Theoretische Informatik II

Prof. Dr. Andreas Goerdt

Sommersemester 2012 – 5. Juli 2012

Definitionen

Definition ($\mathcal{NP}$-hart, $\mathcal{NP}$-vollständig)

Eine Sprache A heißt $\mathcal{NP}$-*hart*, falls für alle $L \in \mathcal{NP}$ gilt $L \leq_p A$.
Eine Sprache A heißt $\mathcal{NP}$-*vollständig*, falls A $\mathcal{NP}$-hart ist und $A \in \mathcal{NP}$ gilt.

- Die $\mathcal{NP}$-harten Sprachen sind mindestens genauso schwierig zu entscheiden, wie alle anderen Sprachen aus der Klasse $\mathcal{NP}$.
- Die $\mathcal{NP}$-vollständigen Sprachen haben die Besonderheit, dass sie selbst in $\mathcal{NP}$ liegen. Sie besitzen damit eine gewisse „Allgemeingültigkeit“ bezüglich der anderen Probleme in $\mathcal{NP}$.

Anmerkung:

Falls(!) wir einen Polynomialzeitalgorithmus für eine $\mathcal{NP}$-vollständige Sprache hätten, dann hätten wir *automatisch* auch einen Polynomialzeitalgorithmus für alle anderen Sprachen aus $\mathcal{NP}$!
Die Frage, ob es einen solchen Algorithmus gibt, ist ein offenes Problem der Theoretischen Informatik.

Doch zunäschst die Frage: Gibt es überhaupt $\mathcal{NP}$-harte bzw. $\mathcal{NP}$-vollständige Sprachen?

Ja, gibt es!

Das SAT-Problem

Das *Erfüllbarkeitsproblem der Aussagenlogik*, kurz *SAT*, wollen wir hier als Sprache formulieren. (Später nehmen wir das dann nicht mehr so genau.)

Definition (SAT-Problem)

- *Eingabe:* Formel F der Aussagenlogik
- *Frage:* Ist F erfüllbar, d.h. gibt es eine Belegung der Variablen von F, so dass F den Wahrheitswert 1 annimt?

Wir haben es hier zunächst nur mit einer Ja/Nein-Frage zu tun. Die erfüllende Belegung ist nicht gefragt.

Das SAT-Problem (2)

Die zugehörige *Sprache* können wir dann wie folgt definieren.

- Aussagenlogische Formeln lassen sich über einem Alphabet Σ sicher geeignet kodieren.
- Alle erfüllbaren Formeln gehören zur Sprache.

$$\text{SAT} = \{\text{code}(F) \in \Sigma^* \mid F \text{ ist eine erfüllbare Formel der Aussagenlogik}\}$$

Satz von Cook

Satz (Cook)

Das Erfüllbarkeitsproblem der Aussagenlogik (SAT) ist $\mathcal{NP}$-vollständig.

Beweis, SAT $\in \mathcal{NP}$

Wir zeigen zunächst, dass SAT $\in \mathcal{NP}$.
Wir betrachten eine nichtdeterministische Turingmaschine (NTM) M, die wie folgt arbeitet:

Satz von Cook (2)

Beweis, SAT $\in \mathcal{NP}$ (Fortsetzung)

1. Die Eingabe lesen und feststellen, welche Variablen in der Formel vorkommen. Wir gehen davon aus, F enthält die Variablen $x_1, x_2, \ldots, x_k$.
2. In einem nichtdeterministischen Teil werden jetzt die Werte für die Variablen gesetzt. M „rät“ quasi die richtige Belegung für F.
   $\rightarrow$ Jetzt existieren praktisch 2^k mögliche Rechnungen, für jede Belegung eine.
3. Nun wird wieder deterministisch der Wahrheitswert von F unter der eben gewählten Belegung berechnet.
   $\rightarrow$ Falls $F = 1$ gilt, akzeptiert M die Eingabe.

Satz von Cook (3)

Beweis, SAT $\in \mathcal{NP}$ (Fortsetzung)

Die einzelnen Schritte lassen sich in Polynomialzeit ($p(|F|)$) ausführen. Falls F erfüllbar ist, d.h. es eine Belegung φ mit $F_{|\varphi=1}$ gibt, dann hat M mindestens eine akzeptierende Rechnung. Es gilt:

$$F \in \mathrm{SAT} \iff \begin{array}{l}\text{es gibt eine Rechnung von M}\\ \text{mit polynomieller Laufzeit,}\\ \text{die } F \text{ akzeptiert.}\end{array}$$

Damit ist SAT $\in \mathcal{NP}$!

