

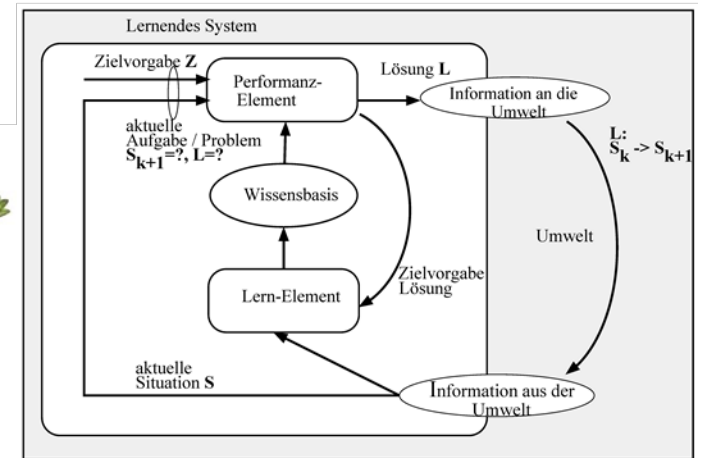
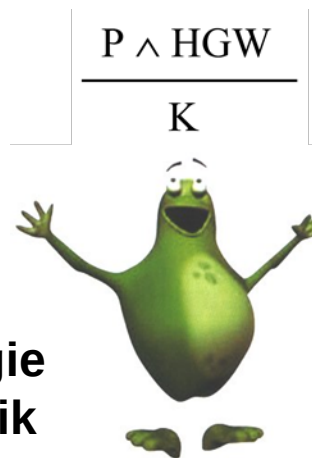
Maschinelle Intelligenz - Verfahren und Innovationspotentiale

Meeting Point Internationalisierung der ClusterAgentur Baden-Württemberg
Künstliche Intelligenz in der Anwendung – Potenziale und Grenzen der Kooperation mit Frankreich und anderen europäischen Ländern
Mittwoch, 22. Januar 2020, Haus der Wirtschaft

Karlsruhe Institute of Technology - Institute for Applied Computer Science (IAI)

Dr. Hubert B. Keller
Leiter Fachgebiet Advanced
Automation Technologies

Karlsruher Institut für Technologie
Institut für Angewandte Informatik



Inhalt

- Vorbemerkungen
- Aristoteles - Erfinder der künstlichen Intelligenz
- Schlüsse ziehen beim Menschen
- Schlüsse ziehen mit Logik
- Neurocomputing
- Maschinen lernen
- Reiz und Reaktion Lernen
- Prämissen und Konsequenzen intelligente Maschinen

Zur Person

- Leiter Fachgebiet Advanced Automation Technologies – A2T
- Leiter Arbeitsgruppe Reliable, safe & secure Software and Systems – RS4
- Forschungsthemen
 - Maschinelle Intelligenz, Sichere Software, Intelligente Sensorik
 - Energiesystem der Zukunft, industrielle Prozesse
 - Dozent Technische Informatik, Echtzeitsysteme, Wissensbasierte Systeme
 - Autor mehrerer Bücher
 - Maschinelle Intelligenz
 - Safety – technische Sicherheit
 - Entwicklung von Echtzeitsystemen
 - Kastel „Sicherheit für Energiesysteme“



Kompetenzzentrum
für angewandte
Sicherheitstechnologie



Komplex und kompliziert

■ Kompliziert (beherrschbar)

- Viele verschiedenartige Elemente
- **Einfach**: ein Spieß mit nur Fleisch
- **Kompliziert**: ein Spieß mit jeweils nur 1 Stück unterschiedlicher Dinge ($n \gg 1$: 1 Stück Fleisch, 1 Stück Paprika, 1 Stück Zwiebel, 1 Stück ...)
- Nur sehr geringe Bezüge zwischen den Elementen (Nachbar links, Nachbar rechts)

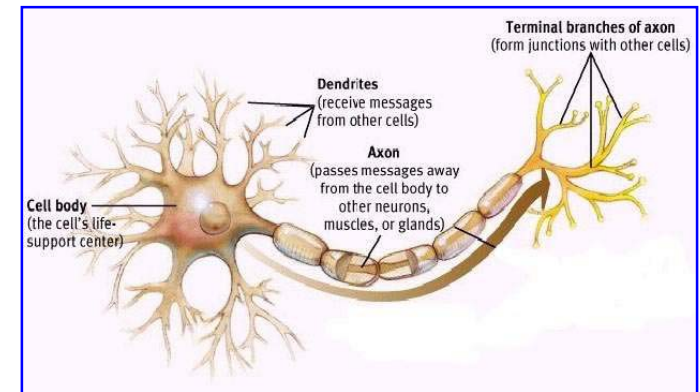
■ Komplex (schwierig, unterschätzt)

- Eine **große Anzahl** ($n \gg 1$) von **gleichartigen** Elementen (Halma Stäbchen)
- Eine sehr **große Anzahl** ($n \gg 1$) von **gegenseitigen Beziehungen**
- Komplex: menschliches Gehirn mit 100 Milliarden Neuronen
- Jedes Neuron ist mit 1.000 anderen Neuronen verbunden
- Alle Neuronen sind über 3 Stufen mit jedem anderen verbunden
- Funktion wird durch Wechselwirkung hervor gebracht

...Komplex und kompliziert

Menschliches Gehirn (ZNS)

- Gehirn enthält etwa 100 Milliarden Neuronen
- Über drei Stufen verschaltet:
 $1 \times 10.000 \times 10.000 \times 10.000$
- Sehr wenig erforscht!

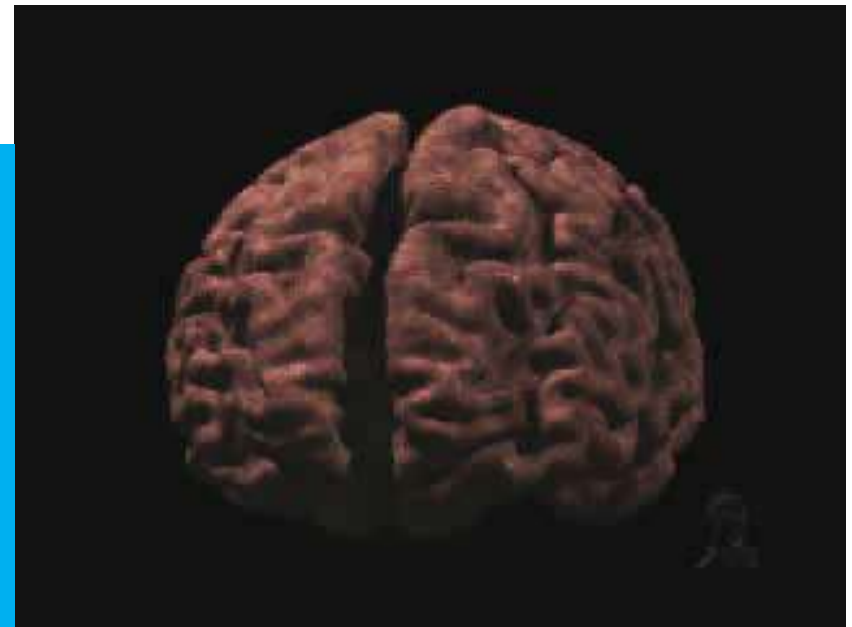


Beispiel Genetik

Code entschlüsselt, Bedeutung unklar
Code steuert sich selbst in Abhängigkeit von Zustand, sich selbst und der Umgebung einschließlich Vererbung von Umgebungsanpassungen

Autonomes Fahren

... später



Artificial Intelligence (AI)

- Intelligenz (welche?) → Machine Intelligence (MI)
 - Symbolische AI – starke AI
 - Wissensverarbeitung (Schlussfolgern deduktiv, induktiv)
 - Logik – klassisch - nichtmonoton
 - Regeln – klassisch - Fuzzy
 - Semantische Netze – Frames - Scripte
 - ...
 - Lernen
 - Algebraische Verfahren – Bacon, Fahrenheit
 - Induktionsbäume (TDIDT), ILP, ...
 - ...
 - Subsymbolische AI – Konnektionismus (Lernmatrix → Steinbuch 1961)
 - Überwacht – Backpropagation, Deep Learning, ...
 - Unüberwacht – Self Organizing Feature Maps (Kohonen)
 - Reinforcement – Bonus/Malus Lernen
 - ...

Aristoteles – Erfinder der künstlichen Intelligenz

Assoziationen-Modell

- Konzept der raum-zeitlichen Nähe
Elemente, welche raum-zeitlich eng zueinander in Bezug stehen werden miteinander assoziiert
Tisch+Stuhl, Schalter ein → Lampe ein
- Konzept der Ähnlichkeit
Elemente, die einander ähnlich sind ebenfalls
Smart+Ferrari (→ Ladekapazität)
- Konzept des Kontrastes (Gegensatz)
aufgrund von direkten Gegensätzen ergeben sich ebenfalls Assoziationen
Tag+Nacht, Sonnenschein+Regen



Aristoteles, 384 v. Chr. in Stageira; † 322 v. Chr. in Chalkis

Schlussfolgern beim Menschen (1)

Logisches Schließen (Modus Ponens)

- Aussage 1 (Dieses ist ein Ball.)
- Aussage 2 (Alle Bälle sind eckig)
- $\Rightarrow$ Schluss (Dieser Ball ist eckig.).

- Schluss ist logisch korrekt - aber inhaltlich falsch.

