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

Gemischtes Programm

Wir werden heute:

1. einen der für die Informatik möglicherweise wichtigsten Zusammenhänge aus dem Bereich der Rekursionstheorie kennenlernen sowie

2. eine wichtige Variante unseres Lernparadigmas durchsprechen.

Hilberts Traum (Hilberts Programm 1921)

- Kann die Arbeit eines Mathematikers mechanisiert werden?
- Was bedeutet das überhaupt?
- Wir brauchen eine Menge von Axiomen und Schlussregeln mit dem
- Ziel: Automatische Herleitung aller Wahrheiten
- Informatik: Deduktionssysteme, Beweissysteme, ...

Peano-Axiome: Ein “beispielhaftes” Axiomensystem

axiomatische Definition der Menge der natürlichen Zahlen $\mathbb{N}$ durch Giuseppe Peano (1889)
eigentlich von Richard Dedekind in “Was sind und was sollen die Zahlen ?” (1888)

1. 0 ist eine natürliche Zahl.
2. Zu jeder natürlichen Zahl n gibt es genau einen Nachfolger n', der ebenfalls eine natürliche Zahl ist.
3. Es gibt keine natürliche Zahl, deren Nachfolger 0 ist.
4. Zwei verschiedene natürliche Zahlen n und m besitzen stets verschiedene Nachfolger n' und m'.
5. Enthält eine Menge X die Zahl 0 und mit jeder natürlichen Zahl n auch stets deren Nachfolger n', so enthält X bereits alle natürlichen Zahlen. *Induktionsaxiomenschema*
(Ist X dabei selbst eine Teilmenge der natürlichen Zahlen, dann ist $X = \mathbb{N}$.)

Peano verwendet dabei die Begriffe 0, Zahl und *Nachfolger*.
Hierin lassen sich nun Begriffe wie Addition, Multiplikation, “kleiner als”, etc. definieren.

Aussagen und Beweise

Warum gilt: “2 ist von 4 verschieden” ? (oder kurz: $2 \neq 4$)

Rückführung auf die soeben eingeführte Schreibweise:
$2 = (0')'$ sowie $4 = (((0')')')'$.
Wegen 3. gilt: $0 \neq 2 = (0')'$, denn 0 ist nicht Nachfolger irgendeiner Zahl, während $(0')'$ offenbar Nachfolger von $0'$ ist.
Damit ergibt sich wegen 4.: $1 = 0' \neq ((0')')' = 2' = 3$.
Nochmalige Anwendung von 4. liefert schließlich: $1' \neq 3'$, also $2 \neq 4$.

Zur Logik

Da hier keine “richtige Logik” vorausgesetzt werden kann, muss das Folgende notwendig skizzenhaft bleiben.

In gewissem Sinne gibt es zwei “Wahrheitsbegriffe” in jeder formalen Logik:

1. Gültigkeit in einem Modell bzw. in allen Modellen
2. Ableitbarkeit in einem gewissen Kalkül

Bsp.: Modus ponens “im Kalkül”: $P \to Q, P \vdash Q$

Eine mögliche Interpretation eines Logik-Kalküls erfolgt über Wahrheitswerte; die Gültigkeit des MP in diesem Modell lässt sich mit einer Wahrheitstafel überprüfen:

P	Q	$\neg P$	$P \to Q (= \neg P \vee Q)$
1	1	0	1
1	0	0	$0, \ldots$

Weitere Bausteine

Zum Nachvollziehen von Gödels Konstruktion benötigt man:

1. ein hinreichend mächtiger Logikkalkül
2. eine elementare Aximatik der natürlichen Zahlen

Beides zusammen soll beschrieben sein als eine aufzählbare Menge AX von Axiomen.
Axiomenschemata wie die Induktion werden also "einzeln aufgezählt".

In dem so entstandenen Kalkül lassen sich:
a) Aussagen formulieren; durch geeignete Kodierungen c_A entsprechen Aussagen gewisse Zahlen;
b) Beweise (im Sinne von Ableitungen im Kalkül) aufschreiben; durch geeignete Kodierungen c_B entsprechen Beweisen gewisse Zahlen. ...

