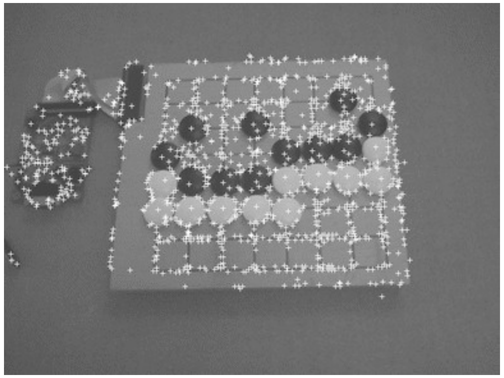
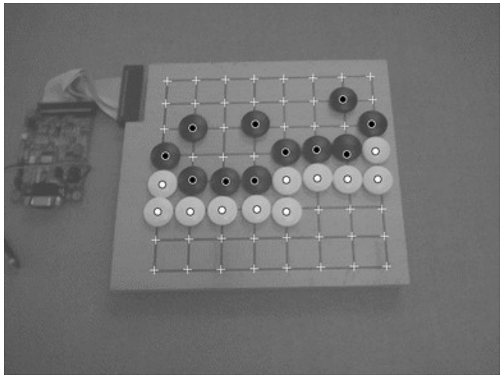
Artificial Intelligence





Alexander K. Seewald

What is Artificial Intelligence?

Systems that think like humans	Systems that think rationally
"The exciting new effort to make computers think... machines with minds, in the full and literal sense" (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning..." (Bellman, 1978)	"The study of mental faculties through the use of computational models" (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason and act" (Winston, 1992)
Systems that act like humans	**Systems that act rationally**
"The art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990) "The study of how to make computers do thinks at which, at the moment, people are better" (Rich and Knight, 1991)	"A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Schalkoff, 1990) "The branch of computer science that is concerned with the automation of intelligent behavior" (Luger and Stubblefield, 1993)

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at

Systems that think like humans

Cognitive Science

1960s Cognitive Revolution: information processing psychology replaced prevailing orthodoxy of behaviourism

Requires scientific theories of brain's internal activities

- Abstraction - level of Knowledge, Assemblies, Neurons...
- Validation - requires predicting and testing behavior of human subjects (top-down = Cognitive Science); and direct identification from neurological data (bottom-up = Cognitive Neuroscience)

Both approaches are distinct from AI but share direction. Much research on visual neuronal correlates of consciousness

Problem: (Prob.)Infeasible even for extremely small organisms

alex@seewald.at
http://www.seewald.at

Systems that think rationally

Laws of Thought

- Normative (or prescriptive) rather than descriptive.
- Aristotle: what are correct arguments / thought processes?
- Several Greek schools developed various forms of logic = notation and rules of derivation for thoughts; may or may not have proceeded to the idea of mechanization.

- Direct line via mathematics and philosophy to modern AI

Problems

- Not all intelligent behavior is related to logical deliberation
- The purpose of thinking = What thoughts should I have?
- Rational thinking is not possible without emotion

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at

Systems that act like humans

The Turing Test

Computing machinery and intelligence [Turing, 1950]

- Can machines think?⇒Can machines behave intelligently?
- Operational test for intelligent behavior = Imitation Game
- Pred. 30% chance for machine to fool lay person for 5mins
- Anticipated all major arguments against AI(!)

Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problems

- Turing test is not reproducible and not constructive
- Chatbots based on (text-)mining terabyte of chat room logs are often judged intelligent by non-experts

Systems that act rationally

<u>Doing the right thing</u>

- Rational behaviour: doing the right thing
- The right thing: which is expected to maximize goal achievement given the available information
- Doesn't necessarily involve thinking, but thinking should be in the service of rational action.

<u>Aristotle (Nicomachean Ethics)</u>

Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good.