Hintergrundwissen:

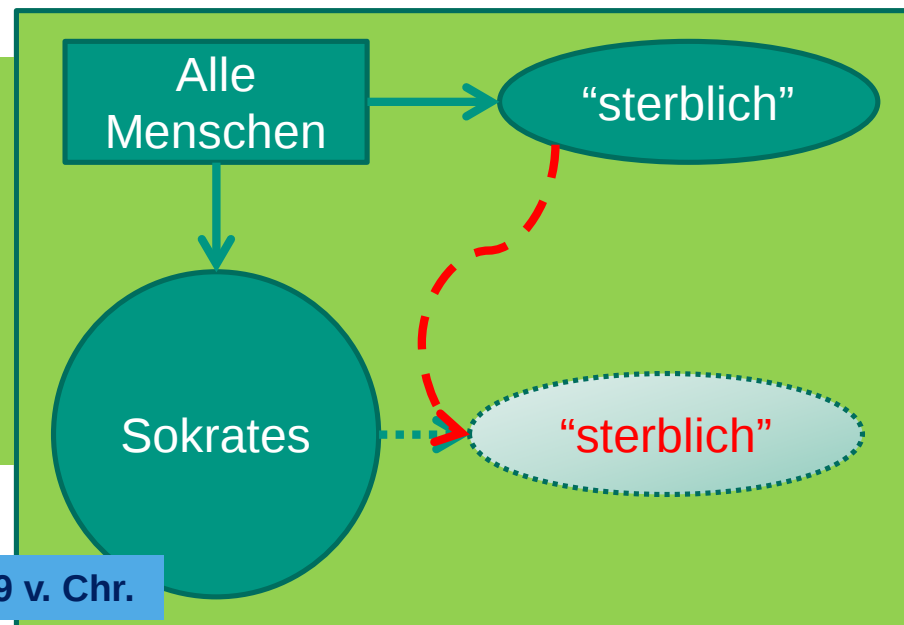
■ Alle Menschen sind sterblich

Prämisse:

■ Sokrates ist ein Mensch

Folgerung:

■ Sokrates ist sterblich

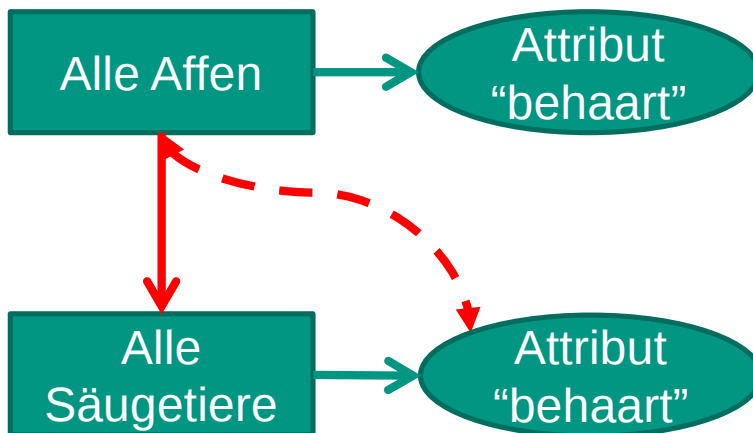


Sokrates, 469 v. Chr. in Alopeke, Athen; † 399 v. Chr.

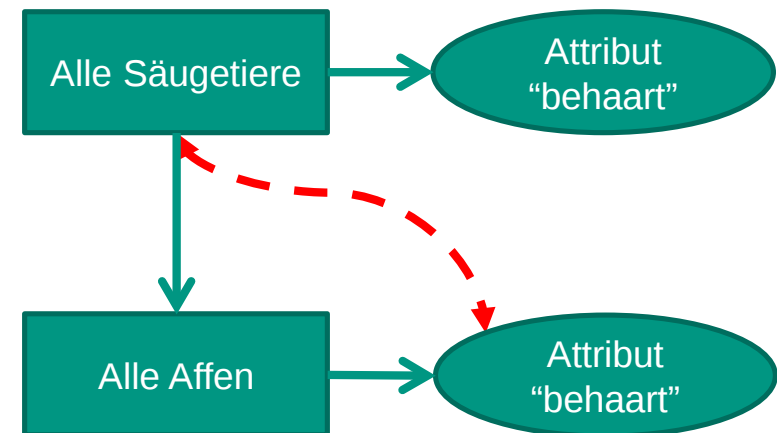
Schlussfolgern beim Menschen (2)

Teddybären, Reifen, ...

- Alle Säugetiere sind behaart
 - Alle Affen sind behaart
- Und jetzt?



Alle Säugetiere sind Affen!
 “Menschen sind eben auch Affen,
 wussten wir ja!”



Menge aller Säuger ist größer
 als Menge aller Affen
 → Alle Affen sind Säugetiere!

Konstruktive Induktion

Schlussfolgern mit Logik (1)

Kommissar Logik hat 3 Verdächtige:

- A, B, C
- Falls A schuldig und B unschuldig ist, so ist C schuldig.
- C arbeitet niemals allein.
- A arbeitet niemals mit C.
- Nur A, B, C kommen als Täter in Frage.

- Wer war es?

George Boole

(* 2. November 1815 in Lincoln, England;
† 8. Dezember 1864 in Ballintemple,
in der Grafschaft Cork, Irland)

Friedrich Ludwig Gottlob Frege

(* 8. November 1848 in Wismar;
† 26. Juli 1925 in Bad Kleinen)

Umformung in Logik

- 1: $A \wedge \neg B \rightarrow C$
- 2: $C \rightarrow A \vee B$
- 3: $\neg (A \wedge C)$
- 4: $A \vee B \vee C$

Umformung für die Maschine

Aus 1: $\neg A \vee B \vee C$

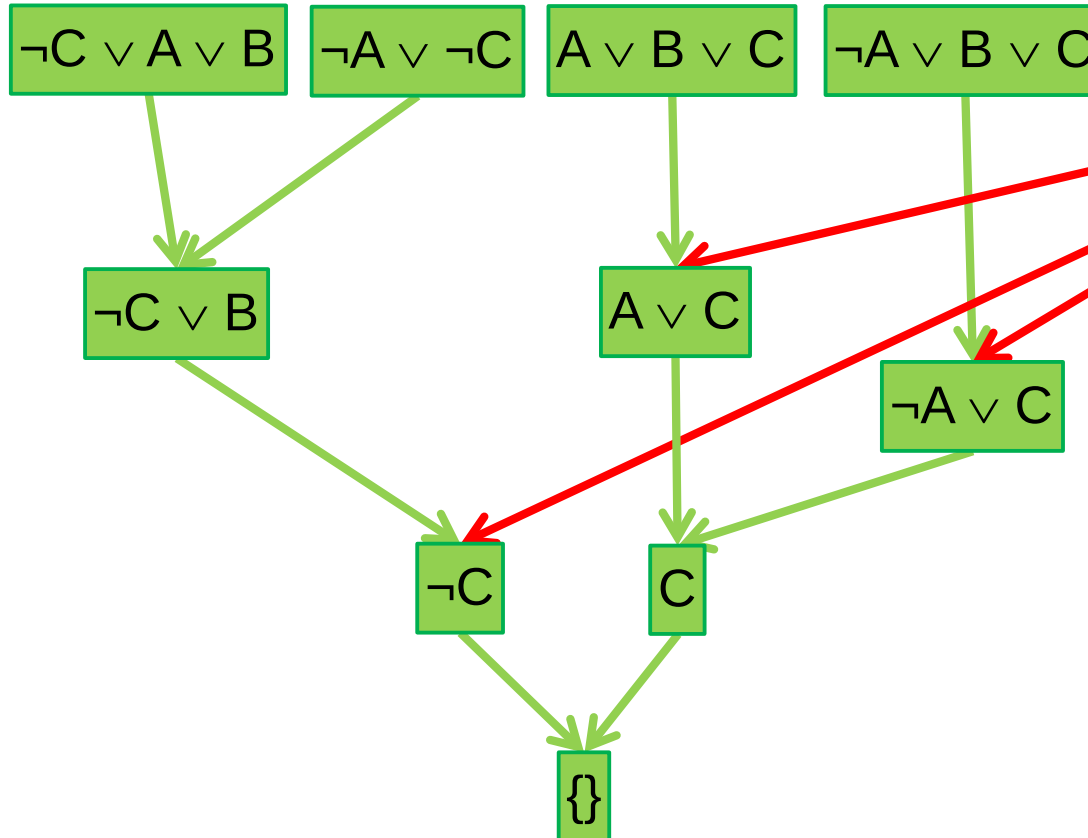
Aus 2: $\neg C \vee A \vee B$

Aus 3: $\neg A \vee \neg C$

Aus 4: $A \vee B \vee C$

Schlussfolgern mit Logik (2) - Frage: War B der Täter?

M: Menge aller Grundaussagen



$\neg B$: Die Gegenbehauptung!

$\neg B$

$\{\}$: Leere Menge, da B wahr in M gegen die Behauptung $\neg B$
 $\rightarrow$ Gegenbehauptung ist falsch!

B muss wahr sein!

Schlussfolgern mit Logik (3)

Resolutionsverfahren in der Logik: Beweis durch Widerspruch!

