Formale Grundlagen der Informatik

9 Entscheidbarkeit von Problemen

Unentscheidbare Probleme
Halteproblem • Reduktionen

Recap: Hierarchiebeziehungen

- $\mathcal{L}(\mathrm{REC}) = \{\, L \mid L$ ist rekursiv $\}$
- $\mathcal{L}(\mathrm{RE}) = \{\, L \mid L$ ist rekursiv aufzählbar $\}$

  $= \{\, L \mid L = G(M)$ für eine TM $M \,\}$

  $= \{\, L \mid L = L(M)$ für eine TM $M \,\}$

Folgerung 8.7: $\mathcal{L}(\mathrm{REG}) \subset \mathcal{L}(\mathrm{REC}) \subseteq \mathcal{L}(\mathrm{RE})$

Teilmenge nach Lemma 8.7;
echt wegen $\{\, a^n b^n \mid n \geq 0 \,\}$

Teilmenge nach Lemma 8.2;
Echtheit heute

Menge *versus* Entscheidbarkeitsproblem

- *Recap:* Eine Menge $L \subseteq \Sigma^*$ ist rekursiv wenn ihre charakteristische Funktion

$$\varphi_L(x) = \begin{cases} 1 & \text{falls } x \in L \\ 0 & \text{falls } x \notin L \end{cases}$$

  für alle $x \in \Sigma^*$ Turing-berechenbar ist.

- Formulierung als **Entscheidbarkeitsproblem**:

  Eingabe: Wort $w \in \Sigma^*$
  Frage: Gilt $w \in L$?

- Umgekehrt definiert jedes Entscheidbarkeitsproblem die Menge aller Eingaben mit Antwort „Ja".

Menge *versus* Entscheidbarkeitsproblem

Beispiel:

Primzahlproblem:

Eingabe: natürliche Zahl n
Frage: Ist n eine Primzahl?

$\Updownarrow$

$\{ n \in \mathbb{N} \mid n \text{ ist Primzahl} \}$

Entscheidbarkeitsprobleme

- *intuitiv:* **Ja/Nein-Frage**, ob Elemente eines Grundbereichs eine gewisse Eigenschaft (Prädikat) haben oder nicht
- **Eingabe:** kodiert als ein Wort w über einem geeigneten Alphabet Σ
  (*z.B.* ein Paar ganzer Zahlen x und y durch bin(x)#bin(y) über {0,1,#})

  Ausgabe: **1** falls w eine gewisse Eigenschaft hat
  0 sonst
  (*z.B.* 1, falls x und y teilerfremd, sonst 0)

- Darstellung durch **Eingabe, Frage** (Ja/Nein-Frage)

Entscheidbarkeit von Problemen

- Sei L die Menge der Wörter über Σ, die die Eigenschaft haben (für die die Ausgabe 1 ist).
  Das **Problem** ist **(algorithmisch) entscheidbar**,
  falls die Menge L entscheidbar (d.h. rekursiv) ist.
  - Also falls die charakteristische Funktion von L Turing-berechenbar ist:

$$\varphi_L(x) = \begin{cases} 1 & \text{falls } x \in L \\ 0 & \text{falls } x \notin L \end{cases}$$

  - Es gibt also eine TM, die *für alle* Eingabewörter über Σ entweder 0 oder 1 ausgibt (f_M ist **totale** Funktion $\longrightarrow$ TM hält bei allen Eingabewörtern!!!)

Wortproblem

- Sei $L \subseteq \Sigma^*$ eine Sprache. Das **Wortproblem** für L lautet:
  Eingabe: Wort $w \in \Sigma^*$
  Frage: Gilt $w \in L$?

➢ Das Wortproblem für eine Sprache L ist genau dann entscheidbar, wenn L *rekursiv* ist.

- Die Entscheidbarkeit des Wortproblems hängt davon ab, wie L spezifiziert werden wird (DEA / NEA / RA / DTM / NTM / …).

Universelles Wortproblem

- **Eingabe:** ein Mechanismus M, der eine Sprache $L(M) \in \Sigma^*$ definiert;
  ein Wort $w \in \Sigma^*$

  Frage: Gilt $w \in L(M)$?

