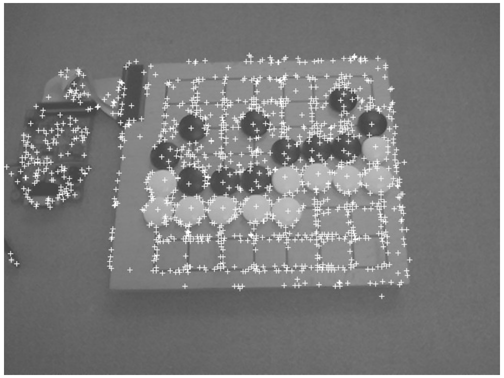
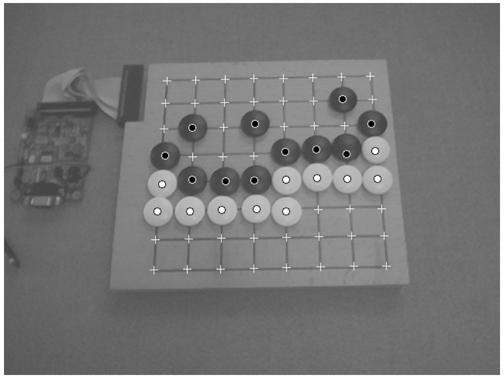
# Artificial Intelligence Crash-Course





Dr. Alexander K. Seewald

SEEWALD SOLUTIONS

# 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)<br><br>"[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)<br><br>"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)<br><br>"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)<br><br>"The branch of computer science that is concerned with the automation of intelligent behavior" (Luger and Stubblefield, 1993) |

# 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

# Systems that act like humans

**<u>The Turing Test</u>**

*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

<u>Problems</u>

- 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


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

# Systems that act rationally

**Doing the right thing**

- 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.

Aristotle (Nicomachean Ethics)

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

Problem

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

# AI Prehistory

| | |
|---|---|
| **Philosophy** | logic, methods of reasoning<br>mind as physical system<br>foundations of learning, language, rationality |
| **Mathematics** | formal representation and proof<br>algorithms, computation, (un)decidability, (in)tractability, probability |
| **Psychology** | adaptation, phenomena of perception and motor control, experimental techniques |
| **Economics** | formal theory of rational decisions |
| **Linguistics** | knowledge representation, grammar |
| **Neuroscience** | plastic physical substract for mental activity |
| **Control Theory** | homeostatic systems, stability<br>simple optimal agent designs |



SEEWALD SOLUTIONS

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

# 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; real-world applications |



SEEWALD SOLUTIONS

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

# Recent AI Successes

**1997** IBM's Deep Blue Chess program beats Kasparov

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

**2007** CMU wins DARPA Urban Challenge, driving 55 miles through an urban environment, avoid hazards and keeping traffic rules

**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

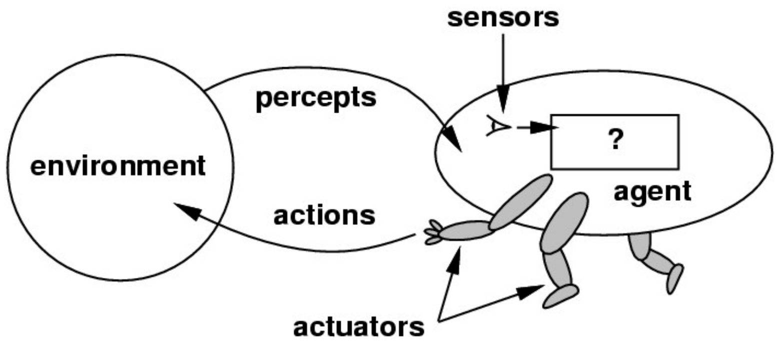
**2011** Major AI research area *Optical Character Recognition* (OCR) is now mainstream (including contextual grammer/spell checker)

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

**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

# Agents and environments




An agent is everything that perceives and acts.

The whole field of AI can be viewed as being concerned with design of intelligent embodied agents.

