

Algorithmisches Lernen Algorithmic Learning



01011101010001010110111
010111000100011

Theorie des Algorithmischen Lernens

Sommersemester 2006

0101110101000101011011101011100
0100011 010 101011011101011100
10

01011101010001010110111010111000100011
01011101010001010110111010111000100011

Gunter Grieser



01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Motivation

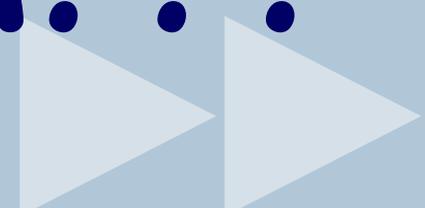
01011101010001010110111
010111000100011

Was ist Lernen???

0101110101000101011011101011100
0100011 01011101010001010110111011100
0100011

01011101010001010110111010111000100011

01011101010001010110111010111000100011



01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Learning

01011101010001010110111
010111000100011

- Learning is any change in a system that allows it to do better the second time on repetition of the same task or another task drawn from the same population.

01011101010001010110111010111000100011

01011101010001010110111010111000100011

Herbert Simon, 1983



Algorithmisches Lernen

Algorithmic Learning



Learning

01011101010001010110111
010111000100011

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

- Learning is the organization of experience.

01011101010001010110111010111000100011
01011101010001010110111010111000100011

Dana Scott, 1983

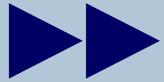
01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Learning

01011101010001010110111
010111000100011

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

- Learning is making useful changes in our minds.

01011101010001010110111010111000100011

01011101010001010110111010111000100011

01011101010001010110111010111000100011

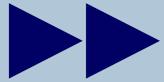
Marvin Minsky, 1985

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Learning

01011101010001010110111
010111000100011

- Learning is constructing or modifying representation of what is being experienced.



Ryszard Michalski, 1986

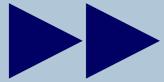
01011101010001010110111010111000100011

00100011



Algorithmisches Lernen

Algorithmic Learning



Learning

01011101010001010110111
010111000100011

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

- We will not state what learning is.

Pieter Adriaans & Dolf Zantinge, 1996... 1998
im Kapitel 2 „What is learning?“

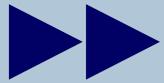
01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Learning

Numerous mathematical models of learning have been proposed during the last three decades. Nevertheless, different models give vastly different results concerning the learnability and non-learnability of objects one wants to learn. Hence, finding an appropriate definition of learning which covers most aspects of learning is also part of the goals aimed at in algorithmic learning theory.

Michael M. Richter, Carl H. Smith, Rolf Wiehagen, Thomas Zeugmann, 1998



Algorithmisches Lernen

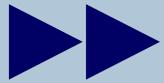
Algorithmic Learning

The Scenario of Inductive Inference

- Irgend etwas soll gelernt werden - aber was ?!
- ... und woraus, d.h. aus was für Information ?
- Soll aus jeglicher Information gelernt werden, oder gibt es "mehr oder weniger" sinnvolle ?
- Und falls man etwas gelernt hat, wie drückt man das aus, mit welchen Mitteln formuliert man das ?
- Wer oder was lernt ?
- Was ist ein "richtiges" Lernergebnis ...
- ... und welche Lernprozesse sind letztendlich erfolgreich ?

Algorithmisches Lernen

Algorithmic Learning



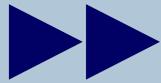
Das Szenario der Induktiven Inferenz

- Klasse zu lernender Objekte
- Informationsfolgen: (Gegen-)Beispiele
- korrekte (Limes-)Spezifikationen
- Lernverfahren
- Hypothesenraum
- Semantik von Hypothesen
- Erfolgskriterium



Algorithmisches Lernen

Algorithmic Learning



Essentials

01011101010001010110111

010111000100011

Wenn man Lernen im Sinne eines solchen Szenarios ernst meint, dann hat man auch unausweichlich mit ein paar Phänomenen zu tun, wie zum Beispiel:

0101110101000101011011101011100

011101011100

0100011

- Es ist im allgemeinen nicht a priori abschätzbar, ob man schon genug Information erhalten hat, ob also der Lernprozeß schon beendet werden kann oder nicht.

01011101010001010110111010111000100011

- Wenn man schon nichts über die "globale" Güte von Hypothesen weiß, dann wird deren "lokale" Güte um so wichtiger.