- Frage: War B der Täter?
- Ansatz für den Computer: Behauptung, er war es nicht! $\rightarrow \neg B$

$(M \cup \neg B)$

5. $\neg B$

6. 1+5: $(\neg A \vee B \vee C) \wedge (\neg B) \vdash \neg A \vee C$

7. 4+5: $(A \vee B \vee C) \wedge (\neg B) \vdash A \vee C$

8. 2+3: $(\neg C \vee A \vee B) \wedge (\neg A \vee \neg C) \vdash \neg C \vee B$

9. 5+8: $(\neg B) \wedge (\neg C \vee B) \vdash \neg C$

10. 6+7: $(\neg A \vee C) \wedge (A \vee C) \vdash C$

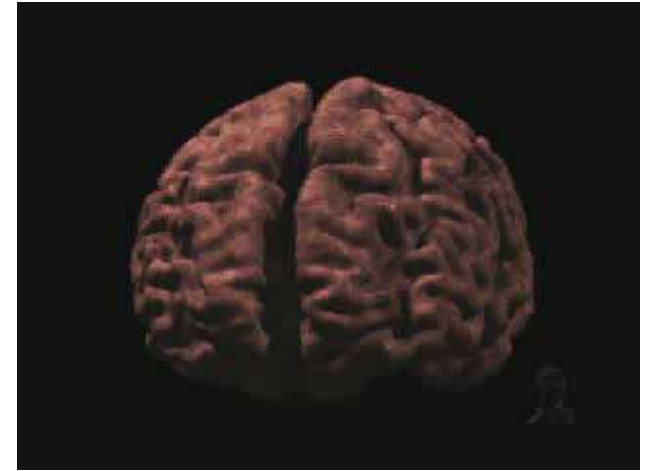
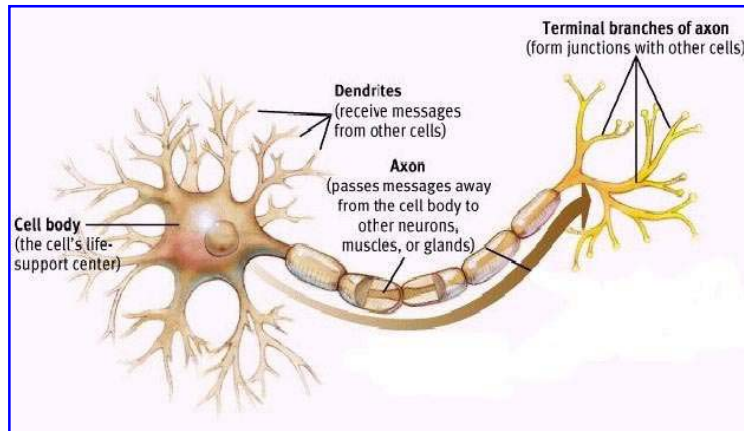
11. 9+10: $(\neg C) \wedge (C) \vdash \{\}$ (Widerspruch!)

John Alan Robinson
(* 1928 in Yorkshire, Großbritannien)

Macht die
Maschine,
nicht wir!

B war es also! (Dies kann der Computer berechnen)

Neurocomputing (1) – Nervensysteme im Computer



$$y_i = x_i$$

$$y_j = f_{\tilde{O}}(net_j - \Theta_j) = \left(1 + \exp(-a \cdot (net_j - \Theta_j))\right)^{-1}$$

$$net_j - \Theta_j = \sum_{i=0}^n y_i \cdot w_{ij}$$

$$y_r = f_{\tilde{O}}(net_r - \Theta_r) = \left(1 + \exp(-a \cdot (net_r - \Theta_r))\right)^{-1}$$

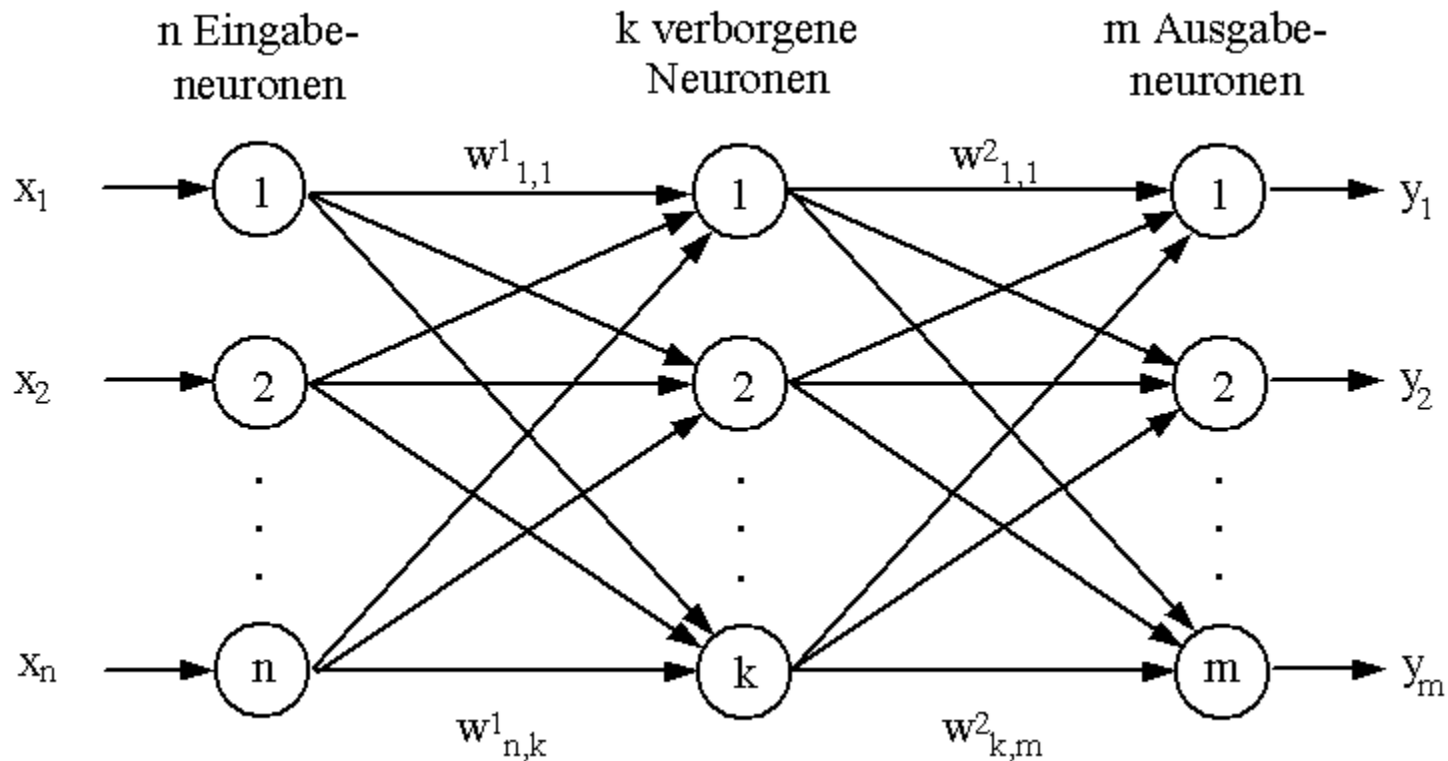
$$net_r - \Theta_r = \sum_{j=0}^k y_j \cdot w_{jr}$$

$$E = \frac{1}{2} \sum_{r=1}^m (e_r)^2 = \frac{1}{2} \sum_{r=1}^m (d_r - y_r)^2$$

Überwachtes Lernen:
Eingang auf Ausgang abbilden
Fehler E berechnen
Über Ableitung (Gradient dE/dy)
mit -1^* anpassen
(R-Prop Verfahren)

Neurocomputing (2)

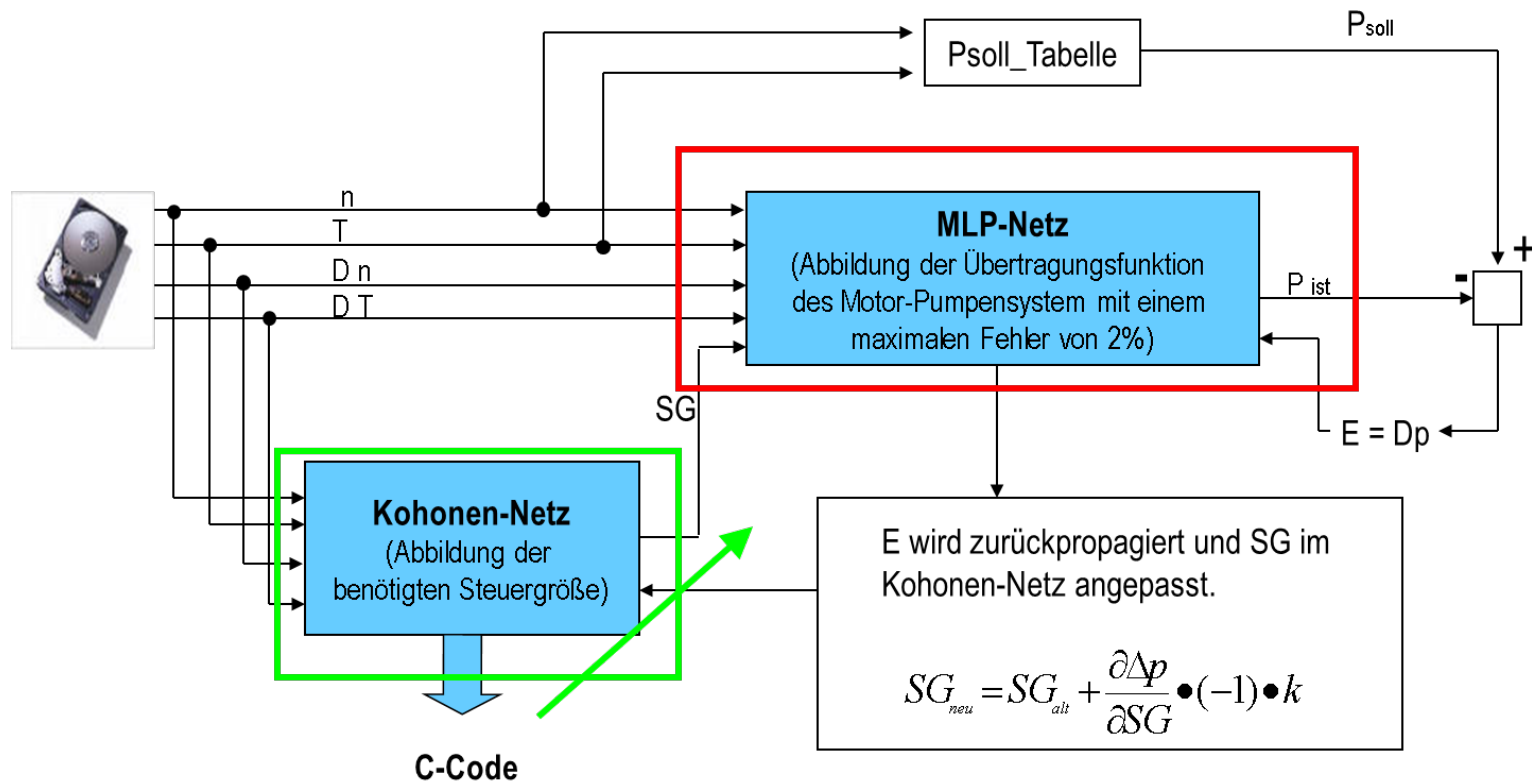
- Optimierung der Ölversorgung eines Hochleistungs-Sportwagens