Agents include humans, robots, softbots, vacuums cleaners...

The agent function maps from percept histories to actions:

$$f: P^* \to A$$

For any given class of environments and tasks, we seek the agent with the best performance. Computational limitations make perfect rationality unachievable.





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

# Example: Vacuum cleaner agent

**Percepts:** clean/dirty, wall, stairs

**Actions:** move, rotate, clean

**Goals:** maximize amount of dirt collected / cleanliness

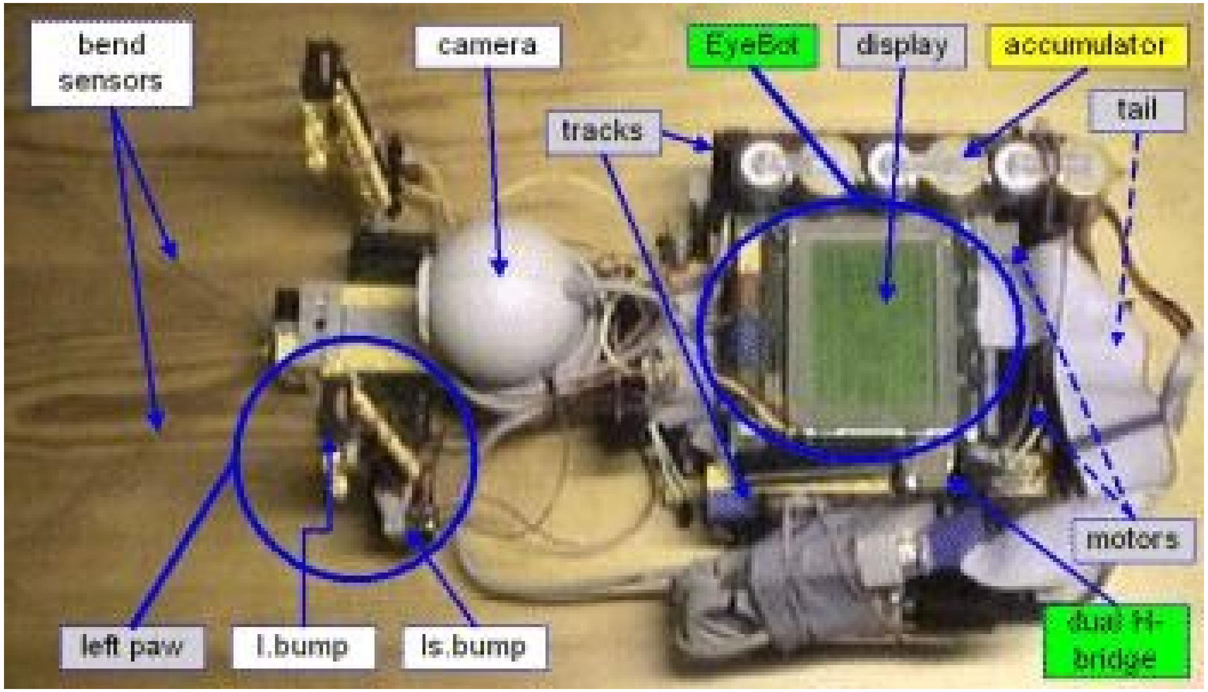
**Environment:** single-level household



SEEWALD SOLUTIONS

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

# Example: RoboCat




RoboCat (Seewald, 1999; Diploma thesis) is an example for a robot with is controlled by emotion and motivation, and shows realistic non-deterministic behaviour. **Video**