01011101010001010110111010111000100011

010111010100010101101110101110

00100011



Algorithmisches Lernen

Algorithmic Learning



Gliederung

1. Motivation und Einführung
 - Einführende Beispiele
 - Szenario der Induktiven Inferenz
 - Charakteristika natürlicher Lernverfahren
2. Lernen formaler Sprachen
3. Lernen rekursiver Funktionen
4. Ausgewählte Anwendungen
5. Zusammenfassung



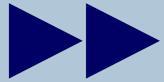
01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Example 1

01011101010001010110111
010111000100011

Angenommen, über eine unbekannte Funktion f wird Information vorgelegt wie z.B.

$$f(0) = 0, f(1) = 1, \dots$$

Um welche Funktion handelt es sich?

$$f(2) = 2, \dots$$

Und jetzt ... ?

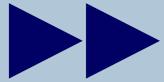
$$f(3) = 720!, \dots$$

Na, noch eine Idee ... ?



Algorithmisches Lernen

Algorithmic Learning



Analysis 1.1

01011101010001010110111
010111000100011

Untersuchen wir das Beispiel

$$f(0) = 0, f(1) = 1, f(2) = 2, f(3) = 720!, \dots$$

mit dem Ziel, Regularitäten aufzudecken, also "zu lernen".

Ein Ansatz könnte wie folgt aussehen:

$$f(3) = 720! = (6!)! = ((3!)!)! = 3!!!$$

Gehen wir weiter in's Detail ...



Algorithmisches Lernen

Algorithmic Learning



Analysis 1.2

01011101010001010110111
010111000100011

Gegeben Information der Form

$$f(x) = y,$$

kann man die Suche nach einer Darstellung

$$f(3) = 720! = (6!)! = ((3!)!)! = 3!!!$$

auffassen als "expressing y in terms of x "
(SUMMERS, 1975 & 1977).



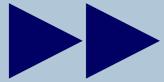
010111010100010101101110101110
010111010100011

Hinweis: Entscheidend sind dabei die verfügbaren bzw. zugelassenen Ausdrucksmittel.



Algorithmisches Lernen

Algorithmic Learning



Analysis 1.3

01011101010001010110111
010111000100011

Wie drückt man das durch Terme aus?!

Gegeben sind

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

$$f(x_1) = y_1, \dots, f(x_n) = y_n .$$

Gesucht sind ein Term $t(x)$ und Substitutionen

0101110101000101011011101011100011000100011
01011101010001010110111010111000 $\sigma_1, \dots, \sigma_n$

mit der Eigenschaft (für $i=1, \dots, n$):

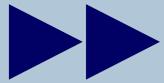
01011101010001010110111010111000100011

$$\sigma_i t = y_i$$



Algorithmisches Lernen

Algorithmic Learning



Analysis 1.4

01011101010001010110111
010111000100011

Wenn man aus den Beispielen

$$f(x_1) = y_1, \dots, f(x_n) = y_n.$$

einen Term $t(x)$ mit Substitutionen $\sigma_1, \dots, \sigma_n$

und der Bedingung $\sigma_i t = y_i$

gefunden hat, dann hat man t "gelernt", ...

01011101010001010110111010111000100011



... was aber nur eine Hypothese sein kann.

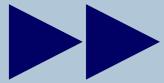
010111010100010111010111000100011

00100011



Algorithmisches Lernen

Algorithmic Learning



Analysis 1.5

01011101010001010110111
010111000100011

Ein - eventuell vorläufiges - Ergebnis besteht darin zu sagen, man hätte aus den Beispielen

$f(0) = 0$, $f(1) = 1$, $f(2) = 2$, $f(3) = 720!$, ...

die Beschreibung

$$g(x,0) = x$$

$$g(x,n+1) = g(x,n)!$$

$$f(x) = g(x,x)$$

der Funktion f "gelernt".



Algorithmisches Lernen

Algorithmic Learning



Beispiel 3.1

01011101010001010110111
010111000100011

Ein Beispieltext (Zeichenkette, Datensatz, ...):

@article (Angluin/Smith83,
author = {Angluin, Dana and Smith, Carl H.},
title = {A Survey of Inductive Inference:
Theory and Methods},
journal = {Computing Surveys},
year = {1983},
volume = {15},
pages = {237--269})



Algorithmisches Lernen

Algorithmic Learning



Beispiel 3.2

01011101010001010110111
010111000100011

... und ein weiteres Beispiel für dasselbe Konzept:

@article (Klette/Wiehagen80,
author = {Klette, Reinhard and Wiehagen, Rolf},
title = {Research in the Theory of Inductive
Inference by {GDR} Mathematicians
- A Survey},
journal = {Information Sciences},
year = {1980},
volume = {22},
pages = {149--169})



Algorithmisches Lernen

Algorithmic Learning

Beispiel 3.3

01011101010001010110111

010111000100011

Was für ein Konzept soll gelernt werden ?!