Neurocomputing (2)

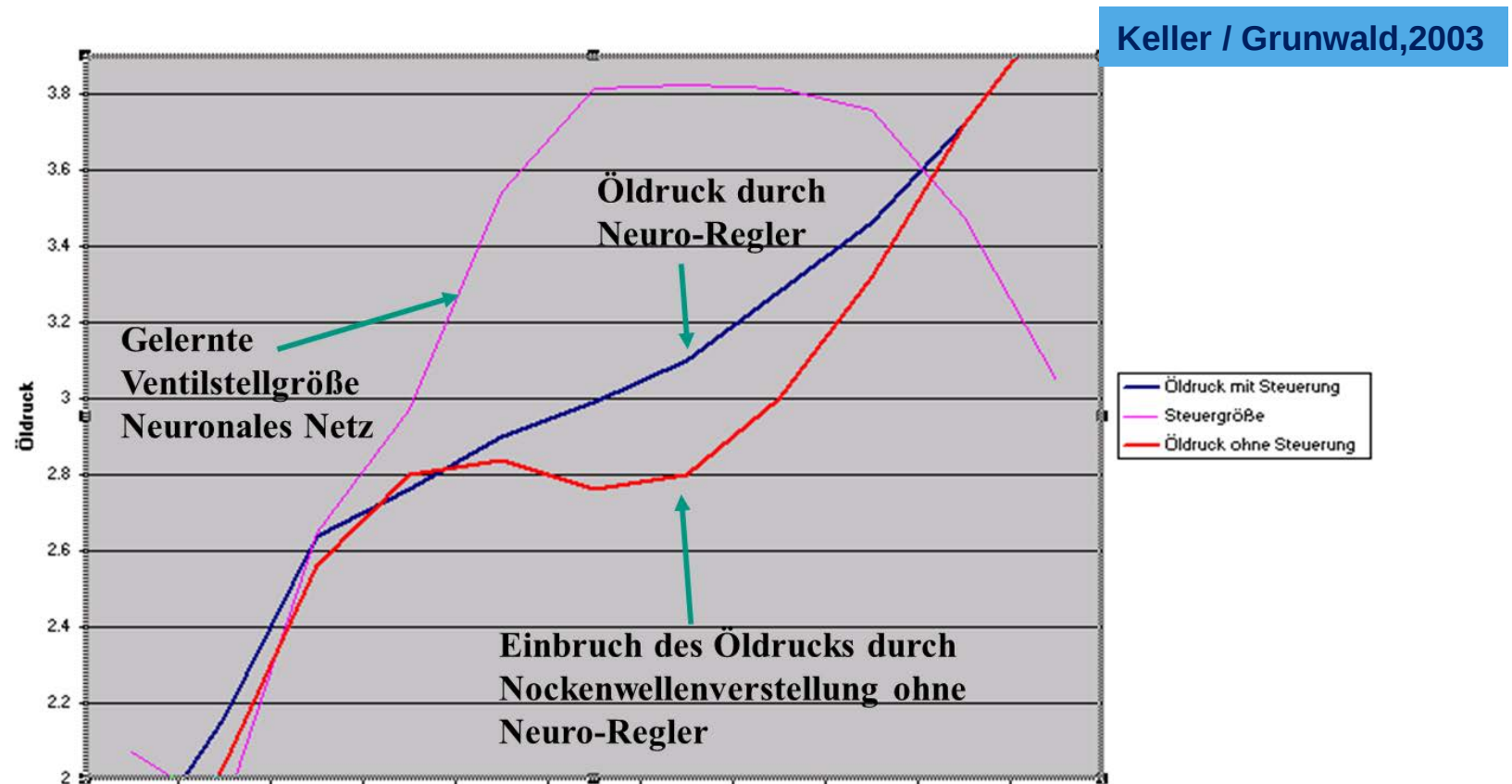
- Optimierung der Ölversorgung eines Hochleistungs-Sportwagens
- MLP+SOFM

Neuronaler Regler lernt Motorverhalten



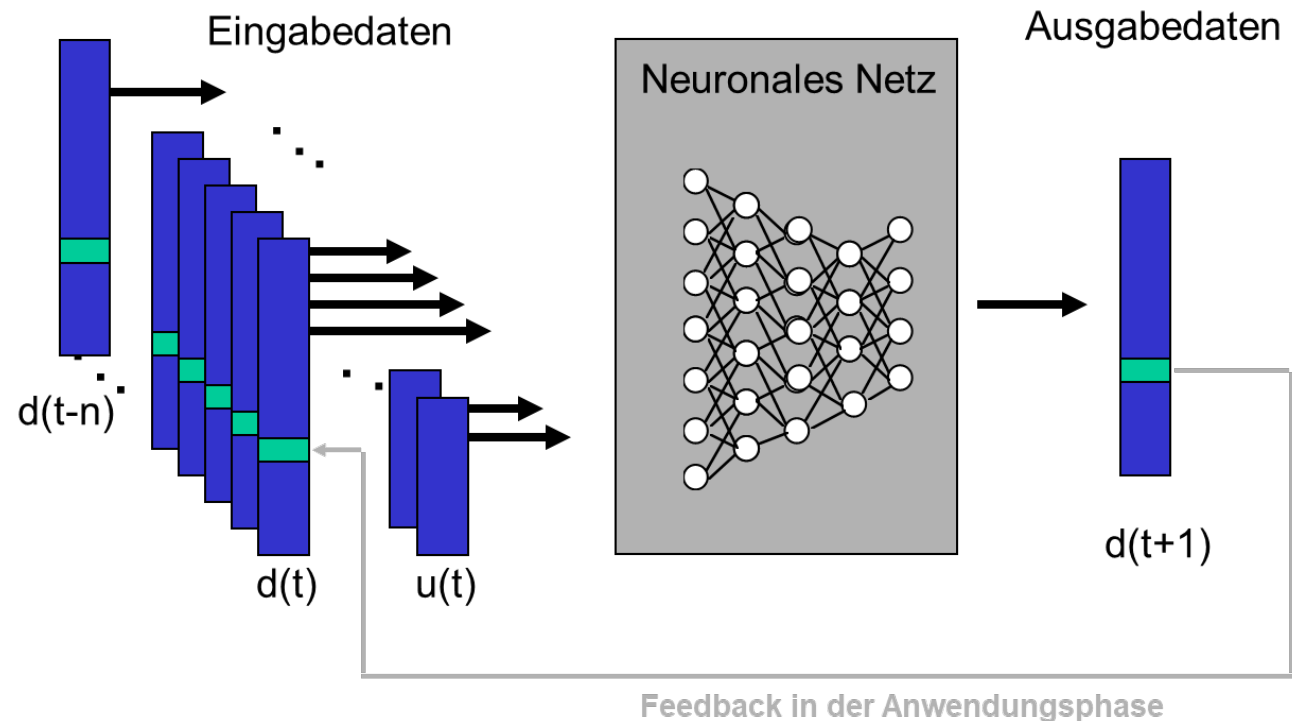
Neurocomputing (3)

- Druckeinbruch durch Nockenwellenverstellung
- Neuronaler Regler lernt Motorverhalten zu kompensieren - frühzeitige Erkennung und Gegenmaßnahme



Neurocomputing (4) - Simulation technischer Prozesse

- Zeitreihen der Prozessgrößen als Eingabe
- Folgezustand aus Ausgabe → Model Predictive Control (MPC) (Müller/Keller 1996)
- Optimierung mit Evolutionären Algorithmen



Neurocomputing (5) - SOFM

Eigenschaften von Tieren klassifizieren

14 Tiere / 13 Eigenschaften, Ähnlichkeiten gruppieren

Variables 13 1 Cases 14 0 0								
	klein	mittelgroß	groß	zwei beine	vier beine	haare		Tierart
01	1	0	0	1	0	0		Taube
02	1	0	0	1	0	0		Huhn
03	1	0	0	1	0	0		Ente
04	1	0	0	1	0	0		Gans
05	1	0	0	1	0	0		Eule
06	0	1	0	1	0	0		Adler
07	0	1	0	0	1	1		Fuchs
08	0	1	0	0	1	1		Hund
09	0	1	0	0	1	1		Wolf
10	1	0	0	0	1	1		Katze
11	0	0	1	0	1	1		Tiger
12	0	0	1	0	1	1		Löwe
13	0	0	1	0	1	1		Pferd
14	0	0	1	0	1	1		Kuh

Neurocomputing (6)