# 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. **Video**



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

# How can we build such agents?

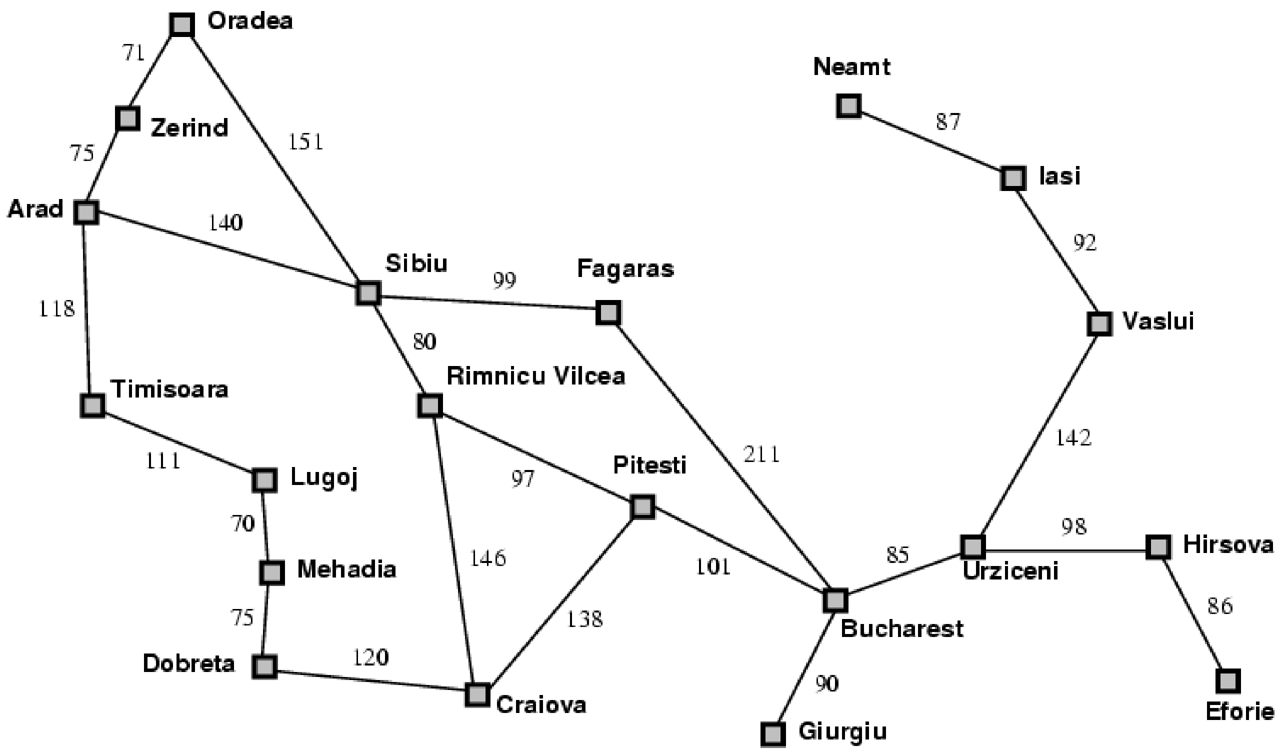
- Search / Problem Solving
- Knowledge and Reasoning; Planning
- Acting under Uncertainty
- Decision Theory
- Communication / NLP

