Rekursions- und Lerntheorie

WiSe 2010/11; Univ. Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

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

Motivation

Lernen ist eine der offensichtlichen Ausprägungen von Intelligenz.

(Mathematische) Modelle für “Lernen” können hilfreich bei der Erklärung von Lernverhalten bei Menschen und auch anderen Lebewesen sein.

$\rightsquigarrow$ Zusammenhänge mit Neurologie, Psychologie, Pädagogik usw.

Können wir “intelligente Maschinen” bauen?

Lerntheorie will helfen, Möglichkeiten und Grenzen dieses Vorhabens im Zusammenhang mit “Lernen” auszuloten.

Denkbare Anwendungen:

— Entwicklung autonomer Systeme

— Unterstützung beim Erkennen von kausalen Zusammenhängen, insbesondere im Bereich der Naturwissenschaften.

Kindliches Lernen, z.B. Spracherwerb

Lehrer (?)

Aspekte des Spracherwerbs

Phonetik: Lautkunde

Lexik(on):

Wortschatz / Wortbildung

Grammatik

(Pragmatik)

Informatik ?

Mathematik ?

Technikbezug ?!

Kindlicher Spracherwerb: manchmal aktiver

Modellierung?!

$\rightsquigarrow$ Betrachtung unterschiedlicher Lernszenarien

insbesondere:

—unterschiedliche Rollen von Lehrer und Schüler,

—soziale Gefüge (Protokolle)

$\rightsquigarrow$ Vorlesung “Lernalgorithmen”

Autonomes System: Antwort auf sich wandelnde Umgebung

$\leadsto$ Notwendigkeit einer formalisierten Sicht des Lernens

Aufgaben aus dem Bereich der Naturwissenschaften

Ein Forscher mag an dem Zusammenhang zweier physikalischer Größen in einem Experiment interessiert sein.
Seine Messreihe liefert ihm Zahlenpaare, z.B.:
$(2, 4)$, $(0, 0)$, $(3, 9)$, $(1, 1)$, $\ldots$
Die bisherigen Ergebnisse könnten den Forscher zu der Hypothese leiten, er habe es mit einem “quadratischen Zusammenhang” der Form (n, n^2) der beiden Messgrößen zu tun.
Es wäre wünschenswert, solcherlei Hypothesen automatisch gestützt zu erzeugen.

$\boxed{\textbf{“Intelligenztest”}}$ spielerische Annäherung ans Sprachenlernen

Frage: Welche Sprache ist (genau) beschrieben durch:

$$\{ab, abb, abbb, abbbb, abbbbb, \ldots\}?$$

Was ist die Struktur dieser Sprache ?

Begriffserklärungen:

Sprache: Menge von Wörtern

Wort: Aneinanderreihung von Zeichen

mathematischer: Element eines frei erzeugten Monoids

Sprachklasse: Menge von Sprachen

$\boxed{\text{“Intelligenztest”}}$

Welche Sprache ist (genau) beschrieben durch:

$$\{ab, abb, abbb, abbbb, abbbbb, \ldots\}?$$

Antwort: ab^+ (?)

D.h.: Erst kommt ein a, dann beliebig viele b’s, mindestens ein b.

“Intelligenztest”

Welche Sprache ist (genau) beschrieben durch:

$$\{ab, abb, abbb, abbbb, abbbbb, \dots\}?$$

andere Antwort: $\{ab^{10n+k} \mid n \geq 0, k = 1, 2, 3, 4, 5\}$ (?)
Also: Das “nächste” Wort wäre $abbbbbbbbbbb = ab^{11}$.
Welche Lösung ist “richtig” ?

Literatur

Unsere Darstellungen werden im Wesentlichen den ersten Kapiteln von dem folgenden Buch entsprechen:
Jain, Osherson, Royer, Sharma: *Systems That Learn*, MIT Press, 1999.

Was benötigt man, um ein Lernparadigma anzugeben?

(1) Welche Wirklichkeiten sind theoretisch möglich?

(2) Welche Art von Hypothesen dürfen geäußert werden?

(3) Von welcher Natur sind die zur Verfügung stehenden Daten über die Wirklichkeit?

(4) Welche Möglichkeiten besitzt der Forscher?