<u>Problem</u>

- Definition of good - must be efficent and fast to check for the agent and still be compatible with complex human definitions

AI History

1943	McCulloch & Pitts: Boolean circuit model of brain
1950	Turing's *Computing Machinery and Intelligence*
1952-69	Look, Ma, no hands! - Phase
1950s	Early AI programs: Samuel's checkers, Newell & Simon's Logic Theorist; Winograd's Blocks World
1956	Dartmouth meeting: Artificial Intelligence adopted
1965	Robinsons complete logical reasoning algorithm
1966-74	AI discovers computational complexity
1969-79	Early development of knowledge-based systems
1980-88	Expert systems industry booms
1988-93	Expert systems industry busts: "AI Winter"
1988-	Resurgence of probability; increase in technical depth
	"Nouvelle AI": ALife, Genetic Algorithms, soft computing
1995-	Agents metaphor: face detection, image keypoints, ...

SEEWALD SOLUTIONS
alex@seewald.at
http://www.seewald.at

Recent AI Successes

2005 Stanford's Stanley wins DARPA Grand Challenge, driving autonomously 131 miles through the desert

2010 Kinect 360° motion sensor uses major AI research area *Machine Learning* to recognize body parts from depth information w/o calibr.

2011 IBM's Watson beats the two greatest Jeopardy! champions

2011 Major AI research area *Optical Character Recognition* (OCR) is now mainstream (including contextual grammer/spell checker)

2011 Major AI research area *Speech Recognition* is now mainstream (e.g Apple iPhone4S Siri, OpenSource project simon listens)

2015 Tesla announces software update to enable self-driving on freeways Self-driving cars now mainstream (but not truly autonomous...)

2016 Google DeepMind AlphaGo beats 9dan Go champion Lee Sedol 4:1

Successes are due to increases in computing power, greater emphasis on solving sub-problems, and collaboration with related fields. Many tasks are still *AI-complete*, human-level intelligence is nowhere in sight.

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at

Example: Invisible Person

The Invisible Person project (1999-2005) with the Technical Museum in Vienna was concerned with the creation of an engaging playful agent. **https://www.youtube.com/watch?v=689MqBd1hoo**

alex@seewald.at
http://www.seewald.at

Example: Stanley

Autonomous robot vehicle which won the DARPA Challenge 2005. Built at Stanford University in about 15 months by a team of around 35 people. Uses Machine-Learned Laser Perception and Speed Strategy.

https://www.youtube.com/watc

Example: AI in Finance

<u>At present</u>

- Estimation of creditworthiness (logistic regression / everyone)
- High-speed trading systems capturing arbitrage (Jane Street Capital)
- 24-7 customer service via chatbots (Personetics, Gridspace; Bank of America: Erica)
- Analyzing commercial loan contracts (JPMorgan COIN)

<u>In the future</u>

- Fully automated 24-7 auditing - flagging suspicious transactions instantly as they occur
- Better estimation of creditworthiness using unstructured data (tweets, facebook posts, snapchat images, ...)

SEEWALD SOLUTIONS
alex@seewald.at
http://www.seewald.at

How does it work?

- **Search / Problem Solving**
- **Knowledge and Reasoning**
- **Acting under Uncertainty**
- **Decision Theory**
- **Communication / NLP**