➢ Eingaben sind also Paare (M, w), wobei M geeignet als Wort kodiert ist

- Die Tupel, die einen Automaten (DEA, NEA, DTM, …) spezifizieren, können als Wörter geschrieben werden, wenn die Tripel $\delta(q_i, a_j) = \delta_{ij}$ hintereinander aufgeschrieben werden.
- *Beispiel:* Bei $Q = \{q_1, q_2, q_3\}$ und $\Sigma = \{a_1, a_2\}$: $\delta_{11}\delta_{12}\delta_{21}\delta_{22}\delta_{31}\delta_{32}$
- *Binäre Kodierungen von Automaten existieren …*

Binär kodierte DTM

- $M = (\{q_1, q_2, \dots, q_k\}, \{a_1, a_2, \dots, a_\ell\}, \{a_1, a_2, \dots, a_\ell, a_{\ell+1}, \dots, a_m\}, \delta, q_1, a_m, \{q_j, q_{j+1}, \dots, q_k\})$
- Kodierungen der Symbole (*intuitiv, aber redundant*):

$q_i \to 0^i1$ für $1 \leq i \leq k,$ (Zustände)

$a_i \to 0^i1^2$ für $1 \leq i \leq m,$ (Bandsymbole)

$\mathrm{R} \to 01^3$ $\quad$ $\mathrm{L} \to 0^21^3$

$(\to 01^4$ $\quad$ $) \to 0^21^4$ $\quad$ $\{ \to 01^5$ $\quad$ $\} \to 0^21^5$

- *Beispiel:* $M = (\{q_1, q_2\}, \{0\}, \{0,*\}, (q_1, 0, R)(q_1,*, R)(q_2, 0, R)(q_2,*, R), q_1,*, \{q_2\})$

$\to$ *Kodierung* $\langle M \rangle = 01^4\, 01^5010^210^21^5 01^501^20^21^5 01^501^20^21^20^21^5$

$01^40101^201^30^21^4\ 01^4010^21^201^30^21^4\ 01^40^2101^201^30^21^4$

$01^40^210^21^201^30^21^4 010^21^2 01^50^210^21^5 0^21^4$

Halteproblem für DTM

- Das universelle Wortproblem für rekursiv aufzählbare Sprachen:
  Eingabe: DTM M, $w \in \Sigma^*$, geschrieben als $\langle M \rangle \langle w \rangle$
  Frage: Gilt $w \in L(M)$?
- … ist äquivalent zu dem Problem
  Eingabe: DTM M, $w \in \Sigma^*$, geschrieben als $\langle M \rangle \langle w \rangle$
  Frage: Hält M bei Eingabe w? ⟶ **Halteproblem für TM**
- Entscheidbarkeit des Halteproblems für DTM ist also die Frage, ob die **universellen Sprache**

$$\boldsymbol{L_u} = \{\langle M \rangle \langle w \rangle \mid w \in L(M)\}$$

  rekursiv ist.

Eigenschaften der universellen Sprache

Satz 9.1. Die universelle Sprache L_u ist rekursiv aufzählbar.

Beweisskizze: Konstruieren eine TM M_u mit $L(M_u) = L_u$ wie folgt:

1. Prüfe, ob das Eingabewort von der Form $\langle M \rangle \langle w \rangle$ ist.
2. Simuliere die Arbeit von M auf der Eingabe w.
3. Halte, wenn M auf der Eingabe w hält.

Dann gilt $w \in L(M)$ gdw. $\langle M \rangle \langle w \rangle \in L(M_u)$.

Da $\langle M \rangle \langle w \rangle \in L_u$ gdw. $w \in L(M)$, akzeptiert M_u die Menge L_u. $\square$

Eine nicht rekursiv aufzählbare Sprache

Satz 9.2. Es gibt eine Sprache, die nicht rekursiv aufzählbar ist.

Beweis: Konstruieren eine Sprache L_d wie folgt.

unendliche Tabelle:

$j \longrightarrow$

$i \downarrow$	1	2	3	4	…
1					
2					
3					
4					
⋮					

i: Index des Wortes w_i in kanonischer Anordnung aller Wörter über $\{0,1\}$

j: falls Binärdarstellung von j (nach Ergänzung einer führenden 0) Code einer TM ist, nennen wir diese TM M_j

Eine nicht rekursiv aufzählbare Sprache

Satz 9.2. Es gibt eine Sprache, die nicht rekursiv aufzählbar ist.