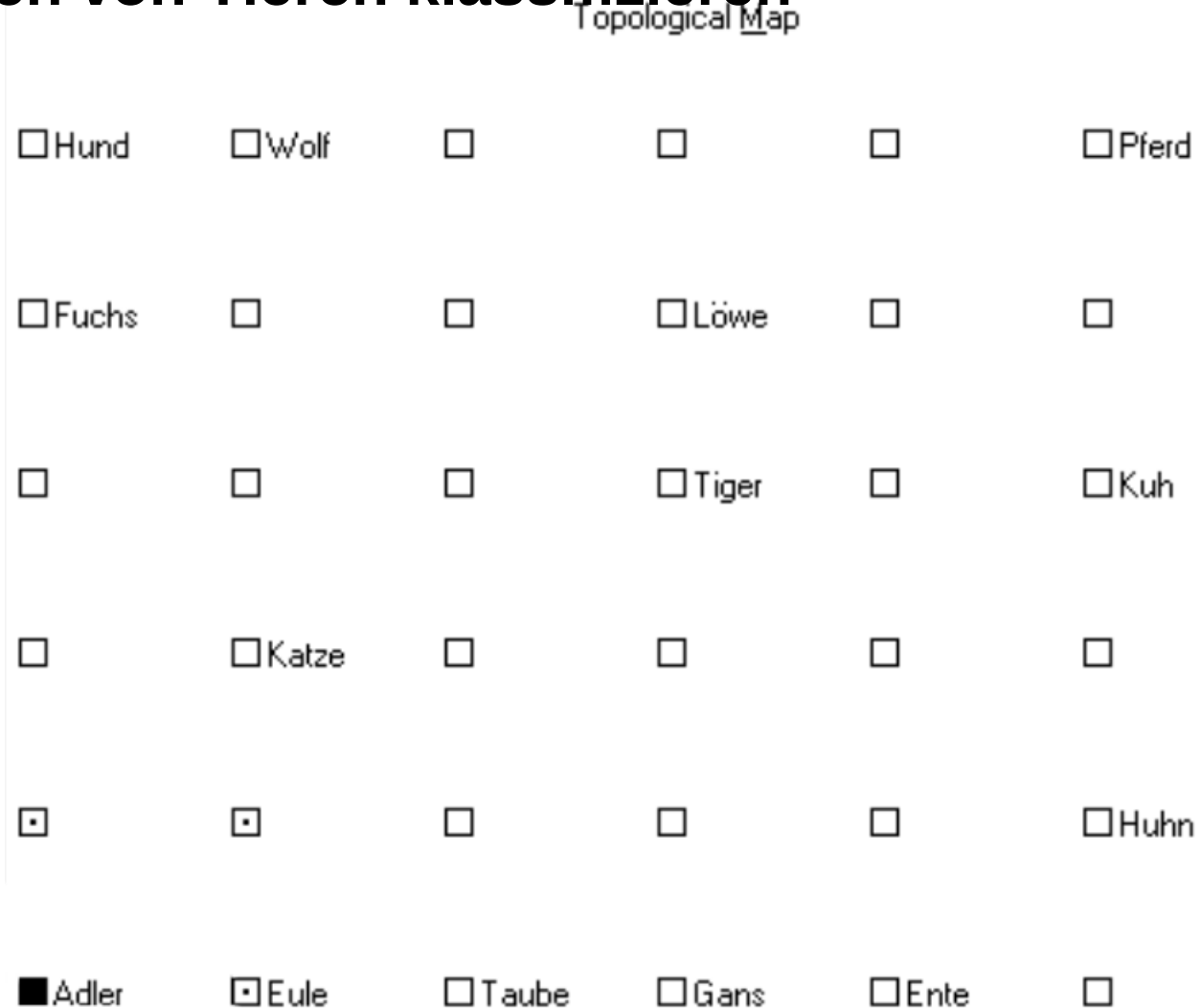
Eigenschaften von Tieren klassifizieren

Data Set Editor									
Variables 13									
	hufe	mähne	federn	jäger	rennt	fliegt	schwimmt	Tierart	
01	0	0	1	0	0	1	0	Taube	
02	0	0	1	0	0	0	0	Huhn	
03	0	0	1	0	0	0	1	Ente	
04	0	0	1	0	0	1	1	Gans	
05	0	0	1	1	0	1	0	Eule	
06	0	0	1	1	0	1	0	Adler	
07	0	0	0	1	0	0	0	Fuchs	
08	0	0	0	0	1	0	0	Hund	
09	0	1	0	1	1	0	0	Wolf	
10	0	0	0	1	0	0	0	Katze	
11	0	0	0	1	1	0	0	Tiger	
12	0	1	0	1	1	0	0	Löwe	
13	1	1	0	0	1	0	0	Pferd	
14	1	0	0	0	0	0	0	Kuh	

Neurocomputing (7)

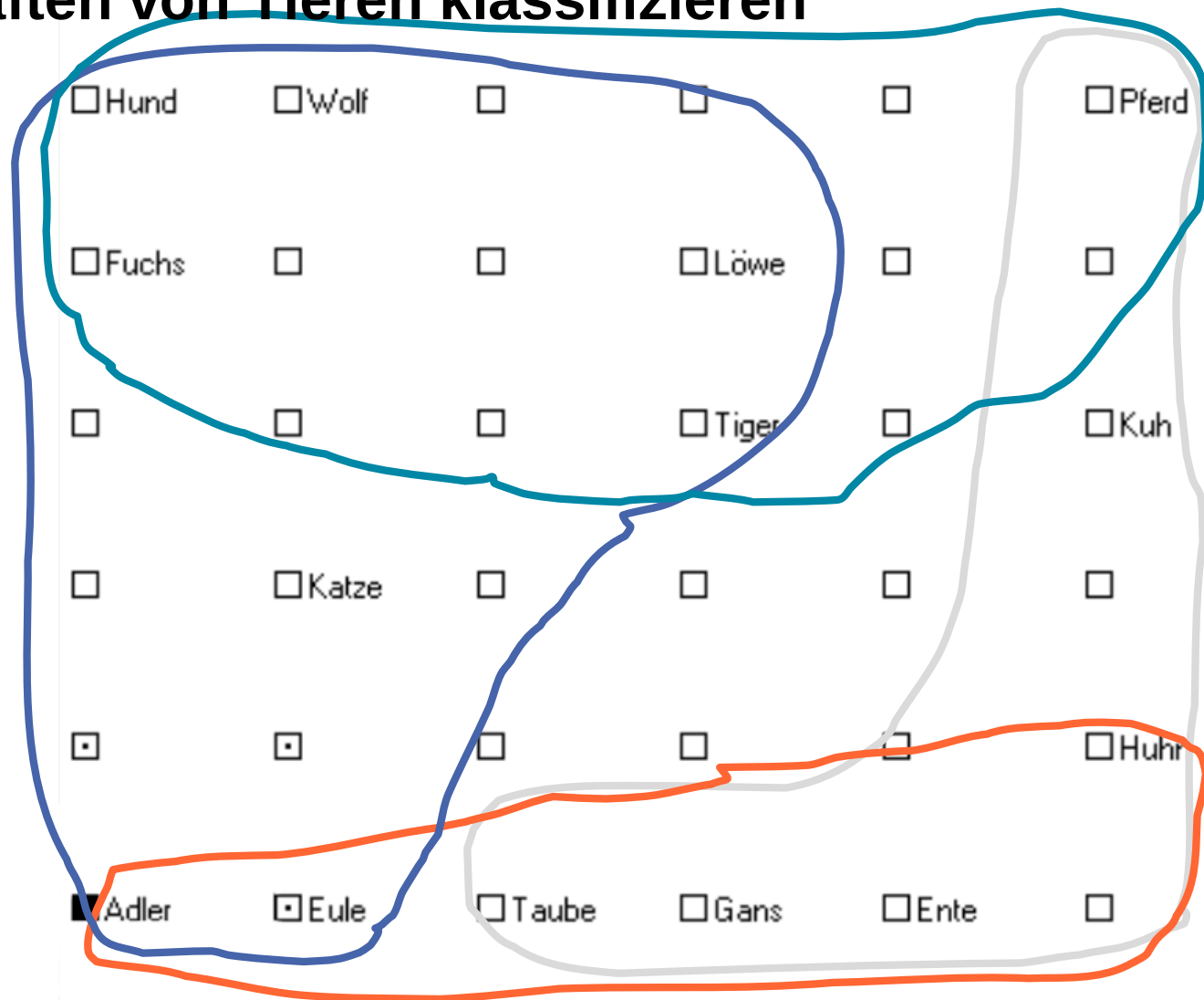
Eigenschaften von Tieren klassifizieren

Einteilung der
Tiere?
Nachbarschafts-
Beziehung



Eigenschaften von Tieren klassifizieren

Topological Map

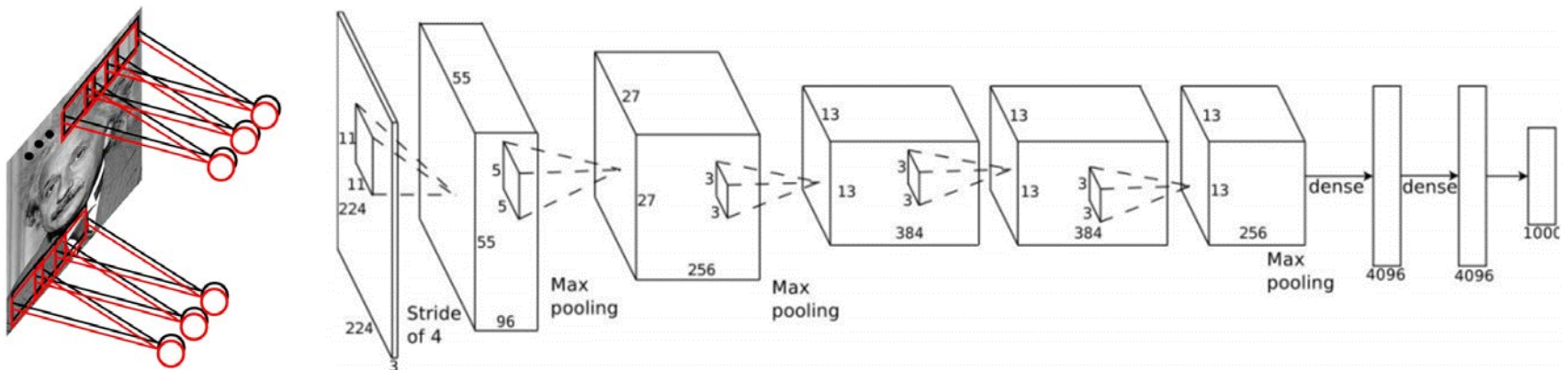


Rennen,
Fliegen,
Jagen,
Friedlich

Teuvo Kohonen (* 11. Juli 1934, Lauritsala, Finnland)

Deep Learning (a lot of hidden layer)

- Deep Learning setzt Rosenblatts Perceptron auf einen neuen Level
- Bildanalysen über neuronale Netze mit vielen Hidden Layern und lokaler Zuständigkeit mit Faltungen (Filter)
- Ziel ist Eigenschaftselemente (**Feature**) von Bildern zu extrahieren, daraus **Objektteile** zusammen zu führen und dann **Objekte** im Bild zu erkennen
- Kann auch für Verhaltensmodellierung eingesetzt werden
- Einsatz für Bildverstehen im autonomen Fahrzeug
- Anwendbar bei definierten Kontexten → Paketverteilanlage
- Komplexe Kontexte → Fahrzeugumfeld → Robustheit kritisch!