(5) Wann gilt die Arbeit des Forschers als erfolgreich?

Speziellere Lernparadigmen

(1) Beim Sprachenlernen beispielsweise werden wir uns auf gewisse Sprachfamilien einschränken.
Aufgrund der bekannten Umrechnungsmöglichkeiten zwischen Wörtern und Zahlen können wir abstrakt “Sprachen” auch als “Zahlenmengen” verstehen und somit Sprachfamilien als Klassen von Zahlenmengen.
(2) Naheliegend sind Gödelnummern sowohl zur Angabe von Sprachen (Wertebereiche partieller Fkt.) als auch für funktionale Zusammenhänge.
(3) Möglich wären endliche Teilmengen bzw. Listen (beim Sprachenlernen).
(4) Am einfachsten wären prinzipiell beliebige “Macht” für den Forscher oder aber die Modellierung des Forschers durch eine Turing-Maschine.
(5) Wir werden einen geeigneten “Grenzwertbegriff” benutzen.

Mögliche Welten: Die Menge $\mathcal{E} \subseteq 2^{\mathbb{N}}$ der aufzählbaren “Sprachen”.

Zur Darstellung von *Beispielströmen*:
Ein *Text* sei irgendeine Abbildung $T : \mathbb{N} \mapsto \mathbb{N} \cup \{\#\}$.
#: *Pausenzeichen*
Der *Inhalt* von T: $\mathrm{Inh}(T) = \{T(i) | i \in \mathbb{N}, T(i) \neq \#\}$.
T heißt *Text für* $L \subset \mathbb{N}$, falls $\mathrm{Inh}(T) = L$.

<u>Hinweis:</u> Für jede nicht-leere Sprache gibt es unendlich viele Texte, denn
(1) die Anordnung ist beliebig, (2) Wiederholungen sind zulässig, und
(3) # kann beliebig eingefügt werden.
Ein Text modelliere den Beispielstrom (aus der “gewählten” Sprache $L \in \mathcal{E}$) in den Lerner, der seine *Hypothesen* in Form von Turingmaschinen oder Programmen äußern kann.

Setze $T[n] = (T(0), \ldots, T(n-1))$.
$SEQ = \{T[n] | T \text{ Text}, n \in \mathbb{N}\}$ ist die Menge aller endlichen Folgen über $\mathbb{N} \cup \{\#\}$.

Die *Länge* von $\sigma \in SEQ$ werde mit $|\sigma|$ bezeichnet.
Bem.: T beginnt mit σ genau dann, wenn $T[|\sigma|] = \sigma$.

Bem.: $T[n]$ ist die Datenmenge, die den Lerner bis zum n-ten Zeitpunkt zur Verfügung steht.

Gehen wir von einer fixierten Nummerierung aller Turingmaschinen aus, so können wir einen *Lerner*, auch *Inferenzmaschine (IM)* genannt, als (zunächst beliebige) partielle Funktion $F : SEQ \multimap \mathbb{N}$ verstehen.
Notation: $F(T[n]) \downarrow$: F ist definert auf $T[n]$, das heißt, der Lerner äußert eine Hypothese nach Beobachtung von $(T(0), \ldots, T(n-1))$.
Zugehörige Sprache: $L(F(T[n]))$.