Beweis: Konstruieren eine Sprache L_d wie folgt.

unendliche Tabelle:

$j \longrightarrow$

$i \downarrow$	1	2	3	4	…
1	0	1	1	0	…
2	1	1	0	0	…
3	0	0	1	0	…
4	0	1	0	0	…
⋮	⋮	⋮	⋮	⋮	⋱

Tabelleneintrag an Zelle (i, j) ist genau dann eine **1**, wenn j Code einer TM und diese TM M_j das Wort w_i akzeptiert;

sonst **0**

Einträge in der Diagonalen sind 1,
wenn $w_i \in L(M_i)$,
sonst 0.

Eine nicht rekursiv aufzählbare Sprache

- Sei $\boldsymbol{L_d}$ die Sprache aller w_i so dass in der Diagonalen in der i-ten Zeile (also an der Tabellenposition (i, i)) eine 0 steht.
- *Annahme*: L_d wird von einer TM akzeptiert.
- Es gelte $L_d = L(M_k)$ für ein $k \geq 1$.
- Falls $w_k \in L_d$, dann ist der Eintrag in der Tabelle an Position (k, k) eine 0. Dann gilt also $w_k \notin L(M_k)$, somit $w_k \notin L_d$. **Widerspruch!**
- Falls $w_k \notin L_d$, dann ist an (k, k) eine 1, also $w_k \in L_d$. **Widerspruch!**
- Folglich gibt es kein solches k. Also ist L_d nicht rekursiv aufzählbar. $\blacksquare$

Eigenschaften der universellen Sprache

Satz 9.3. Die universelle Sprache L_u ist nicht rekursiv.

Beweis: *Angenommen*, die TM E entscheidet $L_u = \{\langle M\rangle\langle w\rangle \mid w \in L(M)\}$.

Konstruieren folgenden Algorithmus, der auf Eingabe w aus $\{0,1\}^*$ so arbeitet:

1. Bestimme die Zahl i, sodass $w = w_i$ in der kanonischen Anordnung.
2. Schreibe die Binärzahl i (potenziell den Code der TM M_i) vor $w = w_i$.
3. Entscheide mit E, ob $0iw \in L_u$ gilt oder nicht.
   - Antwort ist „Ja", falls i eine TM kodiert und $w_i \in L(M_i)$, sonst „Nein".
   - Antwort ist „Ja" gdw. in der Entscheidungstabelle an Position (i, i) eine 1 steht.
   - Antwort ist „Ja" gdw. $w_i \in \overline{L_d}$.

Eigenschaften der universellen Sprache

Angenommen, die TM E entscheidet $L_u = \{\langle M\rangle\langle w\rangle \mid w \in L(M)\}$.

Dann ist $\overline{L_d}$ also rekursiv. **Widerspruch!**

Denn:

- L_d ist nicht rekursiv aufzählbar (Beweis von Satz 9.2).
- L_d ist nicht rekursiv (Lemma 8.2).
- Wenn $\overline{L_d}$ rekursiv wäre, dann auch L_d (Lemma 8.5).
  Also kann $\overline{L_d}$ nicht rekursiv sein. ∎

Folgerung 9.4. Das Halteproblem für DTMs ist unentscheidbar.

Recap: Hierarchiebeziehungen

- $\mathcal{L}$(**REC**) $= \{\, L \mid L$ ist rekursiv $\}$
- $\mathcal{L}$(**RE**) $= \{\, L \mid L$ ist rekursiv aufzählbar $\}$

Folgerung 9.5: $\mathcal{L}(\mathbf{REG}) \subset \mathcal{L}(\mathbf{REC}) \subset \mathcal{L}(\mathbf{RE})$

$\boxed{L_u}$ $\boxed{L_d}$

➢ L_u ist nicht entscheidbar, aber **semi-entscheidbar**:
<u>nur korrekte</u> Eingaben werden identifiziert!

➢ L_d ist nicht semi-entscheidbar.

Reduktionen

- Unentscheidbarkeit des Halteproblems im wesentlichen gezeigt durch Diagonalisierung für L_d.
- Weitere Unentscheidbarkeiten werden meist durch Reduktion von einem bekannten unentscheidbaren Problem (z.B. Halteproblem) gezeigt.

➢ Zusammenhang zwischen den Problemen so herstellen, dass gilt:

Wenn das Problem entscheidbar ist, dann ist auch das Halteproblem für TM entscheidbar (Widerspruch!).

Äquivalenzproblem für DTM

Eingabe: zwei DTM M_1 und M_2

Frage: Gilt $L(M_1) = L(M_2)$?