- Learning

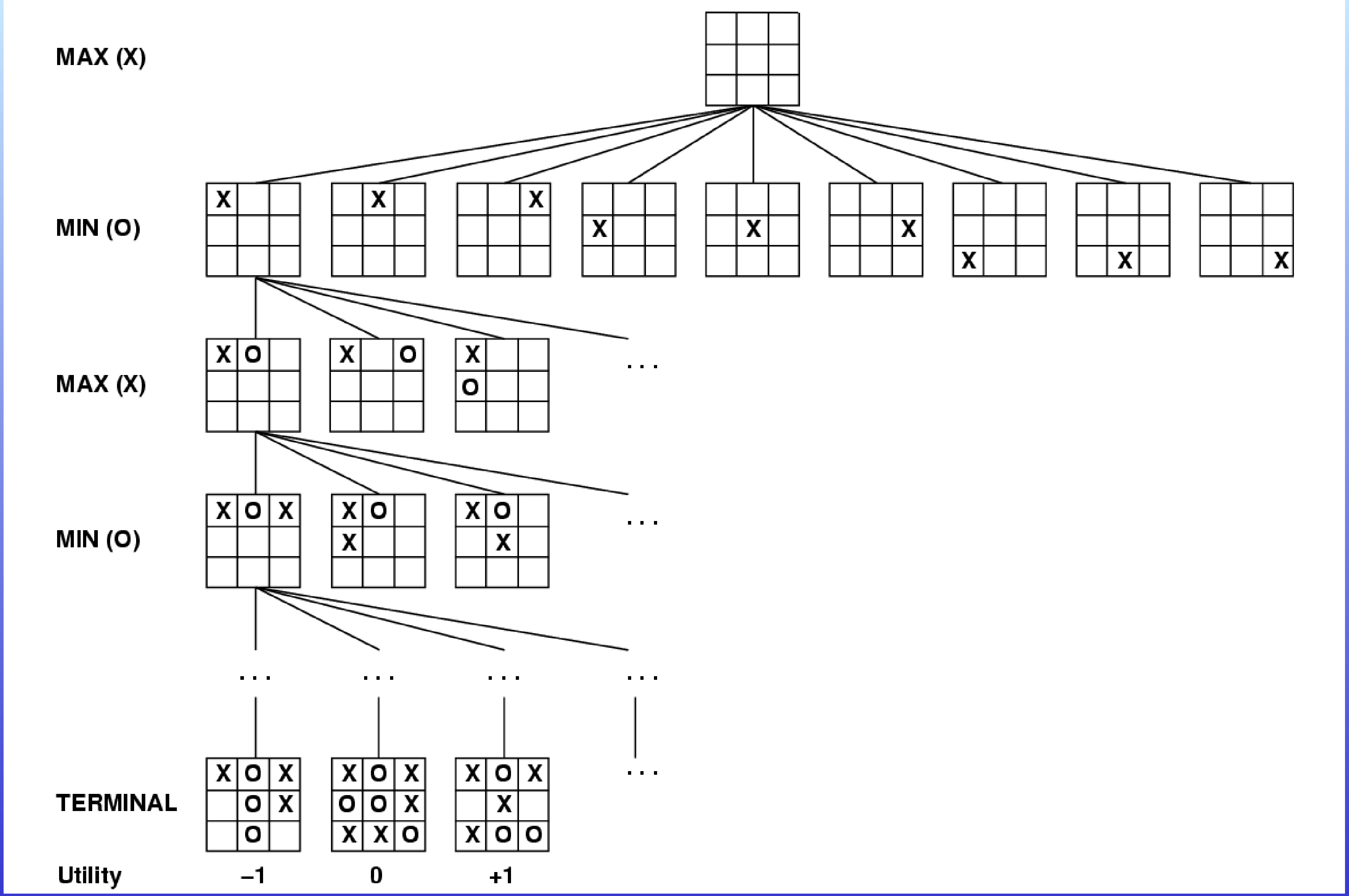

alex@seewald.at
http://www.seewald.at

# 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.



# Example: Game as Search






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

# 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.

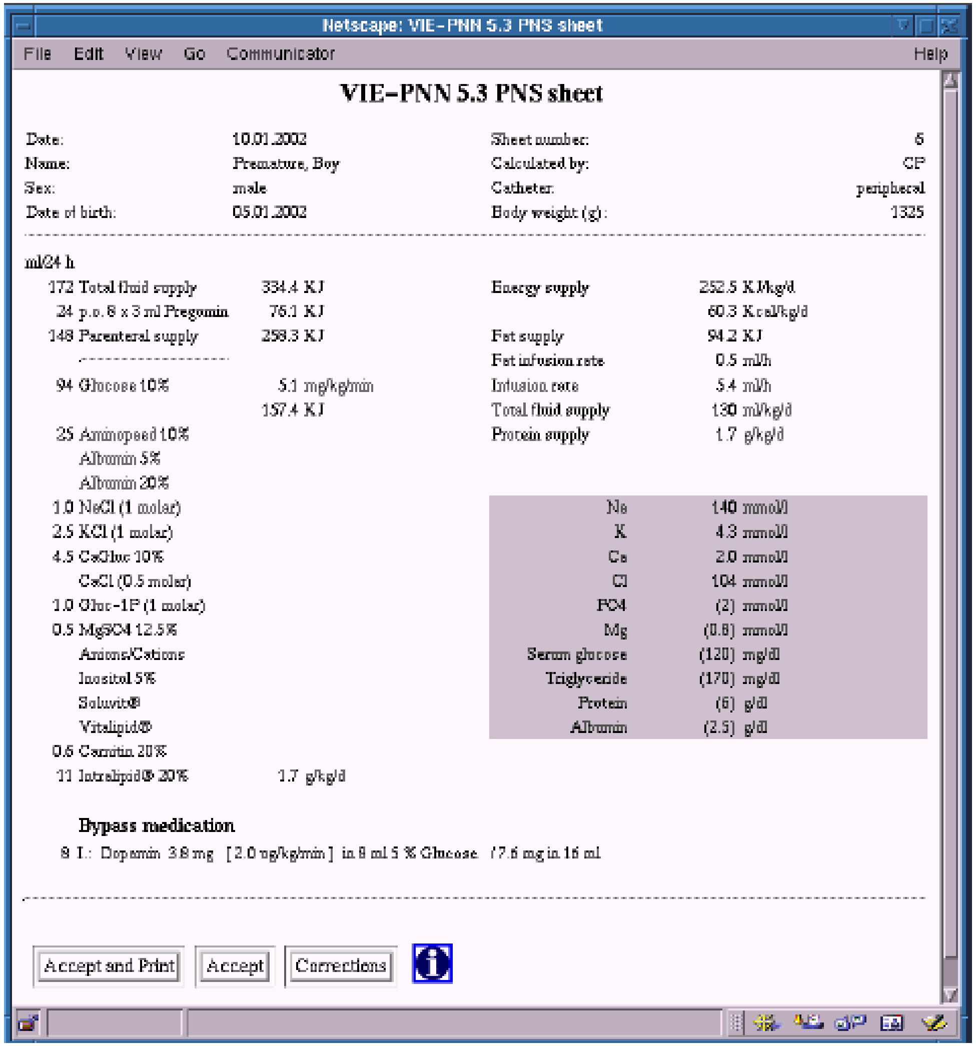
State-of-the-Art are **embedded AI** systems, where AI is used complementary to other programming techniques.



SEEWALD SOLUTIONS

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

# Example: VIE-PNN




- Knowledgebased system for neo-natal nutrition
- Rules derived from expert knowledge.
- HTML-based interface.
- In clinical use for >5 years at AKH Vienna