Deep Learning - Bildverstehen

■ “EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES”

■ → Panda bleibt Panda – für den Mensch

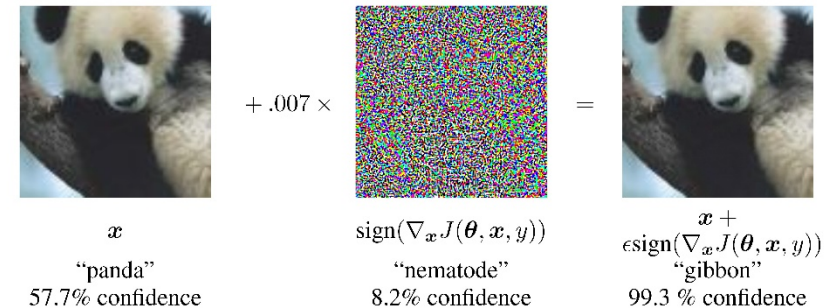


Figure 1: A demonstration of fast adversarial example generation applied to GoogLeNet (Szegedy et al., 2014a) on ImageNet. By adding an imperceptibly small vector whose elements are equal to the sign of the elements of the gradient of the cost function with respect to the input, we can change GoogLeNet’s classification of the image. Here our ϵ of .007 corresponds to the magnitude of the smallest bit of an 8 bit image encoding after GoogLeNet’s conversion to real numbers.

■ “Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition”

■ → eine bunter Brille macht einen anderen Menschen!

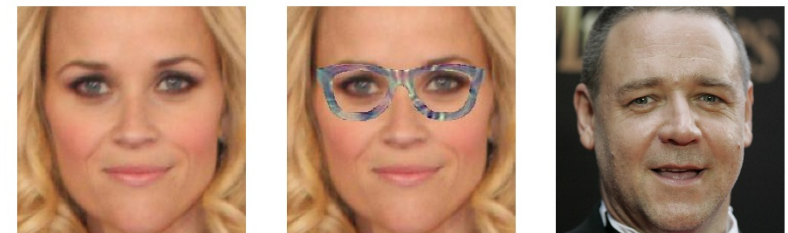
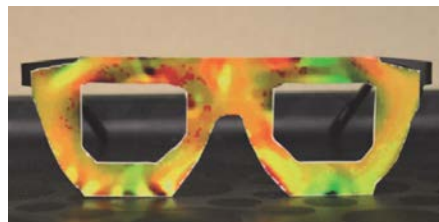


Figure 3: An impersonation using frames. Left: Actress Reese Witherspoon (by Eva Rinaldi / CC BY-SA / cropped from <https://goo.gl/a2sCdc>). Image classified correctly with probability 1. Middle: Perturbing frames to impersonate (actor) Russel Crowe. Right: The target (by Eva Rinaldi / CC BY-SA / cropped from <https://goo.gl/AO7QYu>).

Deep Learning - Bildverstehen

- Deep Learning zur automatischen Erkennung von Situationen und einzelnen Verkehrsteilnehmern
- „Universal adversarial perturbations“ sind für den Menschen irrelevant ...

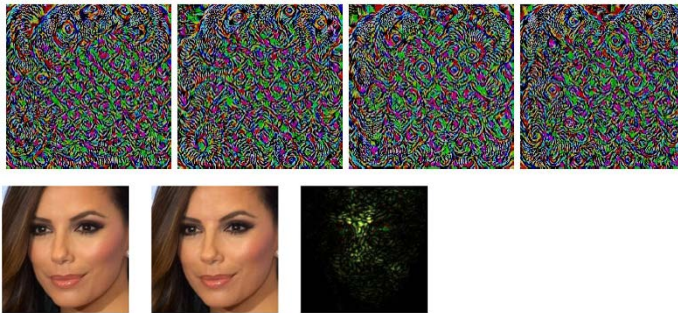


Figure 2: A dodging attack by perturbing an entire face. Left: an original image of actress Eva Longoria (by Richard Sandoval / CC BY-SA / cropped from <https://goo.gl/7QUvRq>). Middle: A perturbed image for dodging. Right: The applied perturbation, after multiplying the absolute value of pixels' channels $\times 20$.

- ... für die Maschine nicht!
→ aus der Wollsocke wird ein indischer Elefant, ... ein Grau-Papagei, ... ein Papagei, ... und so weiter.

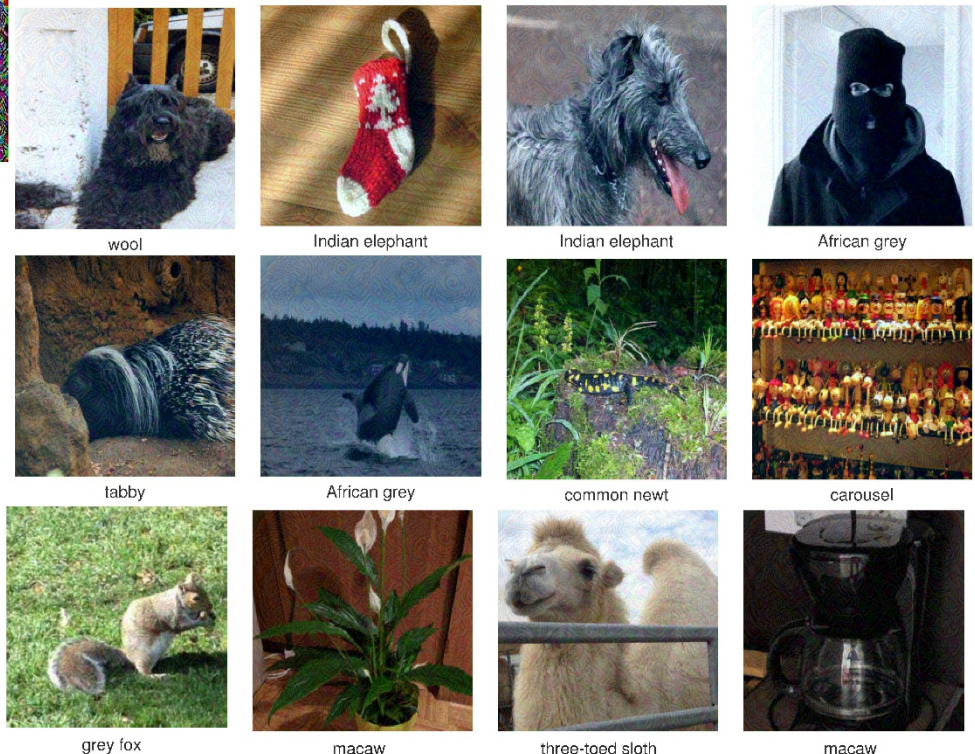


Figure 3: Examples of perturbed images and their corresponding labels. The first two rows of images belong to the ILSVRC 2012 validations set, and the last row are random images taken by a mobile phone camera. We refer to the appendix for the original images.

Deep Learning - Sicherheit Autonomes Fahren

■ Eine reale Anwendung:

Universal Adversarial Perturbations Against Semantic Image Segmentation

Jan Hendrik Metzen

Bosch Center for Artificial Intelligence, Robert Bosch GmbH

janhendrik.metzen@de.bosch.com

Mummadi Chaithanya Kumar

University of Freiburg

chaithu0536@gmail.com

Thomas Brox

University of Freiburg

brox@cs.uni-freiburg.de

Volker Fischer

Bosch Center for Artificial Intelligence, Robert Bosch GmbH

volker.fischer@de.bosch.com

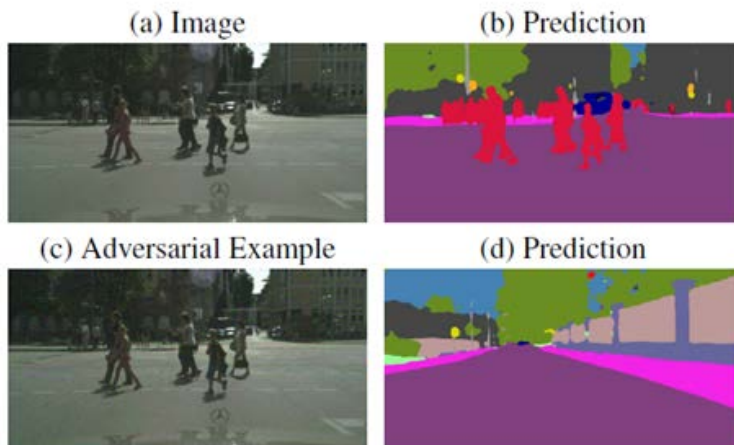
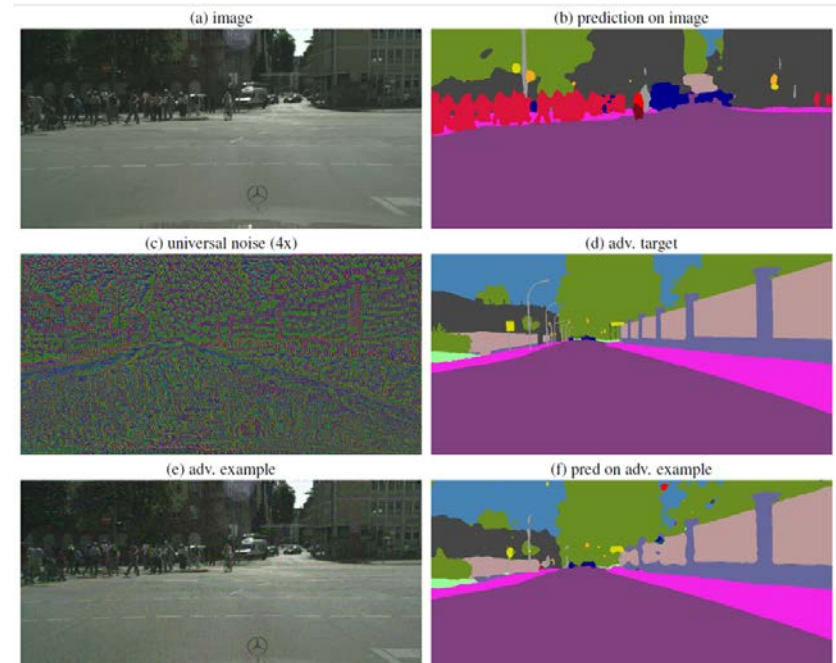


Figure 1. The upper row shows an image from the validation set of Cityscapes and its prediction. The lower row shows the image perturbed with universal adversarial noise and the resulting prediction. Note that the prediction would look very similar for other images when perturbed with the same noise (see Figure 3).