Gödelnummern a): Aussagen sind Zahlen

Sind P und Q Aussagen und gilt bezüglich der i-ten Ableitungsregel $\vdash_i$: $P \vdash_i Q$, so gibt uns das eine binäre Relation R_i zwischen Zahlen mit $(c_A(P), c_A(Q)) \in R_i$.

In "vernünftigen" Kalkülen ist R_i entscheidbar.
Die beweisbaren Aussagen BEW lassen sich daher beschreiben als kleinstmögliche Zahlen(!)-Menge, die $c_A(AX)$ enthält und abgeschlossen ist unter allen Relationen R_i, d.h., falls $n \in BEW$ und $(n, m) \in R_i$, so $m \in BEW$.
BEW enthält also alle "Gödelnummern" von beweisbaren Aussagen.

Gödelnummern b): Beweise sind Zahlen

Ein Beweis β für P lässt sich als Folge von Kalkülschritten begreifen, wobei nur auf Axiome oder bereits bewiesene Aussagen zurückgegriffen werden darf.
Kalkülschritte und Aussagen entsprechen Zahlen, und so eine Folge von Zahlen lässt sich wiederum als Zahl auffassen.
Also gibt es eine binäre Zahlenrelation B, sodass $(c_B(\beta), c_A(P)) \in B$, und es gilt: $n \in BEW$ gdw. es gibt $m \in \mathbb{N}$, sodass $B(n, m)$.

Gödelnummern a1): Variablen sind Zahlen

Die entsprechende Kodierung c_V ist trivial.
Gödelnummern a2): (einstellige) Aussageformen sind Zahlen

Aussageformen enthalten noch Variablen und werden durch Belegung derselben durch Zahlen zu Aussagen.
Aus a) und a1) ergeben sich "Gödelisierungsmöglichkeiten", z.B. c_{AF}.

Betrachte Substitutions-Abbildung $sub : \mathbb{N} \to \mathbb{N}, n \mapsto m$: $n = \langle s, r \rangle$ entspricht Aussageform $c_{AF}^{-1}(s) = F(y)$; substituiere darin y durch r; dies liefert Aussage P mit Gödelnummer $m = c_A(P)$.
Für Zahlen n, m sei $q(n, m)$ die Relation: "n ist keine Gödelnummer für einen Beweis für die Aussage mit der Gödelnummer $sub(\langle m, m \rangle)$."
Wäre für ein gewisses m die Relation $q(n, m)$ für alle n gültig, so wäre die entsprechende Aussage $sub(\langle m, m \rangle)$ nicht im Kalkül ableitbar.

Betrachte nun die spezielle Aussageform $F(y) = \forall n(q(n,y))$.
Diese “bedeutet”: Es gibt keinen Beweis (im Kalkül) für die Aussage mit der Gödelnummer $sub(\langle y,y\rangle)$.
Betrachte die Aussage $P_F = c_A^{-1}(sub(\langle c_{AF}(F), c_{AF}(F)\rangle))$.
P_F entsteht aus $F(y)$ durch Ersetzen von y durch $c_{AF}(F)$.
Ist P_F wahr, so ist P_F innerhalb des Kalküls nicht beweisbar.
Ist P_F falsch, so ist P_F beweisbar innerhalb des Kalküls.

Entweder sind also nicht alle wahren Aussagen innerhalb des Kalküls beweisbar (der Kalkül ist *unvollständig*) oder aber falsche Aussagen sind ableitbar. Daraus kann man einen Widerspruch zur *Konsistenz* (Widerspruchsfreiheit) des Kalküls konstruieren.

Satz (Gödel): Jeder hinreichend aussagestarke konsistente Kalkül der elementaren Arithmetik ist unvollständig.

Weitere Bemerkungen:

Nimmt man an, ein Kalkül “für” elementare Arithmetik gestattet nur die Ableitung von wahren Aussagen (und dies sollte für jeden Kalkül gelten, der “für” die natürlichen Zahlen aufgestellt wird), so muss die oben konstruierte Aussage P_F wahr sein für das “Standardmodell” des Kalküls, nämlich die natürlichen Zahlen.

Somit zeigt dann der Beweis, dass es andere Modelle für so ein Kalkül geben muss, in denen P_F falsch ist.