Satz von Cook (4)

- Wir müssen zeigen, dass *für alle* $L \in \mathcal{NP}$ gilt $L \leq_p SAT$!
- Was wissen wir über L? — Nicht viel.
- L liegt in $\mathcal{NP}$. Also gibt es eine NTM M, die L entscheidet. Und es gibt ein Polynom $p(n)$, das die Rechenzeit von M beschränkt.

Idee

Wir nehmen für ein beliebiges L und ein Wort x die zugehörige NTM M und konstruieren eine Formel F mit der Eigenschaft

$$F \text{ ist erfüllbar} \iff x \in L.$$

Satz von Cook (5)

Zur NTM M:

- Wir nehmen o.B.d.A. an, dass ein einmal betretener Endzustand z_E von M nicht wieder verlassen wird.
- $p(n)$ ist das Polynom für die Rechenzeit von M.
- $x = x_1 x_2 \dots x_n \quad \in \Sigma^*$ ist die Eingabe von M.
- $\Gamma = \{a_1, \dots, a_l\}$ ist das Bandalphabet von M.
- $Z = \{z_1, \dots, z_k\}$ ist die Zustandsmenge von M.

Satz von Cook (6)

Beweis, SAT ist $\mathcal{NP}$-hart

Für die Formel F benutzen wir die folgenden Variablen:

- $\mathrm{zust}_{t,z}$ für $t = 0, \ldots, p(n)$ und alle $z \in Z$.
  Bedeutung: $\mathrm{zust}_{t,z} = 1 \iff$ Nach t Schritten ist M im Zustand z.
- $\mathrm{pos}_{t,i}$ für $t = 0, \ldots, p(n)$ und $i = -p(n), \ldots, p(n)$.
  Bedeutung: $\mathrm{pos}_{t,i} = 1 \iff$ Der Kopf von M befindet sich nach t Schritten an Position i.
- $\mathrm{band}_{t,i,a}$ für $t = 0, \ldots, p(n)$, $i = -p(n), \ldots, p(n)$ und alle $a \in \Gamma$.
  Bedeutung: $\mathrm{band}_{t,i,a} = 1 \iff$ Nach t Schritten steht auf Bandposition i das Zeichen a.
- Das sind polynomiell viele Variablen. Beachte: $|Z|$ und $|\Gamma|$ sind hier Konstanten.

Satz von Cook (7)

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

F stellt eine Reihe von Bedingungen dar, die alle erfüllt sein müssen, damit F wahr wird.

$$F = R \wedge A \wedge U_1 \wedge U_2 \wedge E$$

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

- *R, Randbedingungen:* Für *eine bestimmte Rechnung* gilt zu *jedem Zeitpunkt*:
  - Die NTM befindet sich in *genau einem* Zustand,
  - der Kopf steht an *genau einer* Position und
  - an jeder Bandposition steht *genau ein* Symbol.
  - Für die „genau eine“ Bedungung definieren wir Teilformeln der Art: $G(y_1, \ldots, y_m) = 1 \Longleftrightarrow$ Für genau ein i ist $y_i = 1$.

$$
\begin{aligned}
R \quad = \quad \bigwedge_t \Bigg[& G(\mathrm{zust}_{t,z_1}, \ldots, \mathrm{zust}_{t,z_k}) \wedge \\
& G(\mathrm{pos}_{t,-p(n)}, \ldots, \mathrm{pos}_{t,p(n)}) \wedge \\
& \bigwedge_i \Big(G(\mathrm{band}_{t,i,a_1}, \ldots, \mathrm{band}_{t,i,a_l}) \Big) \Bigg]
\end{aligned}
$$

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

- *A, Anfangsbedingungen:* Die Verhältnisse in der Startkonfiguration von M sehen so aus:
  - M ist im Zustand z_0,
  - der Kopf steht an Position 1,
  - ab Position 1 steht die Eingabe x auf dem Band,
  - sonst stehen überall Blanks ($\square$).