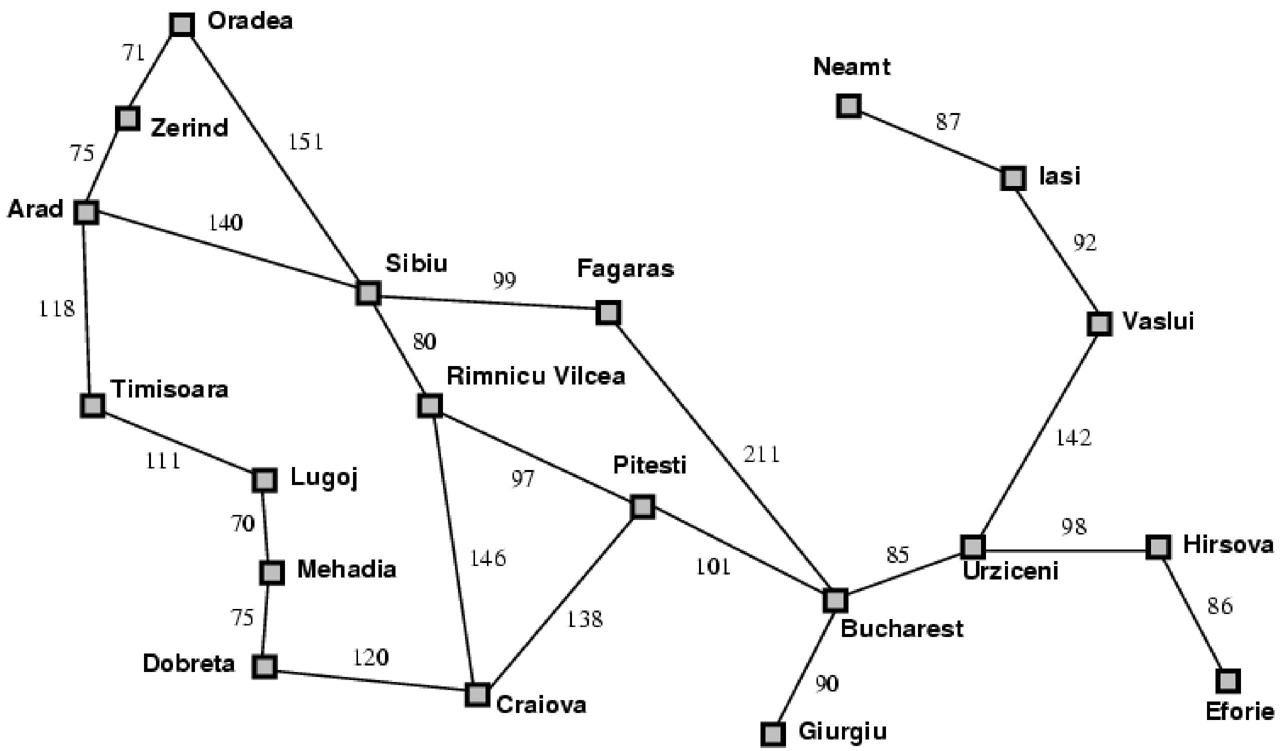
Learning

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at

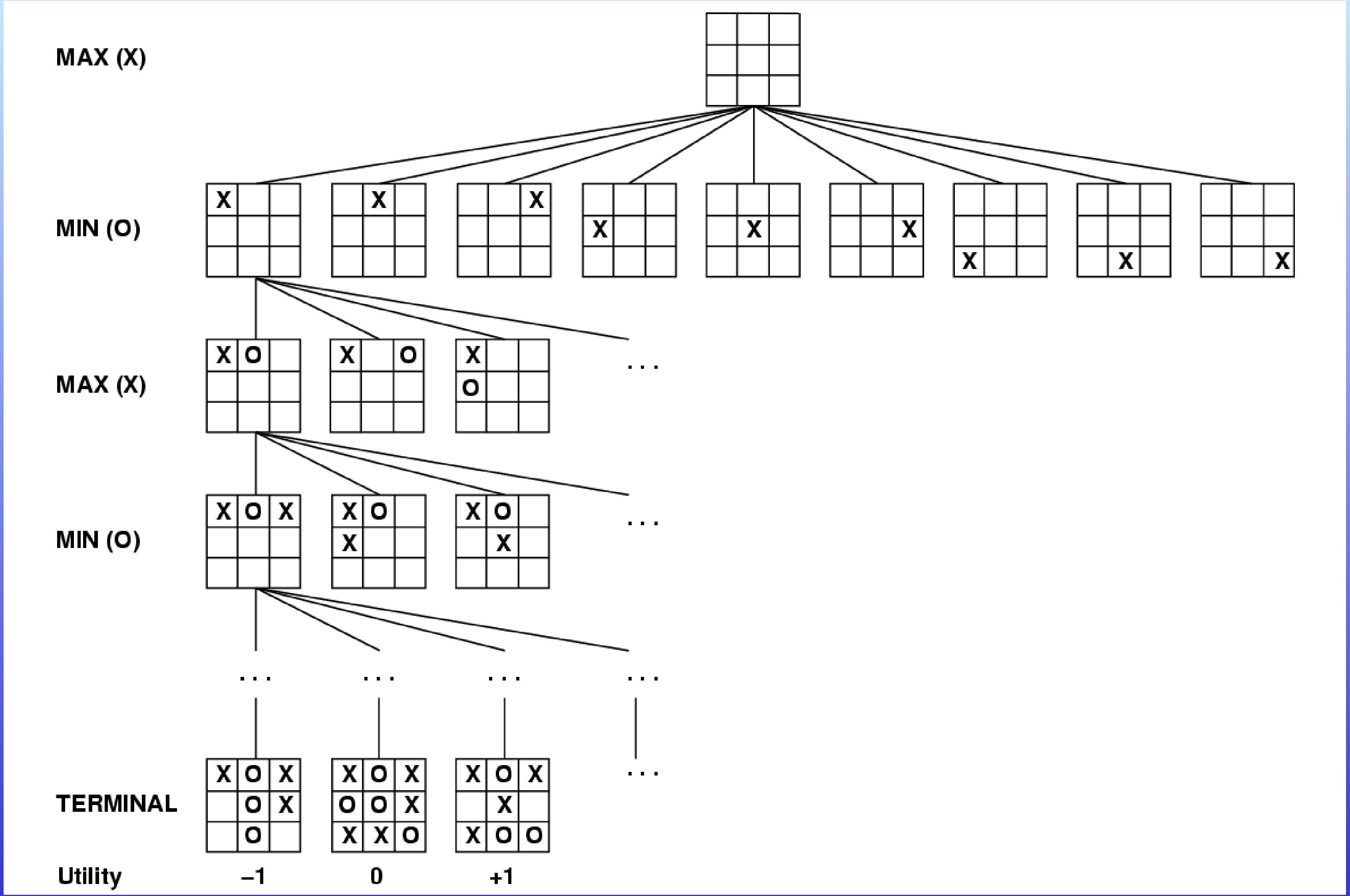
Search / Problem Solving

Search is a central theme in AI. The fastest path through a city; VLSI layout; the correct interpretation of a given sentence; and even general learning - all these can be formulated as search problems.



alex@seewald.at
http://www.seewald.at

Example: Game as Search



Knowledge and Reasoning

Intelligent agents need **knowledge** about the world in order to reach good decisions. Humans use huge amounts of implicit **common-sense knowledge** to solve even tiny tasks. All attempts to model this knowledge have failed.

Constructing **knowledge-based systems** has advantages over programming, but is not feasible for all problems. Modeling relevant knowledge for a task may be infeasible.

State-of-the-Art are **embedded AI** systems, where AI is used complementary to other programming techniques.

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at

Example: RoboSail Systems

- Autopilot for one-person sailing
- Race-proven with various state-of-the-art AI and ML components.
- Human jargon like *gust*, *close-hauled*, *luff* as background knowledge!