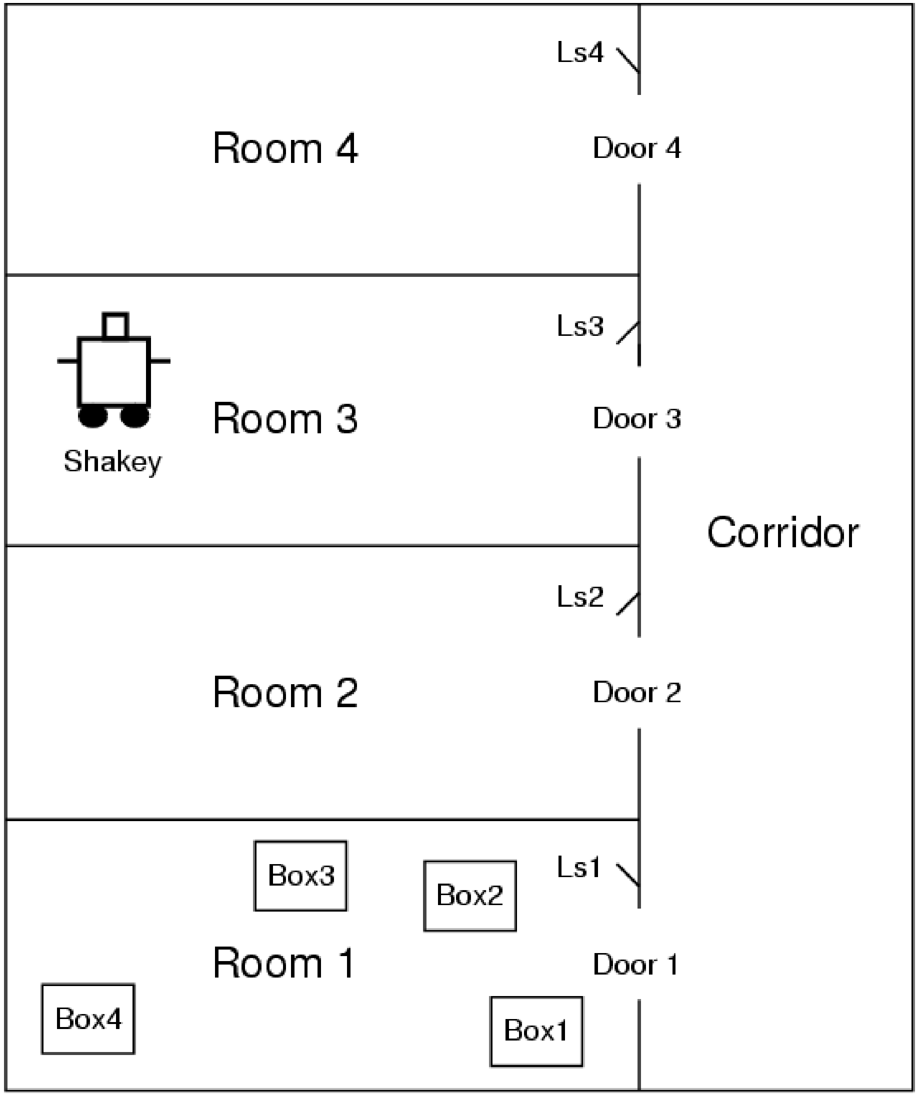
# Planning

**Planning** agents *look ahead* to come up with actions that will contribute to goal achievement. They differ from problem-solving agents in their use of more flexible representations of state, actions, goals, and plans.

Planning systems can be seen as efficient special-purpose reasoning systems designed to reason about actions; or as efficient search algorithms for the space of possible plans.

Automatic planners and schedulers have proven capable of handling complex domains such as spacecraft missions and manufacturing