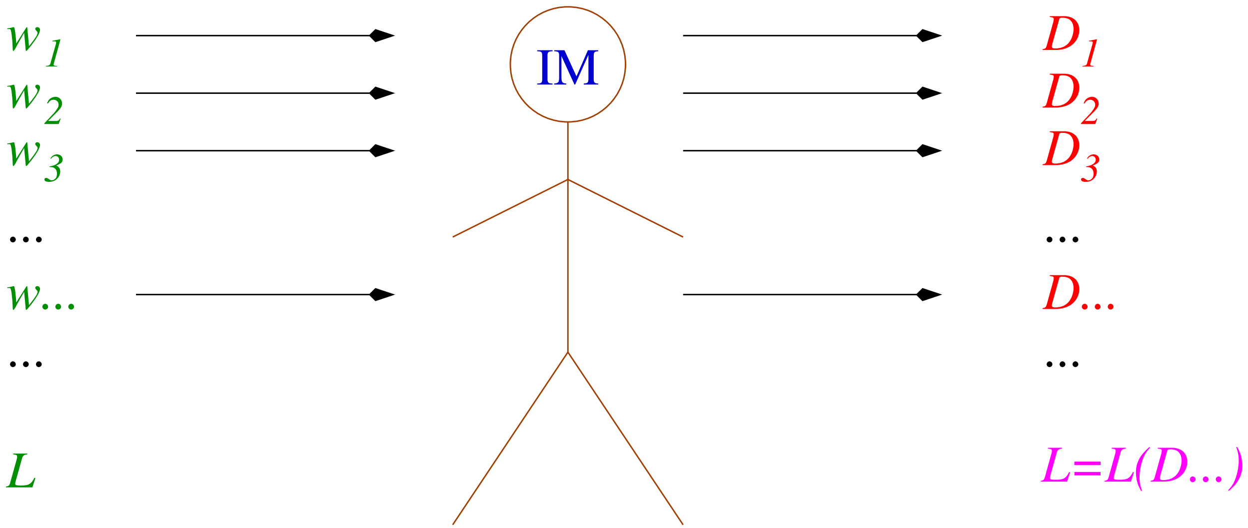
F *konvergiert* auf Text T, falls es einen *Grenzwert* $i \in \mathbb{N}$ gibt, so dass $\exists\, n_0 \in \mathbb{N}\ \forall\, n \geq n_0 : F(T[n]) = i$.
In Zeichen: $F(T) \downarrow= i$.

F *identifiziert* Text T genau dann, wenn
$\exists i \in \mathbb{N} : F(T) \downarrow= i \wedge L(i) = Inh(T)$.

F *identifiziert* $L \in \mathcal{E}$ genau dann, wenn
$\forall$ Texte $T : Inh(T) = L \Rightarrow F$ identifiziert T.

F *identifiziert* $\mathcal{L} \subseteq \mathcal{E}$ genau dann, wenn
$\forall L \in \mathcal{L} : F$ identifiziert L.
(sonst heißt $\mathcal{L}$ *nicht identifizierbar*.)
Anstelle von *Identifikation* spricht man auch vom *Textlernen*.

Golds Lernmodell des Textlernens



Beispiel 1

Betrachte die Klasse $\mathcal{FIN}$ aller endlichen Sprachen.
$\mathcal{FIN}$ ist identifizierbar.

Betrachte folgenden Lerner F:
Zu gegebenem Anfangsstück $T[n]$ von Text T für $L \in \mathcal{FIN}$ extrahiert F zunächst den Inhalt $L_n = \mathrm{Inh}(T[n])$.
Dann sortiert F L_n und erhält Folge $x_1, \ldots, x_r$ mit $r = |\,\mathrm{Inh}(T[n])|$.
Als Hypothese äußert F die Gödelnummer g_n einer TM $T(x_1, \ldots, x_r)$, die bei Eingabe von x schaut, ob x in der Liste $x_1, \ldots, x_r$ enthalten ist.
Wenn ja, hält die TM, andernfalls geht sie in eine Endlosschleife.
Offenbar gilt: $W_{g_n} = \mathrm{Inh}(T[n])$.
Da L endlich ist, konvergiert der Lerner F für jeden Text T von L.
Der Sortierschritt gewährleistet, dass die Gödelnummer, auf die F konvergiert, für jeden Text T von L sogar dieselbe ist.

Beispiel 2

Betrachte die Klasse $\mathcal{L} = \{\mathbb{N} - \{x\} \mid x \in \mathbb{N}\}$.
$\mathcal{L}$ ist identifizierbar.

Zu $\sigma \in SEQ$ sei x_σ die kleinste Zahl, die nicht in $\mathrm{Inh}(\sigma)$ enthalten ist.
Betrachte folgenden Lerner F:
$F(\sigma)$ ist die kleinste Gödelnummer für die Sprache $\mathbb{N} - \{x_\sigma\}$.
Nach Konstruktion ist klar, dass F die Klasse $\mathcal{L}$ identifiziert.
Weniger klar ist, ob dieser Lerner durch eine TM dargestellt werden könnte...

