Rekursions- und Lerntheorie

WiSe 2010/11; Univ. Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

$\boxed{\textbf{Rekursions- und Lerntheorie}}$ Gesamtübersicht

1. Einführung: Grundsätzliche Betrachtungen

2. Berechenbarkeitstheorie: Die Churchsche These (4 VL)

3. Ausblicke auf weitere Ergebnisse der Rekursionstheorie

4. Lerntheorie: Modelle und Aussagen

Exkurs: Was ist eine Zufallszahl?

Mögliche erste Antwort $\rightsquigarrow$ Wahrscheinlichkeitstheorie

anderer Ansatz: *Kolmogorov-Komplexität*

$$K(x) = \mu e(h_e(0) = x)$$

Intuition: $K(x)$ ist die kürzeste Beschreibung von x.
x heißt *zufällig*, falls $x \leq K(x)$.
Dann ist nämlich x selbst die kürzeste Beschreibung von x.

Bemerkungen: Zusammenhänge mit Lerntheorie und Datenkompression:
Nur Regelmäßigkeiten aufweisende Daten (Zahlen) kann man komprimieren.
Man kann nämlich eine komprimierte Datei als Programmtext auffassen
(der dann vom Dekomprimierer interpretiert wird).
Auch Lernalgorithmen schöpfen ihre “Intelligenz” aus dem Erkennen von Regelmäßigkeiten.

Es seien $\mathcal{Z} := \{x \mid x \leq K(x)\}$ die *Zufallszahlen*.

Wie viele Zufallszahlen gibt es?

Satz 1. Es gibt unendlich viele Zufallszahlen.

Beweis: Betrachte die Folge der definierten Werte unter

$$h_0(0), h_1(0), \dots h_n(0).$$

Ist x in dieser Folge nicht enthalten, so $n + 1 \leq K(x)$ nach Def. der Kolmogorov-Komplexität.

Sei $x_n = \min\{x \mid x \notin \{h_0(0), h_1(0), \dots h_n(0)\}\}$ die kleinste dieser x-Zahlen.

Schubfachprinzip $\rightsquigarrow x_n \leq n + 1 \rightsquigarrow x_n \in \mathcal{Z}$.

$\forall n \exists x_n \in \mathcal{Z} : K(x_n) > n \rightsquigarrow$ Beh.

Hinweis: In gewissem Sinne sind Nichtzufallszahlen sogar seltener als Zufallszahlen.

Datenkompression ist so gesehen “fast unmöglich”.

Dass sie in der Praxis doch funktioniert, zeigt:

Die Welt ist nicht beliebig zufällig, sondern verhältnismäßig geordnet.

Eine Eigenschaft der Nichtzufallszahlen $\mathcal{NZ} := \{x \mid x > K(x)\} = \mathbb{N} \setminus \mathcal{Z}$

Satz 2: $\mathcal{NZ}$ ist rek. aufzählbar.

Beweis: Sei T_∞ eine TM, die niemals hält, also $\phi(x) =\uparrow$ für alle x implementiert.

Eine TM T auf x muss nur mit utm für alle $e < x$ schauen, ob $h_e(0) = x$.

Dazu simuliert T den s-ten Schritt von allen T_e (parallel) auf 0 für $s = 0, 1, \ldots$.

Hält die Simulation von T_e im Schritt s_e, so überprüft T, ob die Ausgabe x ist.

Wenn ja, hält die gesamte Simulation.

Wenn nein, wird T_∞ statt T_e simuliert, sobald Schritte $s > s_e$ für T_e simuliert werden sollten.

Somit hält T dann und nur dann, wenn irgendein T_e auf 0 mit Ausgabe x hält.

