

Lernalgorithmen

SoSe 2008 in Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

Lernalgorithmen

Gesamtübersicht

0. Einführung
1. Identifikation (aus positiven Beispielen)
2. Zur Identifikation regulärer Sprachen, mit XML-Anwendung
3. HMM — Hidden Markov Models
4. Lernen mittels Anfragen & zur Roboterorientierung
5. Lernen mit negativen Beispielen
6. PAC-Lernen

Organisatorisches

Vorlesung MO 8.15-9.45, F55

ab der zweiten Woche: MO 8.25-9.55

Publikum Hauptdiplomstudierende und Masterstudierende

Übungsbetrieb in Form von einer “Großen Übungsgruppe”

MI 8.25-9.55, F55

BEGINN: in der zweiten Semesterwoche

Dozentensprechstunde DO, 13-14 in meinem Büro H 410 (4. Stock)

Mitarbeitersprechstunde (Stefan Gulan) DO 13-14 H 413

Scheinkriterien Es wird ein benoteter Schein vergeben nach einer mündlichen Prüfung.

Probleme ? Fragen ?

Klären Sie bitte Schwierigkeiten mit Vorlesungen oder Übungen möglichst **umgehend** in den zur Verfügung gestellten Sprechzeiten.

Wir helfen Ihnen gerne!

... wir sind aber keine Hellseher, die Ihnen Ihre Schwierigkeiten an der Nasenspitze ansehen...

Einführung 1: Einordnung

Was ist Lernen?

Eine Aufgabe für Psychologen, Informatiker, Biologen, Philosophen, ...