Die *Konkatenation* von $\sigma, \tau \in SEQ$ notieren wir durch Hintereinanderschreiben. $\sigma \sqsubseteq \tau$ ($\sigma \sqsubset \tau$) drücke die (echte) *Präfixrelation* aus.
Gilt $\sigma_0 \sqsubseteq \sigma_1 \sqsubseteq \ldots \sqsubseteq \sigma_n \sqsubseteq \ldots$, so bezeichne $\bigcup_{n\in\mathbb{N}} \sigma_n$ den dadurch eindeutig festgelegten Text.

Es seien $L \in \mathcal{E}$, ein Lerner F und $\sigma \in SEQ$ gegeben.
σ heißt *Schlussfolge* (locking sequence) für F auf L genau dann, wenn

(a) $\mathrm{Inh}(\sigma) \subseteq L$,

(b) $L(F(\sigma)) = L$,

(c) $\forall \tau \in SEQ, \mathrm{Inh}(\tau) \subseteq L \Rightarrow F(\sigma\tau) = F(\sigma)$.

Satz 1 (Blum & Blum): Identifiziert F eine Sprache $L \in \mathcal{E}$, so gibt es eine Schlussfolge für F auf L.

Beweis: Gegeben: $L \in \mathcal{E}$ und F. O.E. sei F auf ganz SEQ definiert.
Annahme: Es gibt keine Schlussfolge σ für F auf L. D.h.:

$$\begin{array}{ll} & \forall \sigma \in SEQ,\ Inh(\sigma) \subseteq L \ \wedge \ L(F(\sigma)) = L \\ \Longrightarrow & \exists \tau \in SEQ:\ Inh(\tau) \subseteq L \wedge F(\sigma\tau) \neq F(\sigma). \qquad (*) \end{array}$$

Wir zeigen: Aus $(*)$ folgt die Existenz eines nicht identifizierbaren Textes T für L.
Es sei $U = u(0), u(1), u(2), \ldots$ ein beliebiger Text für L.
Wir definieren induktiv eine Folge $\sigma_n \sqsubset T$.

$n = 0 : \sigma_0 = \emptyset$.
$n \to n+1$:
1. Fall: $L(F(\sigma_n)) \neq L$. Setze $\sigma_{n+1} = \sigma_n u(n)$.
2. Fall: $L(F(\sigma_n)) = L$.
$(*)$ gewährleistet die Existenz von $\tau \in SEQ$, $Inh(\tau) \subseteq L$, $F(\sigma_n\tau) \neq F(\sigma_n)$.
Setze $\sigma_{n+1} = \sigma_n \tau u(n)$.

Es ist $\sigma_i \sqsubseteq \sigma_{i+1}$ für alle i; setze daher $T = \bigcup_{n\in\mathbb{N}} \sigma_n$.
Da U ein Text für L ist, ist auch T ein Text für L.
F konvergiert aber nicht auf T, denn für alle $n+1 \in \mathbb{N}$ ist
entweder $L(\sigma_n) \neq L$
oder es gibt ein $\tau \in SEQ$, $\sigma_n\tau \sqsubseteq \sigma_{n+1}$, $F(\sigma_n\tau) \neq F(\sigma_n)$. $\square$

Satz 2 (Angluin): $\mathcal{L} \subseteq \mathcal{E}$ ist identifizierbar genau dann, wenn
$\forall L \in \mathcal{L} \exists D_L \subseteq L, |D_L| < \infty \;\; \forall L' \in \mathcal{L},\; L' \neq L,\; D_L \subseteq L' \Rightarrow L' \not\subset L.$

Obiges D_L heißt auch *charakteristische (Teil-)Menge* von L.

Beweis: Es sei $\mathcal{L} \subseteq \mathcal{E}$ gegeben.
(1) Angenommen, F identifiziert $\mathcal{L}$.
Dann gibt es zu jedem $L \in \mathcal{L}$ eine Schlussfolge σ_L für F auf L.
Setze $D_L = \mathrm{Inh}(\sigma_L)$.
Zu zeigen: $\forall L' \in \mathcal{L},\; L' \neq L,\; D_L \subseteq L' \Rightarrow L' \not\subset L$.
Betrachte die Situation $D_L \subseteq L'$, $L' \subseteq L$ für zwei $L, L' \in \mathcal{L}$.
Es sei T ein Text <u>für L'</u> mit $T[|\sigma_L|] = \sigma_L$ (wegen $\mathrm{Inh}(\sigma_L) \subseteq L'$).
Da σ_L Schlussfolge, gilt: $L(F(\sigma_L)) = L$, also $L' = L(F(T)) = L$.