Satz 9.6. Das Äquivalenzproblem für DTM ist unentscheidbar.

Vorbereitung der Reduktion vom Halteproblem für DTM:

	Halteproblem	Äquivalenzproblem
Eingabe	$\langle M\rangle\langle w\rangle$	$\langle M_1\rangle\langle M_2\rangle$
Ausgabe ist „Ja“, falls …	$w \in L(M)$	$L(M_1) = L(M_2)$
Ausgabe ist „Nein“, falls …	$w \notin L(M)$	$L(M_1) \neq L(M_2)$

Äquivalenzproblem für DTM

Satz 9.6. Das Äquivalenzproblem für DTM ist unentscheidbar.

Beweis:

- Konstruieren M_1 so,
  - dass alle Wörter außer w sofort abgelehnt werden und
  - bei Eingabe w die Arbeit von M auf w simuliert wird.

$$L(M_1) = \begin{cases} \{w\} & \text{falls } w \in L(M) \\ \emptyset & \text{falls } w \notin L(M) \end{cases}$$

- Konstruieren M_2 so, dass $L(M_2) = \{w\}$.

➢ $L(M_1) = L(M_2)$ gdw. $w \in L(M)$.

➢ Wenn das Äquivalenzproblem entscheidbar ist, dann auch das Halteproblem.

➢ Das Äquivalenzproblem für DTM ist unentscheidbar. ■

Relevanz des Äquivalenzproblems

- *innerhalb der Theorie*:

  z.B. Check, ob Konvertierungen (z.B.
  - von RA in RA oder
  - von DEA in RA oder RA in DEA oder
  - von NTM in DTM …)

  korrekt sind

- *in Anwendungen*:

  z.B. Check, ob eine Softwaremigration geglückt ist …
  - kann für reguläre Sprachen automatisiert werden
  - muss für rekursiv aufzählbare Sprachen in jedem Einzelfall gesondert nachgewiesen werden

Einige unentscheidbare Probleme

- **Leerheit** für DTM:
  - Eingabe: DTM M
  - Frage: Gilt $L(M) = \emptyset$?
- **Regularität** für DTM:
  - Eingabe: DTM M
  - Frage: Gilt $L(M) \in \mathcal{L}(\mathrm{REG})$?
- **Rekursivität** für DTM:
  - Eingabe: DTM M
  - Frage: Gilt $L(M) \in \mathcal{L}(\mathrm{REC})$?

- **Erfüllbarkeit** in der Prädikatenlogik:
  - Eingabe: prädikatenlogischer Ausdruck $\boxed{\forall x . r(x, f(y))}$
  - Frage: Gibt es eine Interpretation der Symbole des Ausdrucks und eine Belegung der freien Variablen, so dass der Ausdruck wahr ist?
- **Hilberts 10. Problem**
  - Eingabe: Polynom in n Variablen mit ganzzahligen Koeffizienten
  - Frage: Gibt es eine ganzzahlige Lösung?

Hilberts 10. Problem (*genauer*)

Eingabe: Ganze Zahl $n \geq 1$, Polynom $p(x_1, x_2, \dots, x_n) = \sum c_{i_1 i_2 \dots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$
mit ganzzahligen Koeffizienten $c_{i_1 i_2 \dots i_n}$

Frage: Gibt es eine Lösung von $p(x_1, x_2, \dots, x_n) = 0$ in $\mathbb{Z}^n$?

Beispiele:

- $p(x, y, z) = 3xyz^2 + 5xy^2 - 4x^2yz = 0$ hat die Lösung (2,1,1)
- $p(x, y, z) = 2x^4y^2 + 3x^2z^2 + 2y^2z^6 - 1 = 0$

  hat keine Lösung, da die ersten drei Summanden 0 oder ≥ 2 sind.

Ein entscheidbarer Spezialfall

- Hilberts 10. Problem ist **entscheidbar** für Polynome in einer Variablen:
$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_k x^k$$

- $p(x) = 0$ gdw. $-c_0 = c_1 x + c_2 x^2 + \cdots + c_k x^k$

➢ Jede Lösung für x muss Teiler von c_0 sein.

➢ Es gibt nur endlich viele Teiler von c_0.

➢ Der Entscheidungsalgorithmus bestimmt diese Teiler und setzt sie in die Gleichung ein. Wird so eine Nullstelle von $p(x)$ gefunden, dann ist die Antwort „Ja“, sonst „Nein“.