Eine naturwissenschaftliche Sicht

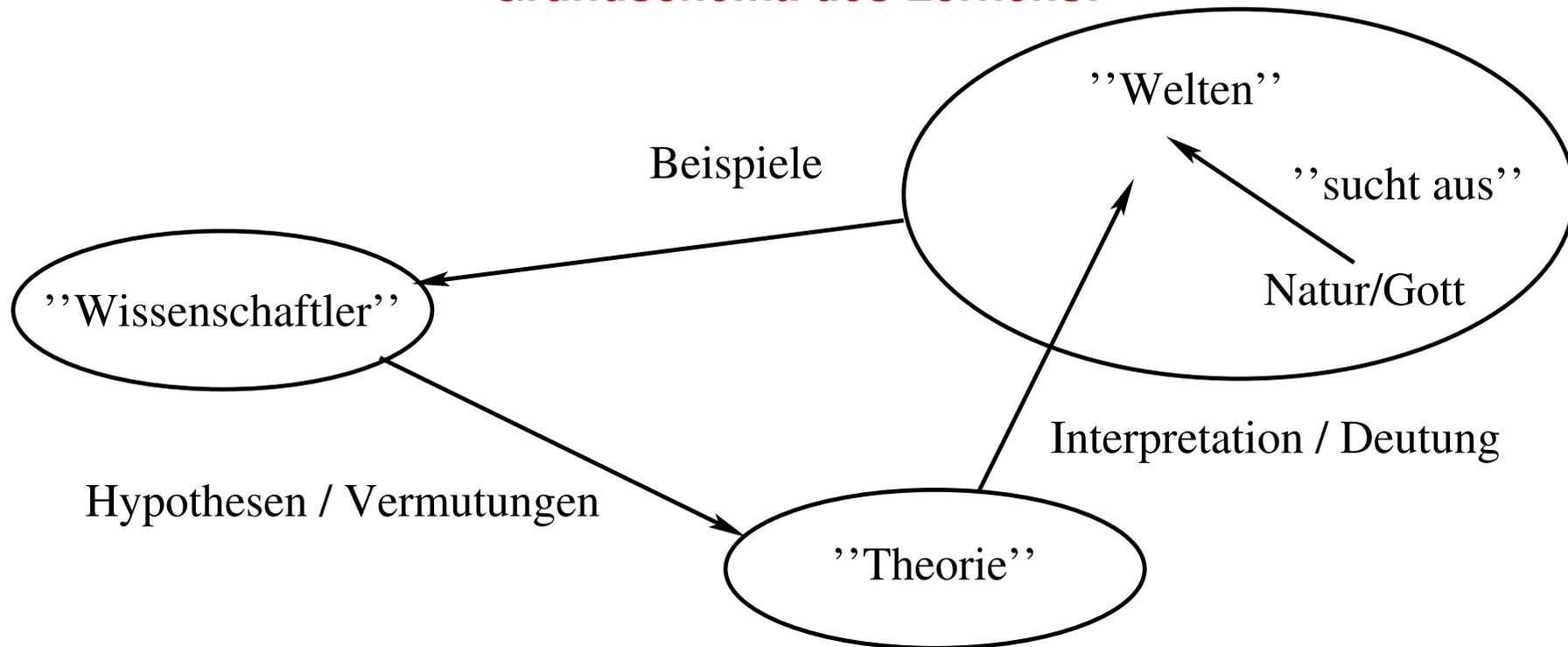
Induktion versus Deduktion

In der Physik gibt es zwei unterschiedliche, sich ergänzenden Vorgehensweisen:

- (A) die Herleitung von allgemeinen Gesetzmäßigkeiten aus durch Beobachtung oder Experiment gewonnenen Beispieldaten. [Induktion]

- (B) das logische Schlußfolgern von Gesetzmäßigkeiten aus postulierten Grundgesetzen (Axiomen). [Deduktion]

Grundschema des Lernens:



Konzeptlernen als **informatische Aufgabe**

Formalisiert sehen wir als Aufgabe des (maschinellen) Lernens, Konzepte zu ergründen, wobei wir uns auf Konzeptklassen $\mathcal{C} \subseteq 2^{\mathbb{N}}$ oder $\subseteq \mathbb{N}^{\mathbb{N}}$ einschränken:

Beispiel 1: $\mathcal{C}_1 = \{M_n \subseteq \mathbb{N} \mid n \in \mathbb{N}, M_n = \mathbb{N} \setminus \{n\}\}$.

Beispiel 2: $\mathcal{C}_2 = \mathcal{C}_1 \cup \{\mathbb{N}\}$.

Beispiel 3: $\mathcal{C}_3 = \{f : \mathbb{N} \rightarrow \mathbb{N} \mid f = \text{id} \vee$
 $(\exists n \in \mathbb{N} : f|_{\{0, \dots, n\}} = \text{id} \wedge \forall m > n : f(m) = m + 1)\}$.

Bemerkung: $\mathcal{C}_3 \hat{=} \mathcal{C}_2$ mit absteigender Sortierung.

Präsentation des Konzepts

- (i) Eine Turingmaschine zählt das gewählte Konzept $C \in \mathcal{C}$ auf, bei Mengen eventuell mit Wiederholungen; in Beispiel 3 in der Reihenfolge der Funktionswerte $f(0), f(1), f(2), \dots$; bei Konzepten $C \subseteq \mathbb{N}$ (oder $C \subseteq \Sigma^*$) auch paarweise $(c, t) \in C \times \{0, 1\}$, wobei $c \in C \Leftrightarrow (c, 1)$ aufgezählt wird und $c \notin C \Leftrightarrow (c, 0)$ aufgezählt wird.
- (ii) Der Wissenschaftler kann sich gezielt Informationen beschaffen (z.B. durch Experimente). Dies kann zur Definition von 2-Personenspielen führen, wobei die Frage der Lernbarkeit zur Frage nach einer Gewinnstrategie wird.

Klassifikation und Adaption

Anwendungenrelevante Aufgabe: Eingaben automatisch zu *klassifizieren*.

Dazu nützliche Herangehensweise:

Erlerne Klassifikationsschema in oben skizzierter Weise aus Beispielen;

~> *Lernphase*

dieses wird — quasi in einer zweiten Phase — als *Klassifikator* benutzt.