Kalküle abstrahieren nämlich immer (nur) gewisse Eigenschaften des “Standardmodells”, besitzen aber oft mehrere andere Modelle, für die man durch Ableitungen im Kalkül ebenfalls wahre Aussagen beweist.

Hinweis: Rolle des Parallelenaxioms in der Euklidischen Geometrie

Eine Formalisierung von Boolos (erschienen in *Mind*, 1994)

Betrachte ein Kalkül, in dem eine Aussage P selbst dargestellt werden kann (Gödelisierung) und ebenso ihre Ableitbarkeit formalisiert werden kann.

$\vdash \Box P$: Es gibt Ableitung für die Ableitbarkeit von P.

Der Kalkül möge Folgendes erfüllen:

(A) Ist $\vdash P$, so auch $\vdash \Box P$.

(B) $\vdash (\Box(P \to Q) \to (\Box P \to \Box Q))$.

(C) $\vdash (\Box P \to \Box\Box P)$.

(D) $\vdash p \leftrightarrow \neg\Box p$ für eine Aussage p.

(A), (B), (C) beschreiben die Möglichkeit des Kalküls, innerhalb des Kalküls über die Ableitbarkeit im Kalkül Aussagen machen zu können.

(D) formalisiert so die soeben konstruierte Aussage P_F.

L1: Falls $\vdash (P \to Q)$, so $\vdash (\Box P \to \Box Q)$.

Bew.: Aus $\vdash (P \to Q)$ folgt mit (A) $\vdash \Box(P \to Q)$ und somit mit (B) die Beh. wegen Modus Ponens.

L2: $\vdash p \to \neg\Box p$.

L3: $\vdash \Box p \to \Box\neg\Box p$ wegen L1 und L2

L4: $\vdash \Box p \to \Box\Box p$ wegen (C)

L5: $\vdash \neg\Box p \to (\Box p \to \bot)$. $\bot$ ist das Falsum; die Aussage ist tautologisch.

L6: $\vdash \Box\neg\Box p \to \Box(\Box p \to \bot)$. (L1 und L5)

L7: $\vdash \Box(\Box p \to \bot) \to (\Box\Box p \to \Box\bot)$ wegen (B)

L8: $\vdash \Box p \to \Box\bot$ (L3, L6, L7 sowie L4 mit Modus Ponens)

L9: $\vdash \neg\Box\bot \to p$. (L8 und (D) mit Kontraposition)

L10: $\vdash \Box\neg\Box\bot \to \Box p$. (L1 und L9)

L11: $\vdash \neg\Box\bot \to \neg\Box\neg\Box\bot$. (L8 und L10 mit Kontraposition)

L12: $\vdash \neg\Box\bot \to \Box\neg\Box\bot$ wegen (A).

Ist der Kalkül widerspruchsfrei, so darf $\vdash \neg\Box\bot$ nicht gelten. Aber natürlich ist $\neg\Box\bot$ wahr, denn sonst würde es ja einen Beweis für das Falsum geben.

Weitere Themen der Lerntheorie

Funktionenlernen mit Anomalien

$g =^n f$ gdw. $|\{x \mid f(x) \neq g(x)\}| \leq n$.

$g =^* f$ gdw. $\exists n(g =^n f)$.

Es sei $a \in \mathbb{N} \cup \{*\}$.

$f \in \mathbf{Ex}^a(M)$ gdw. $M(f) \downarrow$ und $h_{M(f)} =^a f$.

$\mathbf{Ex}^a = \{\mathcal{S} \mid \exists$partiell-rek. $M(\mathcal{S} \subseteq \mathbf{Ex}^a(M))\}$.

Auch beim *Lernen mit Anomalien* können wir uns auf total berechenbare Forscher beschränken.

Wir wollen zeigen:

Satz : $\mathbf{Ex} = \mathbf{Ex}^0 \subsetneq \mathbf{Ex}^1 \subsetneq \mathbf{Ex}^2 \subsetneq \mathbf{Ex}^3 \subsetneq \cdots \subsetneq \mathbf{Ex}^*$.

Selbstbeschreibende Funktionen mit Anomalien

Für $a \in \mathbb{N} \cup \{*\}$ enthalte $\mathcal{ASD}^a$ alle $f \in \mathcal{R}$ mit $h_{f(0)} =^a f$.