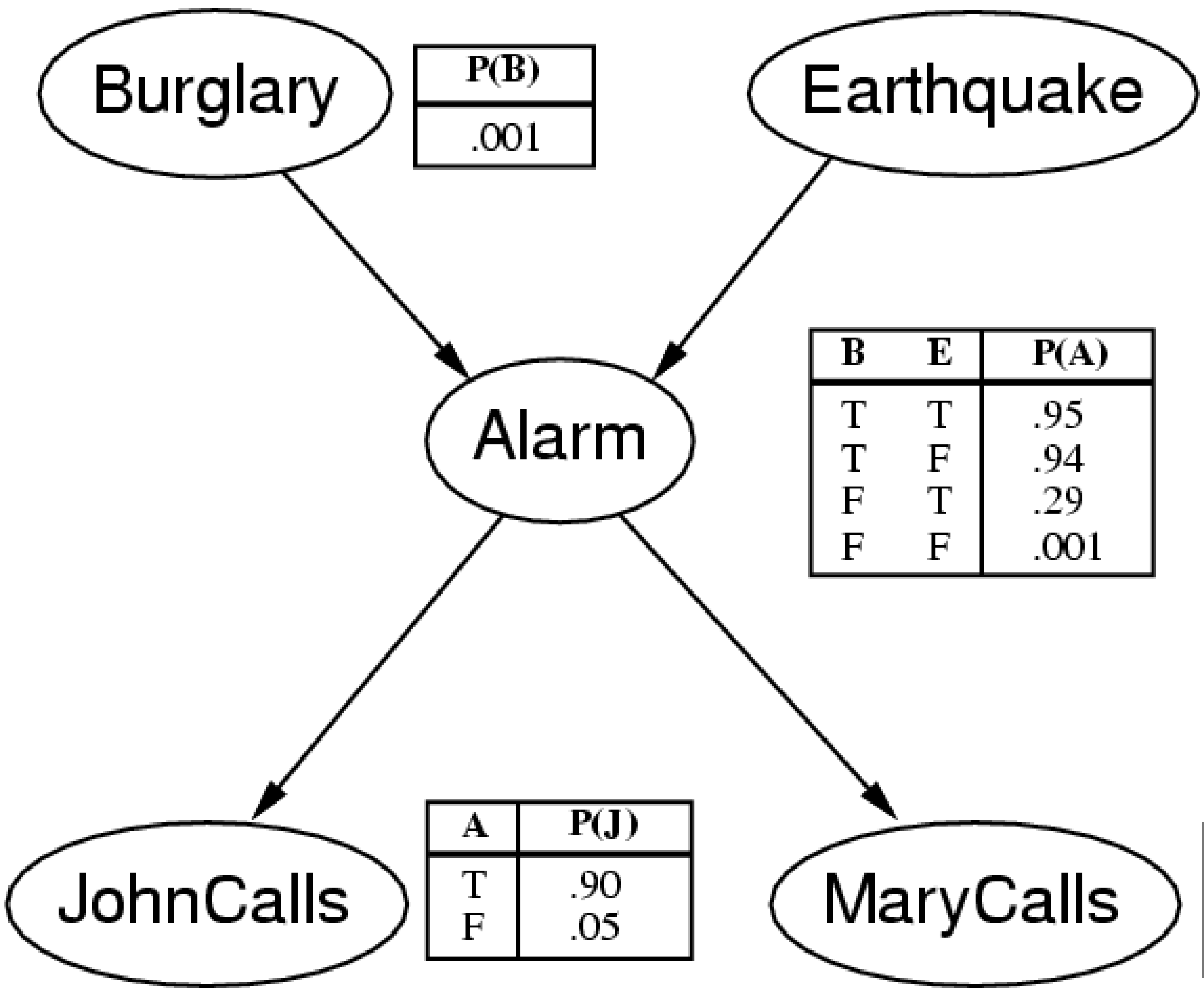
Acting under Uncertainty

Uncertainty is inescapable in complex, dynamic or inaccessible worlds; and means that many simplifications that are possible with deductive inference are no longer valid. **Probability theory** provides a way of summarizing the uncertainty that comes from laziness and ignorance.

Belief networks are a natural way to represent conditional independence information. The links between nodes represent the qualitative aspects of the domain, and the conditional probability tables represent the quantitative aspects.

alex@seewald.at
http://www.seewald.at

Example: Burglar alarm



P(B)
.001

P(E)
.002

B	E	P(A)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J)
T	.90
F	.05

A	P(M)
T	.70
F	.01

alex@seewald.at
http://www.seewald.at

Decision Theory

Simple decision problems can be solved by **decision theory**, which relates what an agent wants (**utility theory**) to what an agent should believe on the basis of evidence (**probability theory**) Utility theory associates a utility value to each state of the agent.

We can use decision theory to build a system that make decisions by considering all possible actions and choosing the one that leads to the best expected outcome. Such a system is known as a **rational agent**.

Decision theory is **normative** - it describes rational behaviour. It is definitely not **descriptive** - people systematically violate the axioms of utility theory.

alex@seewald.at
http://www.seewald.at

Question to the Audience

What would you prefer?

A) 80% chance of winning €4000

B) 100% chance of winning €3000

[Allais, 1953] found that people strongly prefer B)

C) 20% chance of winning €4000

D) 25% chance of winning €3000

[Allais, 1953] found that people strongly prefer C)

Inconsistent human utility theory on monetary value!

$0.8U(\text{€}4000) < U(\text{€}3000)$ and $0.25U(\text{€}3000) < 0.2U(\text{€}4000)$ cannot both be satisfied.

alex@seewald.at
http://www.seewald.at

Communication

Agents need to communicate to each other and to the users. Communication between learning agents is an active research area which sheds light on the development of language in humans.

