Grundlagen der Theoretischen Informatik

Komplexitätstheorie (I)

5.07.2018

Viorica Sofronie-Stokkermans

e-mail: sofronie@uni-koblenz.de

Übersicht

1. Motivation
2. Terminologie
3. Endliche Automaten und reguläre Sprachen
4. Kellerautomaten und kontextfreie Sprachen
5. Turingmaschinen und rekursiv aufzählbare Sprachen
6. Berechenbarkeit, (Un-)Entscheidbarkeit
7. Komplexitätsklassen P und NP

Komplexitätstheorie

Inhalt

- Definition der berühmten Klassen **P** und **NP**.
- Begriff der **Reduktion**: ein Problem (eine Sprache) wird auf ein zweites reduziert. Das erste Problem ist dann höchstens so schwer wie das zweite.
- Der Begriff eines **NP-schweren** Problems.
- Einige Probleme der Graphentheorie: sie sind **NP-vollständig**.
- Die wichtigsten **Komplexitätsklassen** und ihre Struktur.

Komplexitätstheorie

- Die Struktur von PSPACE
- Vollständige und harte Probleme
- Beispiele

Komplexitätstheorie: Motivation

1. Sortieralgorithmen: Bubble Sort, Quicksort, ...
2. Erfüllbarkeitsproblem:
   Gibt es eine erfüllende Belegung für die Variablen $\{x_1, x_2, \ldots x_n\}$?
3. Graphenprobleme:
   Gibt es einen hamiltonschen Kreis in einem Graphen?
   Sind zwei Knoten in einem Graphen voneinander erreichbar?

Komplexitätstheorie

Welche Arten von Komplexität gibt es?

- **Zeit**
- **Speicher**

DTIME und NTIME

Definition [NTIME(T(n)), DTIME(T(n))]

Basismodell: k-DTM $\mathcal{M}$ (ein Band für die Eingabe).

Wenn $\mathcal{M}$ mit jedem Eingabewort der Länge n höchstens $T(n)$ Schritte macht, dann wird sie $T(n)$-**zeitbeschränkt** genannt.
Die von $\mathcal{M}$ akzeptierte Sprache hat **Zeitkomplexität** $T(n)$
(tatsächlich meinen wir $\max(n+1, \lceil T(n) \rceil)$).

- **DTIME**(T(n)) ist die Klasse der Sprachen, die von $T(n)$-zeitbeschränkten DTMs akzeptiert werden.
- **NTIME**(T(n)) ist die Klasse der Sprachen, die von $T(n)$-zeitbeschränkten NTMs akzeptiert werden.

DSPACE und NSPACE

Definition [NSPACE(S(n)), DSPACE(S(n))]

Basismodell: k-DTM $\mathcal{M}$, davon ein spezielles Eingabeband (**offline DTM**).

Wenn $\mathcal{M}$ für jedes Eingabewort der Länge n maximal $S(n)$ Zellen auf den Ablagebändern benutzt, dann heißt $\mathcal{M}$ $S(n)$-**speicherbeschränkt.**
Die von $\mathcal{M}$ akzeptierte Sprache hat **Speicherkomplexität** $S(n)$
(tatsächlich meinen wir $\max(1, \lceil S(n) \rceil)$)

- **DSPACE**(S(n)) ist die Klasse der Sprachen, die von $S(n)$-speicherbeschränkten DTMs akzeptiert werden.
- **NSPACE**(S(n)) ist die Klasse der Sprachen, die von $S(n)$-speicherbeschränkten NTMs akzeptiert werden.

DSPACE und NSPACE

Wieso eine *offline*-Turing-Maschine?

Bandbeschränkung von weniger als linearem Wachstum.

Beispiel Zu welcher Zeit-/Speicherkomplexitätsklasse gehört

$$\mathcal{L}_{\text{mirror}} := \{wcw^R : \ w \in (0+1)^*\},$$

also die Menge aller Wörter, die um den mittleren Buchstaben c gespiegelt werden können?

DSPACE und NSPACE

Beispiel (Forts.)

Zeit: **DTIME**(n+1). Die Eingabe rechts vom c wird einfach in umgekehrter Reihenfolge kopiert. Wenn das c gefunden wird, wird einfach der übrige Teil (das w) mit der Kopie von w auf dem Band verglichen.