Klar: $\mathcal{SD} = \mathcal{ASD}^0$.

Der triviale $\mathcal{SD}$-Lerner zeigt: $\mathcal{ASD}^a \in \mathbf{Ex}^a$.

Wir zeigen das folgende stärkere Ergebnis, aus dem der zuvor festgehaltene Satz folgt:

Lemma: $\forall m \in \mathbb{N} : \mathcal{ASD}^{m+1} \in \mathbf{Ex}^{m+1} \setminus \mathbf{Ex}^m$.

Beweis: Fixiere m. Wir präsentieren eine Stufenkonstruktion für eine zweistellige rekursive Funktion ψ (eigentlich mehrere solche, je nach Unterfall).

Wie schon desöfteren liefert dann das zweite Fixpunkttheorem von Kleene eine Funktion $h_e(x) = \psi(e, x)$ mit $h_e \in \mathcal{ASD}^{m+1} \setminus \mathbf{Ex}^m$.

Wir nehmen wieder an, es gäbe einen $\mathbf{Ex}^m$-Lerner M für $\mathcal{ASD}^{m+1}$ und führen diese Annahme im Folgenden zum Widerspruch.

Konstruktionsdetails

Stufe 0: Setze $\psi(e, 0) = e$.
Sei x_s das kleinste x, sodass $\psi(e, x)$ auf Stufe $(s-1)$ noch nicht definiert wurde, z.B.: $x_1 = 1$.
Für $y \leq m+1$ definiere folgende Hilfsfunktion:

$$\psi_y(e, x) = \begin{cases} \psi(e, x), & \text{falls} \quad x < x_s, \\ y, & \text{falls} \quad x_s \leq x \leq x_s + m, \\ 0, & \text{sonst} \end{cases}$$

$\psi^e(x) = \psi(e, x)$ bzw. $\psi^e_y(x) = \psi_y(e, x)$ seien entsprechende Scharen einstelliger Funktionen.
Führe jetzt die beiden folgenden Fälle “parallel” aus (Schwalbenschwanz):
a. Für $i = 1, 2, \ldots,$ setze $\psi(e, x_s + m + i) = 0$.
b. Suche nach (y, n) mit $y \leq m+1$, $n \geq x_s + m + 1$, $M(\psi^e[x_s]) \neq M(\psi^e_y[n])$.
Ist b. irgendwann erfolgreich und findet (y, n), so setze $\psi(e, x) = \psi_y(e, x)$ für bislang undefinierte Werte $x < n$. Danach gilt: $\psi^e[n] = \psi^e_y[n]$ sowie $x_{s+1} \geq n$. Die nächste Stufe wird anschließend konstruiert usf.
Terminiert b. nicht, so ist dies die letzte Stufe der Konstruktion, und a. wird ψ^e vollständig definieren.

Beweisdetails

Werden unendlich viele Stufen “besucht”, so macht M unendlich viele Meinungsänderungen auf jedem ψ^e. Der zweite Kleenesche Fixpunktsatz liefert ein e mit $h_e(x) = \psi(e, x)$ und $h_e \in \mathcal{ASD}^{m+1} \setminus \mathbf{Ex}^m$.

Andernfalls gibt es eine Stufe s, die niemals verlassen wird.

Betrachte die auf Stufe s definierte Funktion $\psi_y(e, x)$.

Sei $p = M(\psi^e[x_s])$.

Nach Konstruktion (und dem zweiten Kleenesche Fixpunktsatz) gilt für ein geeignetes e:

Für alle $y \leq m + 1$: $\psi_y^e \in \mathcal{ASD}^{m+1}$ und $M(\psi_y^e[n]) = p$.

Da $|\{h_p(x_s + i) \mid i \leq m\}| \leq m$, gibt es wenigstens ein $z \leq m + 1$, sodass $\psi_z(e, x) = z$ sich für alle $x_s \leq x \leq x_s + m$ von $\{h_p(x_s + i) \mid i \leq m\}$ unterscheidet.

Wähle nun dieses $\psi_z^e \in \mathcal{ASD}^{m+1}$.