~> *Anwendungsphase*

Typische Vertreter dieses Zwei-Phasen-Schemas sind *neurale Netze*.

In gewissem Sinne fallen hierunter auch *adaptive Algorithmen*.

Bei ihnen wechseln sich Lern- und Anwendungsphase stets ab, wodurch es zu *Rückkopplungen* (mit evtl. Schwingungen) kommt.

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?!

~> Betrachtung
unterschiedlicher
Lernszenarien
insbesondere:

- unterschiedliche Rollen von
Lehrer und Schüler,
- soziale Gefüge (Protokolle)

Autonomes System: Antwort auf sich wandelnde Umgebung



~> Notwendigkeit einer formalisierten Sicht des Lernens

“Intelligenztest” spielerische Annäherung ans Sprachenlernen

Frage: Welche Sprache ist (genau) beschrieben durch:

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

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

“Intelligenztest”

Welche Sprache ist (genau) beschrieben durch:

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

Antwort: ab^+ (?)

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

“Intelligenztest”

Welche Sprache ist (genau) beschrieben durch:

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

andere Antwort: $\{ab^{10n+k} \mid n \geq 0, k = 1, 2, 3, 4, 5\}$ (?)

Also: Das “nächste” Wort wäre $abbbbbb = ab^{11}$.

Welche Lösung ist “richtig” ?

Einordnung der Vorlesung: Zusammenhänge zu anderen Gebieten

- Statistik, insbesondere Parameterschätzung (\rightarrow Mathematik)
- Numerik, insbesondere Interpolation (\rightarrow Mathematik)
- adaptive Algorithmen
- Kryptographie, Datenkompression
- andere Teilaspekte maschinellen Lernens, z.B.: neurale Netze, fallbasiertes Schließen

Einordnung der Vorlesung: “Hilfsmittel” theoretischer Untersuchungen:

- Rekursionstheorie / Berechenbarkeit
- Komplexitätstheorie / Algorithmik
- Automatentheorie / Formale Sprachen
- (Kolmogorov-Komplexität / Informationstheorie)

Einführung 2: Literatur

Leider gibt es “das Buch” zu dem Thema (noch) nicht.

Colin de la Higuera ist dabei, so ein Buch zu schreiben; möglicherweise ist es fertig, falls Sie sich auf Prüfungen vorbereiten.

Bücher und Skripten

S. Jain, D. Osherson, J. S. Royer, A. Sharma. Systems That Learn. 2. Auflage, MIT Press, 1999.

M. J. Kearns und U. V. Vazirani. An Introduction to Computational Learning Theory. 2. Auflage, MIT Press, 1997.

P. Langley. Elements of Machine Learning. Morgan Kaufmann Publishers, 1996.

T. M. Mitchell. Machine Learning. McGraw-Hill, 1997.

B. K. Natarajan. Machine Learning; a Theoretical Approach. Morgan Kaufmann Publishers, 1991.

K. Morik. Maschinelles Lernen; Skript zur Vorlesung, 12. April 1999; Universität Dortmund.

<http://www-ai.cs.uni-dortmund.de/LEHRE/VORLESUNGEN/MLRN/mlrn.h>

F. Stephan. Einführung in die Lerntheorie.

<http://math.uni-heidelberg.de/logic/fstephan/fstephan.html>

F. Wysotski. Skript zur Vorlesung Maschinelles Lernen, Wintersemester 99/00; Institut für Angewandte Informatik der TU Berlin.

<http://ki.cs.tu-berlin.de/lehre/ml.html>

Lern-Konferenzen

- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Conference on Learning Theory (CoLT)
- European Conference on Learning Theory (EuroCoLT)
- Algorithmic Learning Theory (ALT)

- International Conference on Grammatical Inference (ICGI)
- Principles of Data Mining and Knowledge Discovery (PKDD)
- Discovery Science (DS)
- International Conference on Pattern Recognition (ICPR)
- Syntactical and Structural Pattern Recognition (SSPR)
- ...