Speicher: Die gerade beschriebene Maschine liefert eine Schranke von **DSPACE**(n).

Geht es noch besser?

Ja: **DSPACE**($\lg n$). Wir benutzen zwei Bänder als Binär-Zähler. Die Überprüfung der Eingabe auf das Auftreten von genau einem c benötigt keinen Speicher (kann mit einigen Zuständen getan werden). Als zweites prüfen wir Zeichen für Zeichen auf der linken und auf der rechten Seite: dazu müssen die zu prüfenden Positionen gespeichert werden (sie werden auf den beiden Bändern kodiert).

Man kommt auch mit einem einzigen Zähler aus (und einem Band).

Wichtige Fragen

Wichtige Fragen (1)

Zeit: Wird jede Sprache aus **DTIME**(f(n)) von einer DTM entschieden?

Speicher: Wird jede Sprache aus **DSPACE**(f(n)) von einer DTM entschieden?

Wichtige Fragen

Wichtige Fragen (1)

Zeit: Wird jede Sprache aus **DTIME**(f(n)) von einer DTM entschieden?

Speicher: Wird jede Sprache aus **DSPACE**(f(n)) von einer DTM entschieden?

Die Funktionen f sind immer sehr einfache Funktionen, insbesondere sind sie alle berechenbar. In dieser Vorlesung betrachten wir nur Potenzen $f(n) = n^i$.

Wichtige Fragen

Wichtige Fragen (2)

Zeit/Speicher: Wie sieht es mit **NTIME**(f(n)), **NSPACE**(f(n)) aus?

Zeit vs. Speicher: Welche Beziehungen gelten zwischen
DTIME(f(n)), **DSPACE**(f(n)), **NTIME**(f(n)), **NSPACE**(f(n))?

Wichtige Fragen

Halten, hängen nach n Schritten.

Zeitbeschränkt: Was bedeutet es, dass **eine DTM höchstens n Schritte macht**?

Halten?: Streng genommen müßte sie dann entweder **halten** oder **hängen**. Das bedeutet, im ersten Fall, dass die Eingabe **akzeptiert** wird.

Hängen?: DTM's auf beidseitig unendlichem Band können aber nicht hängen.

Wichtige Fragen

Stoppen nach n Schritten.

Stoppen: Wir wollen unter **eine DTM macht höchstens n Schritte** folgendes verstehen:

- Sie **hält** (und **akzeptiert** die Eingabe) innerhalb von n Schritten.
- Sie **hängt** (und **akzeptiert** die Eingabe **nicht**) innerhalb von n Schritten.
- Sie **stoppt** innerhalb von n Schritten **ohne in den Haltezustand überzugehen**. Auch hier wird die Eingabe nicht akzeptiert.

Antworten

Antworten (informell)

Zeit: Jede Sprache aus **DTIME**(f(n)) ist entscheidbar: Man warte einfach solange wie $f(n)$ angibt. Falls bis dahin kein "Ja" gekommen ist, ist die Antwort eben "Nein".

Speicher: Jede Sprache aus **DSPACE**(f(n)) ist entscheidbar. Denn es gibt nur **endlich viele verschiedene Konfigurationen**. Falls also die DTM nicht terminiert (was passiert), macht man folgendes: man schreibt alle Konfigurationen auf (die komplette Rechnung). Falls die DTM nicht terminiert, muss sie in eine Schleife laufen (eine Konfiguration erreichen, die sie schon einmal hatte). Das lässt sich feststellen.

Antworten

Antworten (informell)

NTM vs. DTM: Trivial sind

- $\mathbf{DTIME}(f(n)) \subseteq \mathbf{NTIME}(f(n))$ und
- $\mathbf{DSPACE}(f(n)) \subseteq \mathbf{NSPACE}(f(n))$.

Versucht man aber eine NTM durch eine DTM zu simulieren, braucht man (vermutlich) exponentiell mehr Zeit.

Für die Speicherkomplexität kann man zeigen:

- $\mathbf{NSPACE}(f(n)) \subseteq \mathbf{DSPACE}(f^2(n))$.

Antworten

Antworten (informell)

Zeit vs. Speicher: Trivial sind

- $\mathbf{DTIME}(f(n)) \subseteq \mathbf{DSPACE}(f(n))$ und
- $\mathbf{NTIME}(f(n)) \subseteq \mathbf{NSPACE}(f(n))$.

