllende Belegung e (nur eine, falls mehr ist sogar besser =¿ mehr Abgründe)
$\bar{e}$ ist Belegung mit dem Hammingabstandn zu e (alle Variablen geflipt).

- Da der Algo immer eine unerfüllte Klausel wählt ist klar, dass mindestens eines der beiden Literale den entgegengesetzten Wert haben muss.

D.h. beim Ändern des richtigen Literals kommt man e um einen Schritt näher, verkürzt somit den Hammingabstand. $\Rightarrow$ $prob$ hierfür $\geq frac12$.

- Belegung $\bar{e}$ (ganz links) führt jeder Flip zu einer besseren Belegung.

$\Rightarrow$ entspricht Random Walk
X := bedeutet die Anzahl der Schleifendurchläufe für erfüllbare Instanzen
$E(X) \leq n^2$
Mit Mavkov-Ungleichung gilt:
$prob(X \geq k * E(X)) \leq \frac{1}{k}$
In unserem Fall:
$prob(X \geq 2n^2) \leq \frac{1}{2}$