Natural language processing techniques make it practical to develop programs that make queries to a database, extract information from texts, translate languages, or recognize spoken words.

In all these areas, there exist programs that are useful, but there are no programs that do a throrough job in an open-ended domain.

Example: Alexa

<u>Amazon's Echo NLP user interface</u>

- Personal assistant with speech-based interface (speech data stored and analyzed in the cloud)
- 25 Million units actively used (only US)
- 15,000 Skills (voice-based apps)
- Skills available by Ally Bank, American Express and several other banks: get bank balance, transfer money, get rates, hear recent transactions....
- No user authentication - anyone can give all commands (if pin is not configured)

MACHINE LEARNING

„The field of machine learning is concerned with the questions of how to construct computer programs that automatically improve with experience.“ (Tom M. Mitchell, 1997)

DATA MINING

„Data Mining is the non-trivial process of identifiying valid, novel, potentially useful, and ultimately understandable patterns in data.“ (Fayyad Piatetsky-Shapiro & Smyth, 1996)

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at

Learning

Learning a function from examples of its inputs and outputs is called **inductive learning**. Learning in the inductive setting is supervised and needs a set of training inputs and outputs.

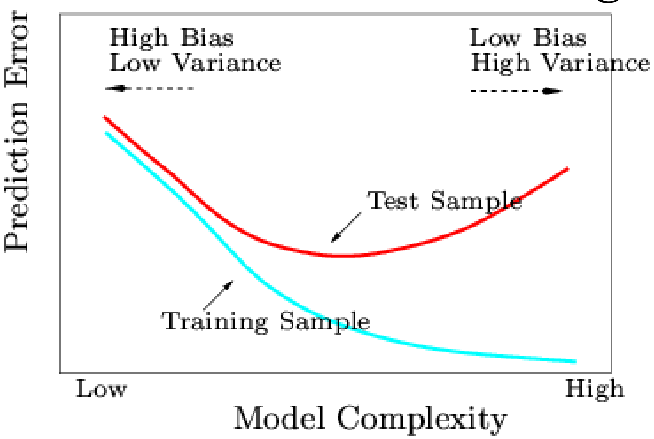
Unsupervised learning uses the structure of training data to infer hidden relationships, which are harder to validate.

Learning in intelligent agents is essential for dealing with unknown environments; and for building agents without prohibitive amount of work. All learning suffers from the **credit assignment** problem = which steps are responsible for a good or bad outcome?

alex@seewald.at
http://www.seewald.at

Bias

"Bias refers to any criterion for choosing one generalization over another other than strict consistency with the observed training instances" (Mitchell, 1980)



Each learning algorithm is biased twofold:

- **language bias** = restricts possible concepts to be learned
- **search bias** = prefers certain models over others

Overfitting occurs when the structure of training data is learned too well; and the generalization performance on unseen data suffers.

Bias is essential to learning!

State-of-the-Art Learning Systems

Simple & Fast Baseline Methods

- OneR, ZeroR

Linear Methods

- Linear & Logistic Regression
- Support Vector Machines with linear & nonlinear kernels

Non-Linear Methods from Machine Learning

- Decision Trees, Random Forest
- Rules Learning (RIPPER)

Non-Linear Methods from Statistics

- NaïveBayes
- Instance-Based learning
- (Convolutional) Neural Networks / Deep Learning

… *many others* …

alex@seewald.at
http://www.seewald.at

Learning Systems - Summary

Characteristic	Lin.& Log.R	SVM	Naïve Bayes	Inst. Based	Dec. Trees	Rule Learn.	Neural Netw.
Natural handling of "mixed" data	–	–	o	–	+	+	–
Handling of MVs	–	–	+	+	+	+	+
Robustness to outliers	–/o	o	o	+	+	+	+
Insensitive to monotone transform.	–	–	–	–	+	+	o
Scalability	+	o/+	o/+	–	+	o	o/+
Robustness concerning irrelevant inputs	–	–	o	–	+	+	+
Interpretability	+	–	o	–	o	+	–
Predictive Power	o	+	–	+	o	o	+