Offenbar kann M ψ_z^e nicht mit m-Anomalien identifizieren. $\square$

Verhaltenskorrektes Lernen

(Txt)Ex-Lerner müssen gegen eine gewisse Gödelnummer einer Funktion oder Menge konvergieren. Diese Bedingung kann man dadurch abschwächen, dass nur gefordert wird, dass “gegen die richtige Funktion” konvergiert wird:

Ein Lerner M identifiziert f nach dem ***Bc**-Kriterium* oder *verhaltenskorrekt* (behaviourally correct), kurz $f \in$ **BC**(M), falls für fast alle n gilt: $h_{M(f[n])} = f$.

Bc $= \{\mathcal{S} \mid (\exists$ part. rek. $M)\mathcal{S} \subseteq$ **Bc**$(M)\}$.

Satz : **Ex**$^* \subsetneq$ **Bc**.

Beweis: Die Inklusion ergibt sich aus der Beobachtung, dass ein **Bc**-Lerner M' einen totalen **Ex***-Lerner M als “Unterprogramm” verwenden kann und dann überprüfen kann, welche Fehler die bisherige Hypothese $h_{M(f[n])}$ macht; M' äußert dann eine Hypothese, die endlich viele Ausnahmen in das Programm $M(f[n])$ einfügt.

Als trennendes Beispiel wähle man: BSD $= \{f \in \mathcal{R} \mid |\{n \in \mathbb{N} \mid h_{f(n)} = f\}| = \infty\}$.

BSD $\in$ **Bc** ist eine nette Übungsaufgabe.

Für den Beweis von BSD $\notin$ **Ex*** benötigt man einen weiteren Fixpunktsatz (auch Operator-Rekursions-Theorem genannt), der uns hier nicht zur Verfügung steht. $\square$

Verhaltenskorrektes Lernen mit Anomalien; sei $a \in \mathbb{N} \cup \{*\}$.

Ein Lerner M identifiziert f nach dem ***Bc**^{a}-Kriterium* oder *verhaltenskorrekt mit a Anomalien*, kurz $f \in \mathbf{BC}^a(M)$, falls für fast alle n gilt: $h_{M(f[n])} =^a f$.

$\mathbf{Bc}^a = \{\mathcal{S} \mid (\exists$ part. rek. $M)\mathcal{S} \subseteq \mathbf{Bc}^a(M)\}$.

Satz : $\mathbf{Bc} = \mathbf{Bc}^0 \subsetneq \mathbf{Bc}^1 \subsetneq \mathbf{Bc}^2 \subsetneq \cdots \subsetneq \mathbf{Bc}^* = \mathbf{Func}$.

Beweis: Die schwache Inklusionskette ist trivial.

Als trennendes Beispiel wähle: $\mathrm{ABSD}^m = \{f \in \mathcal{R} \mid |\{n \in \mathbb{N} \mid h_{f(n)} =^m f\}| = \infty\}$.

Der Beweis benutzt wieder das Operator-Rekursions-Theorem.

Für die letzte Gleichheit muss man zeigen: $\mathcal{R} \in \mathbf{Bc}^*$. (s. nächste Folie) $\square$

Lemma: $\mathcal{R} \in \mathbf{Bc}^*$.

Beweis: Es messe $\Phi_p(y)$ den Zeitbedarf der p-ten Turingmaschine bei Eingabe von y. Betrachte den wie folgt definierten Lerner M:

$$h_{M(\sigma)}(x) = \begin{cases} h_p(x), & \text{falls} \quad p \text{ ist die kleinste Zahl } \leq |\sigma|, \\ & \text{sodass} \quad \max(\{\Phi_p(y) \mid y < |\sigma|\}) \leq x \wedge h_p[|\sigma|] = \sigma. \\ 0 & \text{sonst.} \end{cases}$$

Zu zeigen: Bei Eingabe eines kanonischen Textes von $f \in \mathcal{R}$ konvergiert M gegen eine $*$-Variante von f.

Sei p_0 die kleinste Gödelnummer von f.

Wähle n_0 genügend groß, sodass: (a) $p_0 \leq n_0$ und (b) $(\forall p < p_0) h_p[n_0] \neq f[n_0]$.

Beh.: Für $n \geq n_0$ ist $h_{M(f[n])} =^* f$.