Aber $\mathbf{DSPACE}(f(n)), \mathbf{NSPACE}(f(n))$ sind viel größer.

Bandkompression/Zeitbeschleunigung

Konstante Faktoren werden ignoriert

Nur **die funktionale Wachstumsrate einer Funktion in Komplexitätsklassen zählt**: Konstante Faktoren werden ignoriert.

Bandkompression

Theorem [Bandkompression]
Für jedes $c \in \mathbb{R}^+$ und jede Speicherfunktion $S(n)$ gilt:

$$\textbf{DSPACE}(S(n)) = \textbf{DSPACE}(cS(n))$$
$$\textbf{NSPACE}(S(n)) = \textbf{NSPACE}(cS(n))$$

Beweis Eine Richtung ist trivial. Die andere geht, indem man den Inhalt einer festen Anzahl r ($> \frac{2}{c}$) von benachbarten Zellen auf dem Band als **ein neues Symbol** darstellt. **Die Zustände der neuen Maschine simulieren die alten Kopfbewegungen als Zustandsübergänge** (innerhalb des neuen Symbols). D.h. für r Zellen der alten Maschine werden nun nur maximal 2 benutzt: im ungünstigsten Fall, wenn man von einem Block in den benachbarten geht.

Zeitbeschleunigung

Theorem [Zeitbeschleunigung]

Für jedes $c \in \mathbb{R}^+$ und jede Zeitfunktion $T(n)$ mit $\lim_{n\to\infty} \frac{T(n)}{n} = \infty$ gilt:

$$\mathbf{DTIME}(T(n)) = \mathbf{DTIME}(cT(n))$$

$$\mathbf{NTIME}(T(n)) = \mathbf{NTIME}(cT(n))$$

Beweis Eine Richtung ist trivial. Der Beweis der anderen läuft wieder über die Repräsentation einer festen Anzahl r ($> \frac{4}{c}$) benachbarter Bandzellen durch **neue Symbole**. Im Zustand der neuen Maschine wird wieder kodiert, welches Zeichen und welche Kopfposition (der simulierten, alten Maschine) aktuell ist.

Wenn die alte Maschine simuliert wird, muss die neue nur 4 statt r Schritte machen: 2 um die neuen Felder zu drucken und weitere 2 um den Kopf auf das neue und wieder zurück auf das alte (im schlimmsten Fall) zu bewegen.

Wachstumsrate von DTIME und DSPACE

Wachstumsrate von DTIME, DSPACE

Es sei $T(n)$ eine Zeitfunktion mit $\lim_{n\to\infty} \frac{T(n)}{n} = \infty$, und es sei $S(n)$ eine Speicherfunktion.

(a) Es sei $f(n) = \mathbf{O}(T(n))$. Dann gilt: $\mathbf{DTIME}(f(n)) \subseteq \mathbf{DTIME}(T(n))$.

(b) Es sei $g(n) = \mathbf{O}(S(n))$. Dann gilt: $\mathbf{DSPACE}(g(n)) \subseteq \mathbf{DSPACE}(S(n))$.

Wachstumsrate von DTIME und DSPACE

Definition [P, NP, PSPACE]

$$\begin{aligned} \mathbf{P} &:= \bigcup_{i \geqslant 1} \mathbf{DTIME}(n^i) \\ \mathbf{NP} &:= \bigcup_{i \geqslant 1} \mathbf{NTIME}(n^i) \\ \mathbf{PSPACE} &:= \bigcup_{i \geqslant 1} \mathbf{DSPACE}(n^i) \end{aligned}$$

Wachstumsrate von DTIME und DSPACE

Definition [P, NP, PSPACE]

$$\begin{aligned} \mathbf{P} &:= \bigcup_{i \geqslant 1} \mathbf{DTIME}(n^i) \\ \mathbf{NP} &:= \bigcup_{i \geqslant 1} \mathbf{NTIME}(n^i) \\ \mathbf{PSPACE} &:= \bigcup_{i \geqslant 1} \mathbf{DSPACE}(n^i) \end{aligned}$$

Intuitiv

- Probleme in **P** sind effizient lösbar, jene aus **NP** können in exponentieller Zeit gelöst werden.
- **PSPACE** ist eine sehr große Klasse, weit größer als **P** oder **NP** .