(2) Annahme: $\forall L \in \mathcal{L} \exists D_L \subseteq L(|D_L| < \infty) \forall L' \in \mathcal{L},\ L' \neq L,\ D_L \subseteq L' \Rightarrow L' \not\subseteq L$. [+]
Definiere einen Lerner F' wie folgt:
$F'(\sigma)$ sei das kleinste $i \in \mathbb{N}$, so dass für ein $L \in \mathcal{L}$ gilt:
(a) $L(i) = L$,
(b) $D_L \subseteq Inh(\sigma) \subseteq L$.
Gibt es solch ein i nicht, so sei F' undefiniert.
Zu zeigen: F' identifiziert $\mathcal{L}$.
Wir fixieren $L' \in \mathcal{L}$ und einen Text T für L'.
Es gibt dann ein $n_0 \in \mathbb{N}$, so dass $D_{L'} \subseteq Inh(T[n_0]) \subseteq L'$.
Sei i_0 der kleinste Index für L'. Nach Konstruktion von F' gilt:
F' konvergiert gegen i_0 auf T (und identifiziert so T), falls:
$\forall j < i_0, L = L(j) \in \mathcal{L} \forall m \in \mathbb{N}, D_L \subseteq Inh(T[m]) \exists n' > m \forall n \geq n' : Inh(T[n]) \not\subseteq L$.
Sei also $j < i_0$ und $m \in \mathbb{N}$ gegeben mit $L = L(j) \in \mathcal{L}, D_L \subseteq Inh(T[m])$.
Wegen $L \neq L'$ und $D_L \subseteq Inh(T) = L'$ gibt es wegen [+] ein $x_0 \in L' \setminus L$.
T ist ein Text für L', so dass es ein n' gibt mit $x_0 \in Inh(T[n])$ für alle $n \geq n'$.
Daher ist $Inh(T[n]) \not\subseteq L$. $\square$

Möglichkeiten und Grenzen des Textlernens

Folgerung: $\mathcal{FIN}$ ist identifizierbar.

Eine Sprachklasse, die alle endlichen sowie eine unendliche Sprache enthält, heißt *superfinit*.
Folgerung (Gold): Keine superfinite Sprachklasse ist identifizierbar.

Folgerung: Die Klasse der regulären Sprachen ist nicht identifizierbar.

Warum sind dies einfache Folgerungen ?

Hinweis: Im ersten Fall: Wähle als L vom Satz von Angluin eine unendliche Sprache.

Einordnung und Ausblick

Wir haben hier den Einstieg in die Lerntheorie begonnen.

Bislang haben wir lediglich das Sprachenlernen untersucht und auch keinerlei rekursionstheoretische (oder anders geartete) Einschränkungen für die Lerner (Forscher) behandelt. Dies ist Gegenstand der folgenden Vorlesungen.

Dafür benötigen wir noch einen weiteren Satz aus der Rekursionstheorie, der im Folgenden nachgeliefert wird.

Der zweite Fixpunktsatz von Kleene

Satz 3: Ist ψ eine partiell-rekursive zweistellige Funktion, so gibt es einen Index e mit: $h_e(y) = \psi(e, y)$ für alle y.

<u>Beweis:</u> Das smn-Theorem liefert speziell: Es gibt eine rekursive Funktion $s_1^1 : \mathbb{N}^2 \to \mathbb{N}$, sodass für alle $p, x, y \in \mathbb{N}$, $h_{s_1^1(p,x)}(y) = h_p(x, y)$.

Es gibt einen Index d mit $h_d(x, y) = \psi(s_1^1(x, x), y)$ für die vorgelegte Funktion ψ.

Betrachte $e = s_1^1(d, d)$.

$\rightsquigarrow h_e(y) = h_{s_1^1(d,d)}(y) = h_d(d, y) = \psi(s_1^1(d, d), y) = \psi(e, y)$ für alle y.