• basierend auf Zeichenketten

- eine CHOMSKY-Grammatik

- Text-Muster

...

0101110101000101011010111000100011

• basierend auf Relationen (relationalen DB)

- Datenabhängigkeiten (welcher Art?)

...

Algorithmisches Lernen

Algorithmic Learning



Analysis 3.1

01011101010001010110111
010111000100011

Ein erster intuitiver (naiver ...?!) Ansatz für ein Muster
(engl.: Pattern) p_1 :

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

@article (x_1/x_2x_3 ,
author = { x_1, x_4 and x_2, x_5 },
title = { x_6 },
journal = { x_7 },
year = {19 x_3 },
volume = { x_8 },
pages = { x_9 -- x_{10} })



01011101010001010110111
0100011

010111010100010101101110101110
0100011



Algorithmisches Lernen

Algorithmic Learning



Criticism of the Naïve Approach

01011101010001010110111

010111000100011

Die Wahl von p_1 beruht auf einer Reihe von impliziten Annahmen, sozusagen auf "Vorurteilen". Wieso?!

0101110101000101011011101011100

0100011 0101110101000101011011101011100

0100011

- Zum Beispiel ist " $x_1/x_2 \wedge x_3$ " spezifischer als " $x_1/x_2 x_3$ ".

- Sowohl $\{x_{61}$ of Inductive Inference $x_{62}\}$

- als auch $\{x_{61}$ Theory $x_{62}\}$

- ist viel genauer als

$\{x_6\}$



Algorithmisches Lernen

Algorithmic Learning

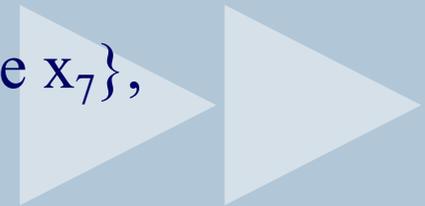


Analysis 3.2

01011101010001010110111
010111000100011

Eine Hypothese, die weniger intuitiv, aber akurater die Beispiele widerspiegelt, ist p_2 :

@article (x_1/x_2 x_3 ,
author = { x_1, x_4 and x_2, x_5 },
title = { x_6 of Inductive Inference x_7 },
journal = { $x_8 x_9$ },
year = {198 x_3 },
volume = { x_{10} },
pages = { x_{11} -- x_{12} 69})



Algorithmisches Lernen

Algorithmic Learning



Analysis 3.3

01011101010001010110111
010111000100011

Eine konkurrierende Hypothese, unvergleichbar zu p_2 ,
ist p_3 :

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

@article (x_1/x_2 x_3 ,
author = { x_1, x_4 and x_2, x_5 },
title = { x_6 Theory x_7 },
journal = { $x_8 x_9$ },
year = {198 x_3 },
volume = { x_{10} },
pages = { x_{11} -- x_{12} 69})



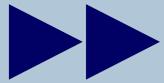
01011101010001010110111
0100011

010111010100010101101110101110
0100011



Algorithmisches Lernen

Algorithmic Learning



Discussion

01011101010001010110111
010111000100011

Sowohl die Hypothese p_1 als auch die Hypothese p_2 charakterisiert die zwei Beispiele nicht so gut wie möglich.

0100011 0101110101000101011011101011100
0100011

- Zum Beispiel ist "journal = $\{x_{81}ox_{82}tix_{83} Sx_{91}ex_{92}s\}$ " spezifischer als "journal = $\{x_8 x_9\}$ ".

- Die Teilstruktur "title = $\{x_{61}A Survey x_{62}\}$ " ist sinnvoll, falls leere Substitutionen zugelassen werden.

01011101010001010110111010111000100011

00100011

1110



Algorithmisches Lernen

Algorithmic Learning



Short Digression

01011101010001010110111

010111000100011

Wie vernünftig sind "vernünftige" Ideen, d.h. konkret:

Um wieviel wächst die Ausdruckskraft von Patterns, wenn man leere Substitutionen zulässt?!

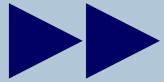
Hat diese Verallgemeinerung des Ansatzes irgendeine Auswirkung auf die Problematik, solche Patterns zu lernen?!