Setze $x_n = \max(\{\Phi_p(y) \mid y < n\})$ und betrachte $x \geq x_n$.

Nach Konstruktion von M gilt: $h_{M(f[n])}(x) = h_{p_0}(x)$ wegen $x \geq x_n$,

und nach Wahl von p_0 gilt: $f(x) = h_{M(f[n])}(x)$.

Also ist die Hypothese höchstens für $x < x_n$ (also endlich oft) falsch. $\square$

Bemerkungen zum Sprachenlernen

Man kann die soeben diskutierten Lernvarianten auch für das Sprachenlernen definieren.

Hierbei ergeben sich folgende wichtige Ergebnisse:

1. **TxtEx** $=$ **TxtEx**$^0 \subsetneq$ **TxtEx**$^1 \subsetneq$ **TxtEx**$^2 \subsetneq \cdots \subsetneq$ **TxtEx***.
2. **TxtBc** $=$ **TxtBc**$^0 \subsetneq$ **TxtBc**$^1 \subsetneq$ **TxtBc**$^2 \subsetneq \cdots \subsetneq$ **TxtBc***.
3. **TxtBc** und **TxtEx** sind unvergleichbar.

Punkt 3 zeigt einen deutlichen Ergebnisunterschied zum Funktionenfall.

Man kann 1. z.B. zeigen, indem man auf die entsprechenden Funktionenergebnisse zurückgreift durch Verwendung der folgenden Codierung einer rekursiven Funktion f: $L_f = \{\langle n, f(n)\rangle \mid n \in \mathbb{N}\}$; für $\mathcal{S} \subseteq \mathcal{R}$ sei $\mathcal{L}_{\mathcal{S}} = \{L_f \mid f \in \mathcal{S}\}$.

Dann diskutiere man $\mathcal{L}_{\mathsf{ASD}^m}$.

Details siehe Kap. 6 in STL

Weitere Themen der Lerntheorie

Motivation häufig: Erzielung “praktisch relevanterer” Ergebnisse.

— Inferenz von Approximationen (im wahrscheinlichkeitstheoretischen Sinne) Kap. 7 STL

Hinweis: Das Thema “PAC-Lernen” wird in der Lernalgorithmenvorlesung zumindest angerissen.

— verrauschte Eingaben; Eingaben in besonderer Ordnung; mehrere Informationsquellen Kap. 8 STL

— Beschränkungen der Hypothesengröße Kap. 10 STL

— Beschränkungen der Hypothesenwechsel Kap. 12 STL

Weitere Themen der Lerntheorie

Motivation manchmal: vertiefte Querbezüge zur Rekursionstheorie

— Orakel-TM als Lerner Kap. 11 STL
— robustes Lernen Kap. 13 STL: Invarianz der "Lernbarkeit" einer Sprache gegenüber von rekursiven Transformationen

Hinweis: Solche Invarianz dient auch dazu, Gegenbeispiele wie die selbstbeschreibenden Funktionen auszuschließen, denn für die Transformation $\Psi(f) = g$ mit $g(x) = f(x+1)$ gilt:
$\mathcal{R} = \{\Psi(f) \mid f \in \mathcal{SD}\} \notin$ **Ex**.

Diese Studien hängen auch (historisch) mit der (von Fulk widerlegten) These von Gold zusammen, derzufolge alle (robusten) **Ex**-lernbaren Funktionenklassen durch *Enumeration* gelernt werden können, d.h., durch einen Lerner, der stets die kleinste Gödelnummer äußert, deren Funktion mit den bisherigen Daten kompatibel ist. Jede aufzählbar indizierbare Funktionenklasse kann so gelernt werden, aber $\mathcal{SD}$ z.B. nicht.

Rückblick

Wir haben gesehen:

Rekursions- und Lerntheorie sind aufs Engste miteinander verwoben.

Ähnliche enge Bezüge weisen Mathematische Logik und Rekursionstheorie auf.

Der Gödelsche Unvollständigkeitssatz ist nur ein Beispiel.

Daher Empfehlung: VL Mathematische Logik im nächsten Semester.

Nutzen Sie auch unsere (Pro-)Seminar-Angebote: Thema Logik (!)