$$A = \mathrm{zust}_{0,z_0} \wedge \mathrm{pos}_{0,1} \wedge \Big(\bigwedge_{j=1}^{n} \mathrm{band}_{0,j,x_j} \Big) \wedge \Big(\bigwedge_{j=-p(n)}^{0} \mathrm{band}_{0,j,\square} \Big) \wedge \Big(\bigwedge_{j=n+1}^{p(n)} \mathrm{band}_{0,j,\square} \Big)$$

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

- U_1, *Übergangsbedingungen:* Beim Übergang vom Zeitpunkt t zu $t+1$, wenn M einen Rechenschritt macht, gilt folgendes:
  - An der Kopfposition wird ein Zeichen geschrieben,
  - der Kopf bewegt sich nach links ($d=-1$), rechts ($d=1$) oder nicht ($d=0$) und
  - der Zustand ändert sich.

$$U_1 = \bigwedge_{t,i,a} \Big[(\mathrm{zust}_{t,z} \wedge \mathrm{pos}_{t,i} \wedge \mathrm{band}_{t,i,a}) \rightarrow \bigvee_{(z',a',d)\in\delta(z,a)} (\mathrm{zust}_{t+1,z'} \wedge \mathrm{pos}_{t+1,i+d} \wedge \mathrm{band}_{t+1,i,a'}) \Big]$$

Beachte: Durch die Implikation ist bei „falschem“ Zustand, Position, Symbol auf dem Band der entsprechende Teil von U_1 automatisch wahr.

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

- *U_2, Übergangsbedingungen:* An den Bandpositionen, an denen sich der Kopf *nicht* befindet, ändert sich das Band auch nicht!

$$U_2 \quad = \quad \bigwedge_{t,i,a} \Big[(\neg \mathrm{pos}_{t,i} \wedge \mathrm{band}_{t,i,a}) \rightarrow \mathrm{band}_{t+1,i,a} \Big]$$

- *E, Endebedingung:* Irgendwann muss die NTM einmal einen Endzustand erreichen. Da wir annehmen können, dass der Endzustand nicht mehr verlassen weden kann, muss sich die TM speziell zum Zeitpunkt $p(n)$ in einem Endzustand befinden.

$$E \quad = \quad \bigvee_{z \in E} \mathrm{zust}_{p(n),z}$$

Satz von Cook (12)

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

Was haben wir gezeigt?

- *angenommen* $x \in L$*:*
  - Dann gibt es eine akzeptierende Rechnung der NTM der Länge $p(n)$.
  - Wir wählen alle Variablen so, „wie wir uns das gedacht haben.“
  - Dann haben alle Teilformeln von F und damit auch F den Wert 1. F ist also erfüllbar.

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

- *angenommen F ist erfüllbar:*
  - Dann gibt es eine Belegung φ der Variablen, so dass F wahr ist.
  - φ erfüllt R, d.h. wir können für jedes t die Variablen $\mathrm{zust}_{t,z}, \mathrm{pos}_{t,i}, \mathrm{band}_{t,i,a}$ sinnvoll als Konfiguration von M nach t Schritten interpretieren.
  - φ erfüllt auch A, d.h. zum Zeitpunkt $t = 0$ befindet sich M in ihrer Startkonfiguration. Daraus ergibt sich auch die Eingabe x.
  - φ erfüllt U_1 und U_2, d.h. es wird immer eine der möglichen Nachfolgekonfigurationen genommen.
  - φ erfüllt E, d.h. M „landet“ in einem Endzustand.

  Damit haben wir eine akzeptierende Rechnung der NTM rekonstruiert. Es gilt also $x \in L$.

Beweis, SAT ist $\mathcal{NP}$-hart (Fortsetzung)

Damit die Reduktion funktioniert, muss sich die Formel F in Polynomialzeit konstruieren lassen. Das gelingt sicher, falls F eine polynomiell beschränkte Länge hat. Für die Teilformeln gilt:

$$\begin{aligned} |R| &= O(p(n)^3) \\ |A| &= O(p(n)) \\ |U_1| &= O(p(n)^2) \\ |U_2| &= O(p(n)^2) \\ |E| &= O(1) \end{aligned}$$

Damit ist $|F| = O(p(n)^3)$. Für die Länge von R brauchen wir noch, dass die Teilformeln $G(y_1, \ldots y_m)$ die Länge $O(m^2)$ haben. ($\rightarrow$ Übungsaufgabe.) $\square$