Allgemeiner: Wie weit reicht der "gesunde Menschenverstand", und von welcher Stelle an wird Wissenschaft unverzichtbar?



Algorithmisches Lernen

Algorithmic Learning



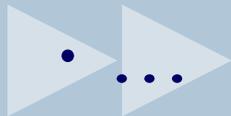
Postulates of Naturalness

01011101010001010110111
010111000100011

Übersicht: In dieser Vorlesung können nur ein paar Beispiele behandelt werden. Eine recht umfassende Darstellung findet man schon im Artikel JANTKE/BEICK, "Combining Postulates of Naturalness in Inductive Inference".

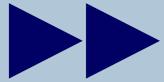
Wir diskutieren:

- Consistency
- Conservative Mind Changes
- Finiteness
- Incremental Hypothesis Construction



Algorithmisches Lernen

Algorithmic Learning



Consistency

01011101010001010110111
010111000100011

Ein induktives Lernverfahren arbeitet konsistent genau dann, wenn jede durch dieses Verfahren generierte Hypothese die Eigenschaft hat, daß sie die Information, auf der sie basiert, korrekt widerspiegelt.

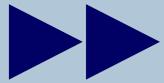
01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Ein Beispiel zur Illustration

Zu lernen sind Textpatterns (Strings aus Terminalen und Variablen).

Es werden Beispiele w_1, w_2, w_3, \dots vorgelegt.

Wir betrachten folgendes Lernverfahren:

$$M(w_1) := w_1$$

$$(1) \quad M(h_n, w_{n+1}) := \text{merge}(h_n, w_{n+1}) \text{ , falls } |w_{n+1}| = |h_n|$$

$$(2) \quad M(h_n, w_{n+1}) := h_n \text{ , falls } |w_{n+1}| > |h_n|$$

$$(3) \quad M(h_n, w_{n+1}) := w_{n+1} \text{ , falls } |w_{n+1}| < |h_n|$$

merge geht von links nach rechts durch die Zeichenketten und setzt neue Variablen in der Hypothese, falls nötig. (genaue Def. Später)

Algorithmisches Lernen

Algorithmic Learning



Ein Beispiel zur Illustration

aabab

01011101010001010110111

Hypothese: aabab

ababb

Hypothese: aXYXb

0101110101000101011011101011100

0100011 0101110101000101011011101011100
0100011

aaa

Hypothese: aaa

baaa

Hypothese: aaa

01011101010001010110111010111000100011

01011101010001010110111010111000100011

bba

Hypothese: XXa

010111010100010101101110101110

00100011

01011101010001010110111010111000100011

aab

Hypothese: XXY



Algorithmisches Lernen

Algorithmic Learning



Discussing Details

01011101010001010110111
010111000100011

Eine Lehre

Wir neigen bei der Behandlung komplizierter Fragen dazu, implizit bestimmte Annahmen zu machen, welche uns das Nachdenken erleichtern oder "dem gesunden Menschenverstand entsprechen", die aber unzulässig sein können ...



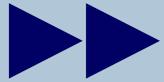
01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Conservativeness

01011101010001010110111
010111000100011

Ein induktives Lernverfahren arbeitet konservativ genau dann, wenn jeder durch dieses Verfahren vollzogene Hypothesenwechsel dadurch motiviert ist, daß die neu vorgelegte Information mit der bisherigen Hypothese nicht konsistent war.



01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Incremental Hypothesis Construction

01011101010001010110111
010111000100011

Ein induktives Lernverfahren arbeitet inkrementell genau dann, wenn es nur das jeweils vorgelegte Beispiel in die letzte Hypothese einarbeitet.

01011101010001010110111010111000100011
01011101010001010110111010111000100011



01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Finiteness

01011101010001010110111
010111000100011

Ein induktives Lernverfahren arbeitet finit genau dann, wenn es ein zugehöriges Entscheidungsverfahren gibt, das zu jedem möglichen vorgelegten Informationsangebot feststellt, ob die generierte Hypothese bereits richtig ist und demzufolge endgültig beibehalten werden kann.

01011101010001010110111010111000100011

010111010100010101101110101110
00100011



Algorithmisches Lernen

Algorithmic Learning



Observation ... Theorem

Frage: Wie ist das Verhältnis der Anforderungen zueinander?

0101110101000101011011101011100
0100011 0101110101000101011011101011100
0100011

Beispielantwort: 010001010110111010111000100011

Wenn ein Lernverfahren finit arbeitet, dann kann es in ein konsistentes Verfahren transformiert werden.

Die Umkehrung gilt nicht.