Nichtzufallszahlen sind simpel: Vorbereitungen

Erinnerung: Eine rekursiv aufzählbare Menge A heiße *simpel*, wenn $\mathbb{N}\setminus A$ unendlich ist und jede unendliche rek. aufzählbare Menge schneidet. Simple Mengen sind nicht m-vollständig.

utm-Eigenschaft $\leadsto$ $\phi : \langle e, n\rangle \mapsto$ "n-tes Element in Standardaufzählung von W_e" ist berechenbar.
Hierbei sei *Standardaufzählung* Folgendes:
Simuliere, für $s = 0, 1, 2, \ldots$ und $j = 0, \ldots s - j$, den $(s - j)$-ten Schritt von T_e, angesetzt auf j.
So werden nacheinander alle Elemente von W_e gefunden, was die Aufzählung festlegt.
smn-Eigenschaft $\leadsto$ es gibt rek. Fkt. r, sodass $h_{r(e,n)}(0) = \phi(\langle e, n\rangle)$.

Satz 3. $\mathcal{NZ}$ ist simpel.

Beweis: Zu zeigen bleibt wegen Satz 1 und 2:
$\mathcal{NZ}$ schneidet jede unendliche rek. aufzählbare Menge W_e.
Betrachte ϕ und r wie soeben.
Wäre $r(e,n) < \phi(\langle e,n\rangle)$, so $\phi(\langle e,n\rangle) \in \mathcal{NZ}$.
Betrachte rek. Fkt. $t(n) = \max\{r(e,n) \mid e \leq n\}$.
Bei Eingabe von n, beobachte nun Standardaufzählung von W_e solange, bis ein Element $y_{e,n}$ größer als $t(n)$ aufgezählt wird.
Def. $W_{g(e)} = \{y_{e,n} \mid n \in \mathbb{N}\}$.
Konstruktion leicht modifizierbar, sodass sich Standardaufzählung ergibt.
Die vorigen Zeilen zeigen: g ist rekursiv.
Mit W_e ist auch seine Teilmenge $W_{g(e)}$ unendlich.
Für $n > g(e)$ gilt: $r(g(e),n) \leq t(n) < h_{r(g(e),n)}(0) = \phi(\langle g(e),n\rangle) = y_{e,n}$.
Also ist $\phi(\langle g(e),n\rangle) \in W_e \cap \mathcal{NZ}$.

Berrys Paradoxon

Zu jeder Zahl n betrachte folgende Zahl:

“Die kleinste Zahl, die sich nicht mit n Zeichen beschreiben lässt.”

Zum Aufschreiben von n in diesem Satz benötigen wir $\log(n)$ Zeichen (z.B. Binärcodierung).

Also gilt für fast alle n: Der obige Satz benötigt $c + \log(n) < n$ Zeichen, ist also kürzer als die Zahl n selbst angibt.

Der letzte Beweis verwendet diese Art von Argumentation!

Orakel-Turingmaschinen

Ein *Orakel* O ist im Folgenden eine bel. Menge natürlicher Zahlen oder Wörter.

Normale TM werden erweitert um ein *Orakelband*:
Auf dieses darf nur geschrieben werden.
Durch Übergang in einen speziellen *Orakelzustand* z_{O} wird gefragt, ob der Bandinhalt in O liegt.
Entsprechend wird in dem (speziellen) Zustand z_{O}^0 bzw. z_{O}^1 weitergemacht.
Der Bandinhalt des Orakelbandes ist jetzt wieder gänzlich gelöscht.

Akzeptanz von Wörtern bzw. Berechnung von Funktionen wird nun wie bisher definiert.
$\rightsquigarrow$ *rekursiv (aufzählbar) bzgl. / relativ zu Orakel* O;
partiell rekursiv bzgl. / relativ zu Orakel O

Relativierte Berechnungen

Man kann jetzt “Berechenbarkeit bzgl. O” studieren.
Alle bislang gemachten Überlegungen gelten auch bzgl. jedes Orakels, oder kurz *relativiert*.

So lautet beispielsweise der relativierte Satz von Post:

Satz (Post): Es sei O eine Menge. Eine Menge A ist rekursiv bzgl. O gdw. A ist rek. aufzählbar bzgl. O und $\mathbb{N} \setminus A$ ist rek. aufzählbar bzgl. O.

Man vollziehe den Beweis hierzu nach!

Turing-Reduktion

Eine Menge A heißt *T-reduzierbar* auf die Menge B gdw. A ist rekursiv bzgl. B.
Beispiel: m-Reduktionen sind T-Reduktionen mit nur einem Orakelaufruf.