alex@seewald.at
http://www.seewald.at

Example: Mastercard Decision Intelligence

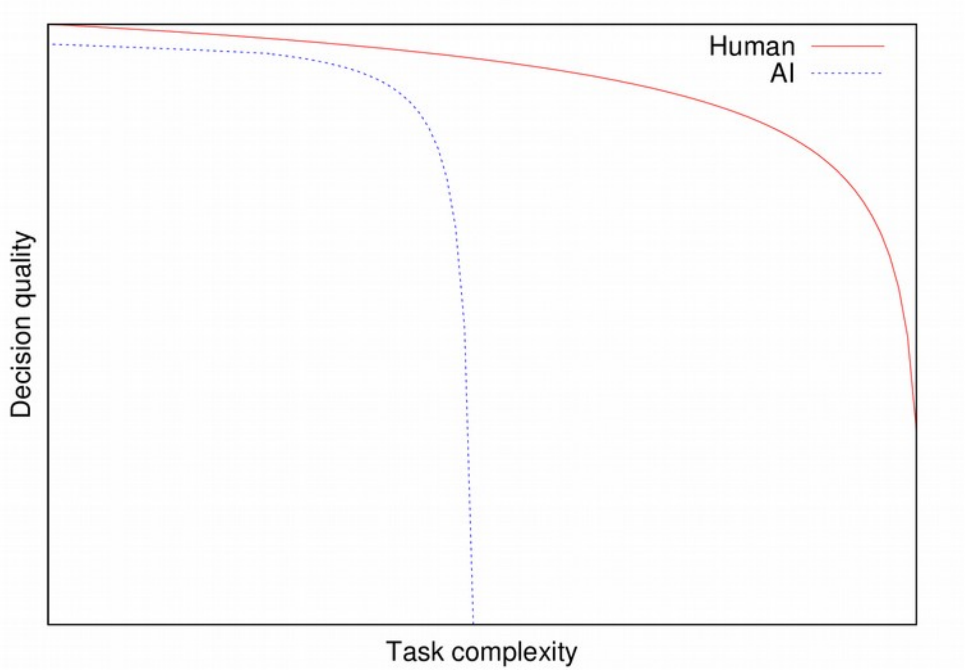
- Generates one predictive real-time fraud score for every credit card transaction
- Distinguishes normal and abnormal spending behaviours from historical data
- Data points used: IP-address of purchasing device, device identification, phone number, email adress
- Modifies normal fraud model towards higher transaction approvability

Goal: To reduce number of transactions falsely declined without increasing overall fraud incidence

alex@seewald.at
http://www.seewald.at

AI Caveats

- AI errors are qualitatively different from human error (less generalization accuracy, "catastrophic failure" / no graceful degradation)
- Most successful systems (deep learning) are least understandable - very hard to characterize error modes!
  - Very subtle correlations with outcome will be found by ML algorithms. This makes it hard to apply non-understandable algorithms to historical data



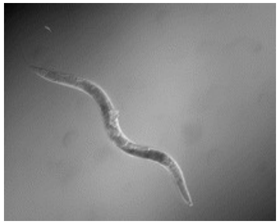
Two ways towards true „strong“ AI

Top-Down

- Build embodied agents (e.g. toy robots)
- Must learn from experience - no ad-hoc code!
- Complex behavior, evolving over time

Bottom-Up

- Analyze simple lifeforms with very small nervous systems (C. elegans: 300 nerve cells)
- Slowly scale up to human-level intelligence



Ex.: Visual images from fMRI cortex activity

Reconstructing visual experiences from brain activity evoked by natural movies. Shinji Nishimoto, An T. Vu, Thomas Naselaris, Yuval Benjamini, Bin Yu & Jack L. Gallant. Current Biology, published online September 22, 2011.

- Record brain activity using fMRI while viewing videos (several hours per person)
- Build statistical model to predict brain activity from video stream using machine learning techniques
- Apply to 18 million seconds of random videos (YouTube)
- Record brain activity for a different stream (90min)
- Recreate viewed video by averaging the 100 streams with the most similar activity - showing the best 30 seconds...

https://www.youtube.com/watch?v=nsjDnYxJ0bo

SEEWALD SOLUTIONS

alex@seewald.at
http://www.seewald.at