# Example: Shakey



# Acting under Uncertainty

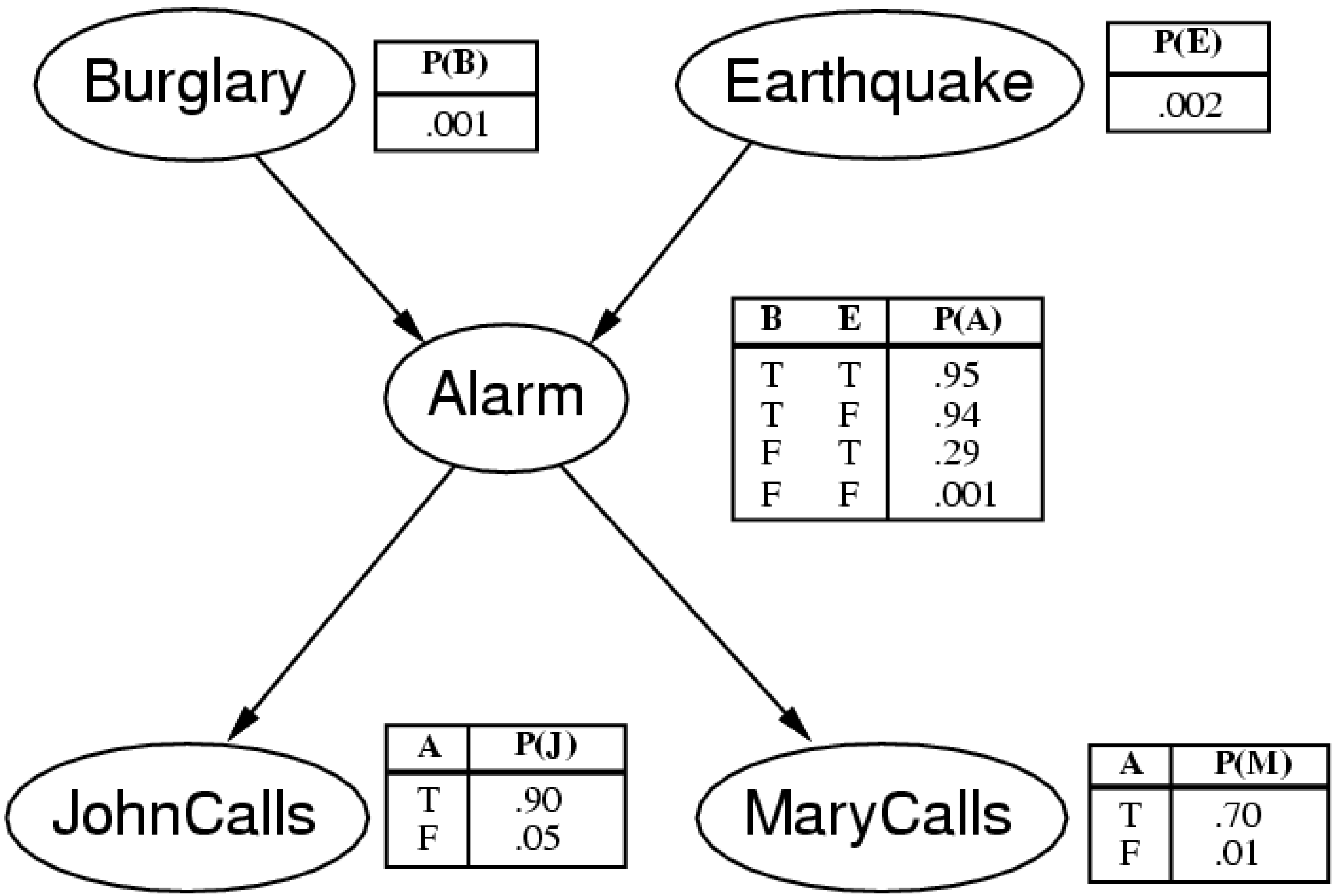
Uncertainty is inescapable in complex, dynamic or inaccessible worlds; and means that many simplifications that are possible with deductive inference are no longer valid. **Probability theory** provides a way of summarizing the uncertainty that comes from laziness and ignorance.

**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 |

# 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.

# 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.

# Learning (1)

**A large variety of learning systems is available which can learn...**

- A state evaluation function to play checkers
- A belief network to model sleep stages
- A function to predict insurance risks
- Logic programs to determine cancerogenity
- Association rules in supermarket basket analysis
- Time-dependent models of speech
- Shapes of biological objects (e.g. erythrocytes)
- The color and texture of a walkway in a park

...



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

# Learning (2)

**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?

**Reinforcement learning** is an active research topic, and computationally very expensive. Temporal difference learning and Q-Learning are common learning algorithms.

**Genetic algorithms** achieve reinforcement by increasing the proportion of successful functions. They achieve generalization by mutating and cross-breeding programs.



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!**



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

# But learning is still hard! Why?

**Inductive learning is inherently risky**

- There is no safe way to predict the future.
- Bias is essential, but may be wrongly chosen.

**No Free Lunch!**

- Theoretically, it is not possible to learn anything.
- Practically, the world shows an enormous variety of patterns. Life has adapted over billions of years to take advantage of these specific patterns.



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

# What would be needed for 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
- *May never work*

**Bottom-Up**

- Analyze simple lifeforms with very small nervous systems (e.g. C. elegans: 300 nerve cells)
- Slowly scale up to human-level intelligence
- *May not yield a true understanding of intelligence at all & may never work*

# 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...

**Video**