Schreibweise: $A \leq_T B$: A ist T-reduzierbar auf B.

Schreibweise: $A \equiv_T B$ gdw. $A \leq_T B$ und $B \leq_T A$.

$\leq_T$ ist eine *Quasiordnung* auf $\mathbb{N}$ (bzw. E^*), aufgefasst als Gödelnummern, also eine reflexive und transitive Relation.

Das Mengensystem $\{B \mid B \equiv_T A\}$ heißt auch *T-Grad* (von A).

Turing-Vollständigkeit

Satz 4. Ist A rek. aufzählbar, so $A \leq_T \mathcal{K}$.

<u>Beweis:</u> Dies folgt aus der schon gezeigten entsprechenden Aussage für m-Reduktionen.

Eine rek. aufzählbare Menge A heißt *Turing-vollständig* gdw. $\mathcal{K} \leq_T A$.

Mitteilung: Ein rek. aufzählbare Menge A ist T-vollständig gdw. es gibt eine bzgl. A partiell rek. Fkt. f ohne Fixpunkte, d.h.: $\forall x : W_x \neq W_{f(x)}$.

Turing-Grade

Ein *rek. aufzählbarer Turing-Grad, kurz im Folgenden: T-Grad*, ist eine Äquivalenzklasse Turing-äquivalenter Mengen (also Zahlen), die wenigstens eine rek. aufzählbare Menge enthält.

Beispiele:
$\mathbf{0}$: Der T-Grad der rekursiven Mengen
$\mathbf{0}'$: Der T-Grad von $\mathcal{K}$.

Posts ursprüngliche Frage: Gibt es weitere (rek. aufzählbare) T-Grade?
Dekker konnte zeigen:
Ein (rek. aufz.) T-Grad $\neq \mathbf{0}$ enthält eine simple Menge.
Simple Mengen lösen im Allgemeinen also Posts Problem nicht.

Leisten die uns bekannten simplen Mengen es vielleicht doch das von Post Gewünschte?

Eine rek. aufzählbare Menge A heißt *effektiv simpel* wenn ihr Komplement unendlich ist und es außerdem eine rekursive Fkt. s gibt mit:

$$W_e \subseteq \mathbb{N} \setminus A \implies |W_e| \leq s(e).$$

1. Posts simple Menge ist effektiv simpel.
   Aufgabe!

2. $\mathcal{NZ}$ ist effektiv simpel.
   Nämlich: Enthält W_e wenigstens $t(g(e))$ viele Elemente (z.B., weil W_e unendlich ist), so enthält $W_{g(e)}$ wenigstens $g(e)$ viele Elemente (per def. von g im Beweis) und daher eine Nichtzufallszahl. Enthält W_e nur Zufallszahlen, so folgt mit $s(e) = t(g(e))$: $|W_e| \leq s(e)$.

Der Satz von Martin

Satz 5. Jede effektiv simple Menge ist T-vollständig.

Beweis: Sei A effektiv simpel. Wir haben:

$$W_e \subseteq \mathbb{N} \setminus A \implies |W_e| \leq s(e).$$

Definierte f wie folgt:

$W_{f(e)} = \{$die ersten $s(e) + 1$ Elemente nicht in $A\}$.

Per def. ist f berechenbar bzgl. A.

Per def. hat f keinen Fixpunkt: Wäre $W_e = W_{f(e)}$, so $W_e \subseteq \mathbb{N} \setminus A$, aber $|W_e| = s(e) + 1$.

Nach obiger Mitteilung ist A T-vollständig.

Einordnung und Ausblick

Wir werden hier den Weg zur Lösung des Postschen Problems abbrechen.
Dieser würde uns über hypersimple, hyperhypersimple usf. Mengen führen.
Allein schon die Bezeichnungen deuten an, dass dieser Weg recht steinig ist.
Mitteilung: Es gibt weitere rek. aufzählbare Turing-Grade.

Orakelberechnungen werden uns auch noch im Weiteren beschäftigen.
Ganz analoge Überlegungen finden sich auch in der Komplexitätstheorie.