Top Down Inductive Decision Tree - TDIDT

Lernen zu klassifizieren am Beispiel der Simpsons

- Wer ist männlich (Male)?
- Wer ist weiblich (Female)?

Person	Hair Length	Weight	Age	Class
 Homer	0"	250	36	M
 Marge	10"	150	34	F
 Bart	2"	90	10	M
 Lisa	6"	78	8	F
 Maggie	4"	20	1	F
 Abe	1"	170	70	M
 Selma	8"	160	41	F
 Otto	10"	180	38	M
 Krusty	6"	200	45	M

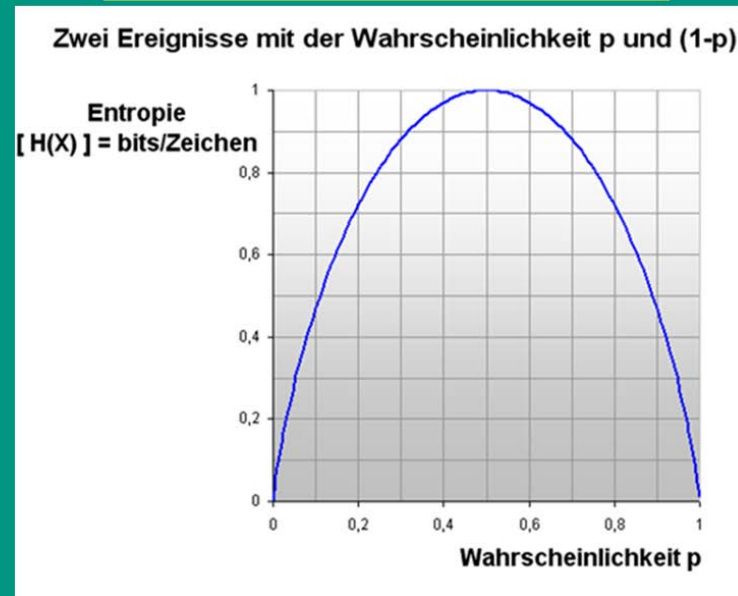
Top Down Inductive Decision Tree - TDIDT

Lernen zu klassifizieren am Beispiel der Simpsons

- Wer ist männlich (Male)?
- Wer ist weiblich (Female)?

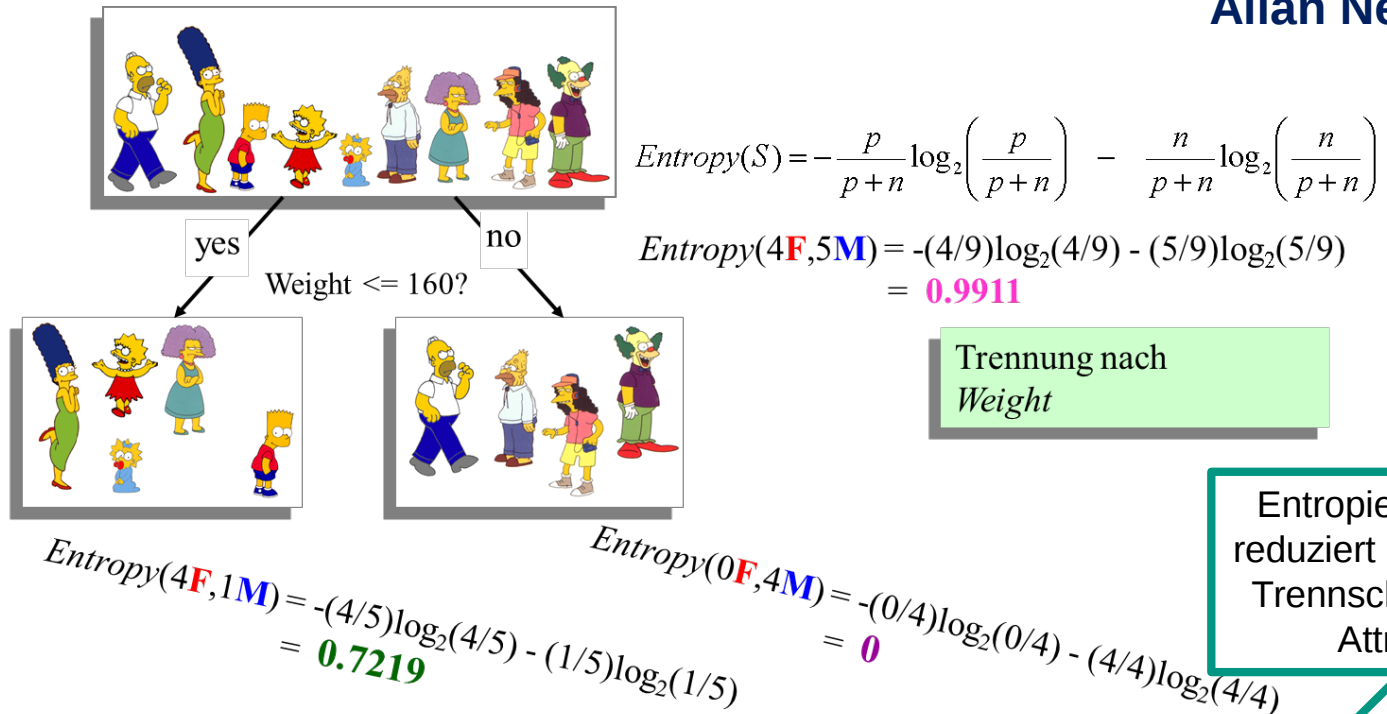
Verwendung der Entropie:

- Je chaotischer, desto höher
- Je geordneter, desto niedriger



- Komplexe Rechnerei (→ Entropie vor und nach Kriterium)

Allan Neymark



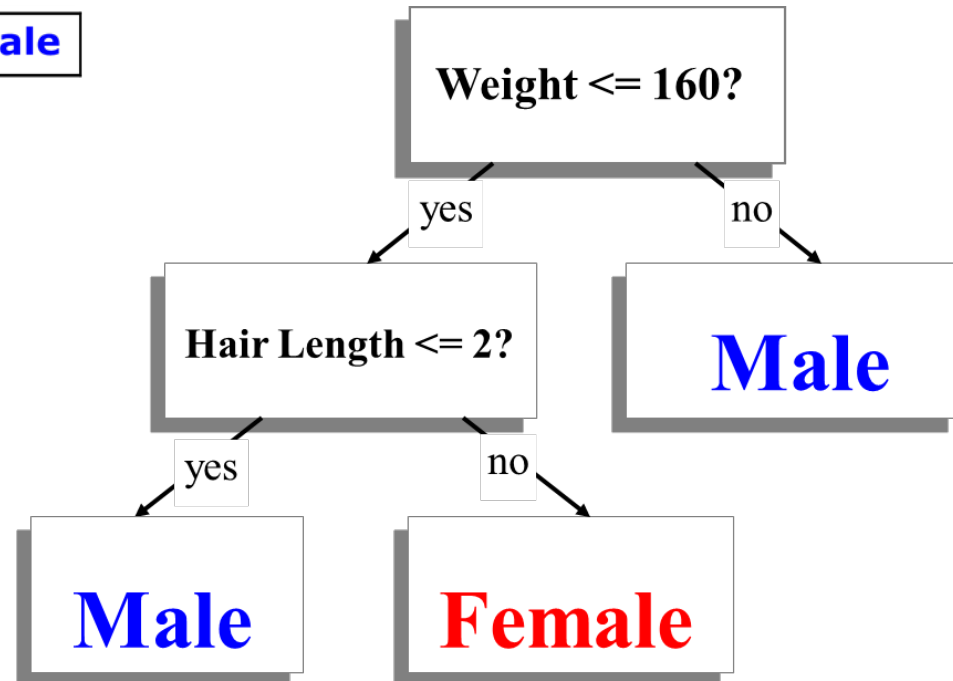
Entropie maximal reduziert = maximale Trennschärfe aller Attribute

$$Gain(A) = E(\text{Current set}) - \sum E(\text{all child sets})$$

$$Gain(\text{Weight} \leq 160) = 0.9911 - (5/9 * 0.7219 + 4/9 * 0) = 0.5900$$

TDIDT

	Comic	8"	290	38	→ Male
--	-------	----	-----	----	--------



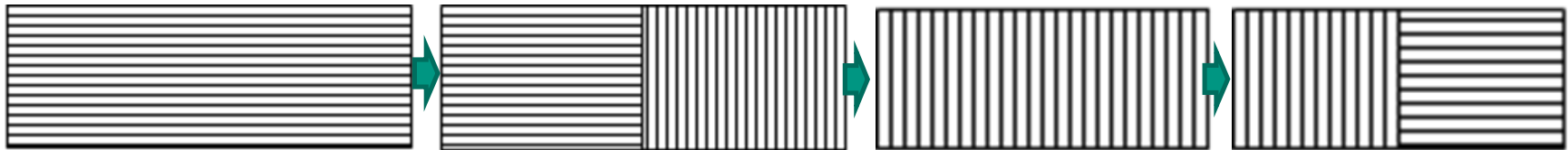
Rules to Classify Males/Females:

If *Weight* greater than 160, classify as **Male**

Elseif *Hair Length* less than or equal to 2, classify as **Male**

Else classify as **Female**

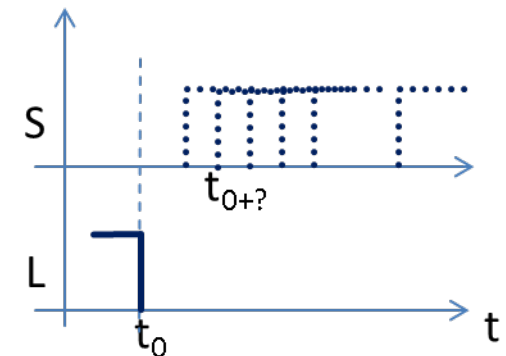
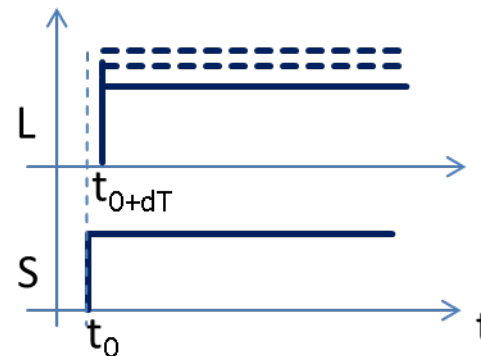
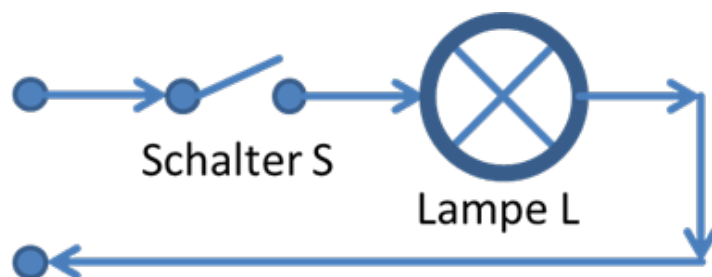
Reiz und Reaktion Lernen (Licht und Schatten)



→ Und nun?
Merkmale und
zeitliche Folge



Keller, 1998/2000



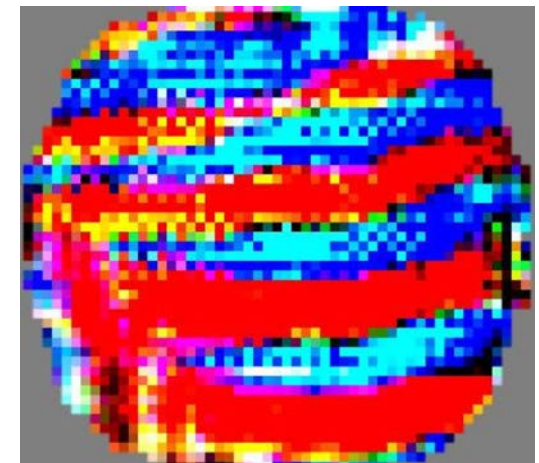
Zeitliche Koinzidenz Schalter-Lampe System (**S→L**)
ganz nach Aristoteles – raumzeitliche Nähe

C3R – Children's Cognitive Learning Behaviour for Causal Reasoning about dynamic Systems (Keller/Weinberger 1994)

KI und Cyber Security und KI

- KI in der Cyber Security
 - Analyse und Lernen von Normalverhalten
 - Analyse und Erkennung von Anomalien
(NISTIR Report 8219, Securing Manufacturing Industrial Control Systems: Behavioral Anomaly Detection)
 - Bitstring / symbolische Elemente zur statischen Analyse
 - Klassifikation im Zugang (Security)
 - ...
- Cyber Security von KI
 - Manipulierbarkeit von KI Verfahren
 - Datenmanipulation (Bilder, Werte, Fake News, ...)
 - Bilderkennung im autonomen Fahren (siehe Bildverstehen)
 - Klassifikationen, Handlungsableitung
 - ...

Farbmuster MPI Tübingen



Wissen Experten was sie wissen?

Realität komplexer Problemstellungen

- Lohausen, Dörner:
Versuchspersonen sollen als Bürgermeister einer Kleinstadt (komplexes System) agieren.
- Die Anforderungen waren hoch, die Versagensquote ebenfalls.
- Die Versuchspersonen (Manager aus der Industrie) konnten kein anwendbares mentales Modell ableiten.
- Die vorhandenen Heuristiken waren zur Beherrschung der Komplexität nicht adäquat genug.
- Die heuristische Kompetenz, d. h. die Fähigkeit des Menschen mit einer solchen Problemstellung umgehen zu können, hängt dabei auch von der jeweiligen Erfahrung ab.

Menschliche Expertise

Kühlhausexperiment

- Reichert:
Einfaches dynamisches System mit Totzeitverhalten (Kühlhaus) wurde an Versuchspersonen zur Führung übergeben mit der Maßgabe, eine bestimmte Temperatur sicherzustellen.
- Es zeigte sich, dass auch hier Probleme in der heuristischen Kompetenz auftraten.
- Erkenntnis:
 - „**die Fähigkeit, das erworbene Wissen zu verbalisieren, in keinen klaren Zusammenhang zu der Steuerungsfähigkeit steht**“.
 - Bei einem korrektem internen heuristischen und entsprechend richtig angewandten mentalen Modell ist das verbalisierte Erklärungsmodell grundverkehrt.
 - Implizit gelernte und durchgeführte Problemlösung ist nicht „verstanden“!!

Intelligenz von Experten

- Experten benötigen bis zu 10 Jahren an Erfahrung für Aufbau von Wissen durch die Beobachtung des Verlaufs der Werte von Systemvariablen über der Zeit und die Integration dieser Beobachtungen in bisherige Erfahrungen.
- „Man muss wohl davon ausgehen, dass (vom Bediener) Wissen über das System erworben und verwendet werden kann, ohne dass dieses unmittelbar verbalisierbar ist.“ (Reichert).
- Hochqualifizierte Experten stützen sich auf Hypothesen, die sie im Verlauf der Problemlösung evaluieren (kontextabhängiges Agieren, in der aktuellen Situation wird das mentale Modell geändert).
- Theoretisches Expertenwissen ist nicht unbedingt Handlungswissen.
- Theoretische Kenntnisse über das Systemverhalten (z. B. in Grenzbereichen) werden bei der (Handlung nicht unbedingt umgesetzt).
- Bsp.: Reaktorunglück in Tschernobyl

Menschliche Intelligenz

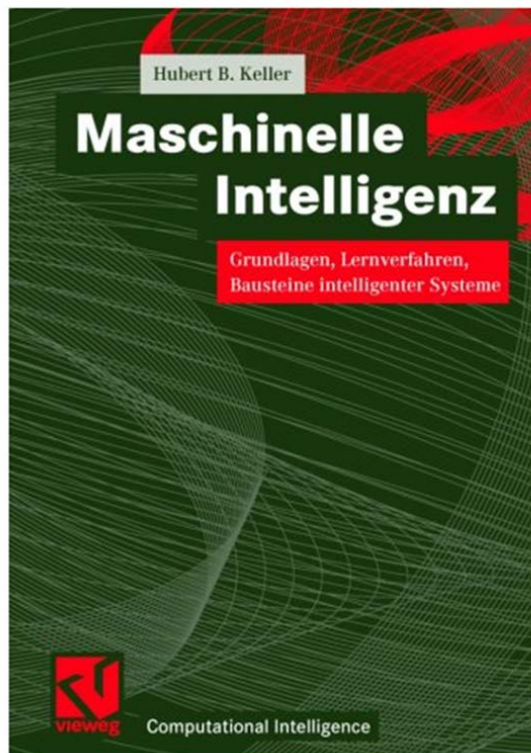
Intelligenz ist (rational, sozial, emotional, ...)

- die zielorientierte Entwicklung von Werkzeugen mit entsprechend intendierten Gebrauch, deren Vorhaltung und permanenten Optimierung
- Modell vom Ich, Umwelt und Werkzeug (Differenz)
- Abstraktion über Handlungen
- Fähigkeit der Reflektion und Moral bei Entscheidungen!

**Wenn die Informatik *LERNENDE* Maschinen baut,
falls das je funktioniert,
WAS und *WIE* lernen diese Maschinen dann und
wissen wir, was diese Maschinen dann tun werden?**

Prämissen und Konsequenzen zum Bau intelligenter Maschinen

- Menschliches Handeln ist komplex
(Emotion, Motivation, Aufmerksamkeit usw.)
- Innere Vernetzung der Variablen ist komplex
(Signalchemie, Zellvorgänge, Zellverbände, Wahrnehmung, ...)
- Genetische Determination, Konditionierung durch Umwelt, Unterbewohner, Selbststeuerung von Denkprozessen (Denken beeinflusst Denken)
- “Was kann ich” → “Was will ich” → “Was brauche ich dazu”
(Modell vom **ICH**, Modell der **Umwelt** und Modell der **Werkzeuge**)
- Intelligenz ist ein sehr frühe Lernprozess, beherrscht nicht jeder gleich gut
(Stufen 4+5: Piaget, Lerntheoretiker)



**Danke für Ihre
Aufmerksamkeit!**



Zum Nachlesen:

Hubert B. Keller, Maschinelle Intelligenz - Grundlagen,
Lernverfahren, Bausteine intelligenter Systeme, Vieweg
Verlag 2000

Biologische Psychologie, Springer-Lehrbuch, Niels
Birbaumer, Robert F. Schmidt

Dörner diverse Bücher

Kontakt: Hubert.Keller@KIT.EDU