Session #13 & 14

CMSC 409: Artificial Intelligence

http://www.people.vcu.edu/~mmanic/

Virginia Commonwealth University,
Fall 2023,
Dr. Milos Manic
(mmanic@vcu. edu)

1

CMSC 409: Artificial Intelligence

Session # 13 & 14

Topics for today

- Announcements
- Previous session review
- Learning - the effect of transfer function
- Robotics and Artificial Intelligence
  - *A little bit of history*
  - *Etymology of the words and three laws of robotics*
  - *Definitions of robot*
- Intelligence in robots
- Artificial vs. Biological Intelligence
  - *Engineering counterpart to biological intelligence*
  - *Control and regulation*
  - *Feedback control system*

2

CMSC 409: Artificial Intelligence

Session # 13 & 14

Topics for next time…

- Robot learning
  - *Types of learning*
  - *Learning controllers*
  - *Adaptive learning*
  - *Reinforcement learning*
- Reinforcement learning (RL)
  - *Learning through reinforcement*
  - *Box pushing example (textbook)*
  - *Obelix - sensors equipped;*
  - *Credit assignment problem*
  - *Behavior based software architecture*
- Examples (jumping frogs, tower of Hanoi)

3

CMSC 409: Artificial Intelligence

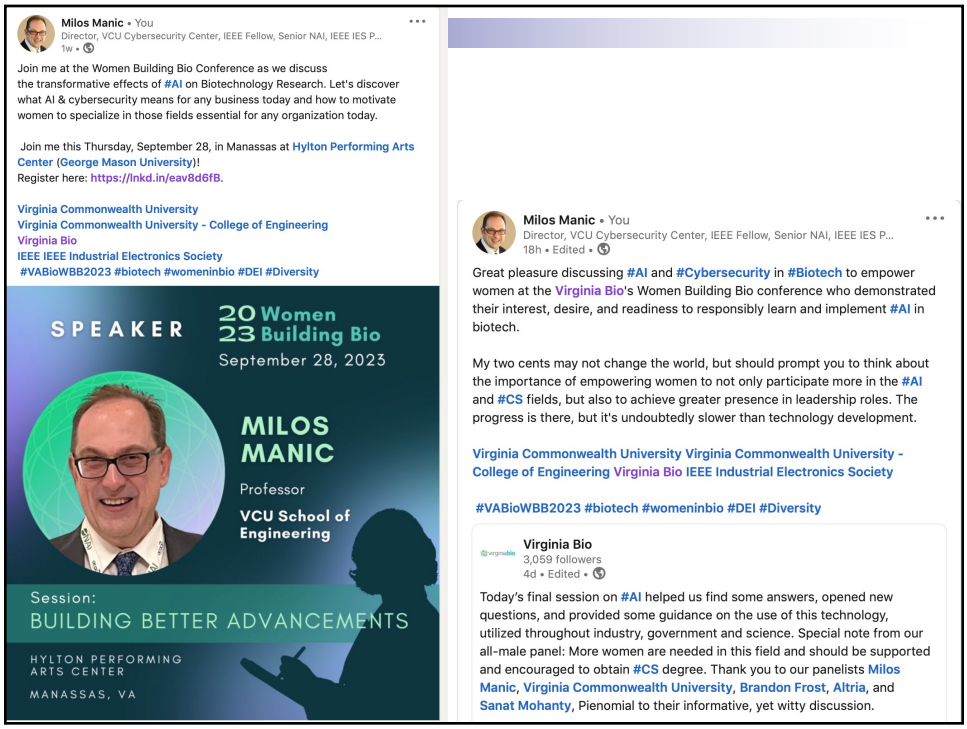
Announcements Session # 13 & 14

- Canvas
  - *New slides posted*
- Office hours zoom
  - *Zoom disconnects me after 45 mins of inactivity. Feel free to chat me via zoom if that happens and I will reconnect (zoom chat welcome outside of office hours as well)!*
- Project #2
  - *Deadline Oct. ~~3~~ 4(noon); Review a week from the deadline.*
- Midterm exam (in-class)
  - *Oct. 19 (Thu); prep examples are posted*
- Paper (optional)
  - *The 2nd draft due Oct. 10 (noon)*
  - *Literature review and updated problem description (check out the class paper instructions for the 2nd draft)*
- Subject line and signature
  - *Please use [CMSC 409] Last_Name Question*

4


Cybersecurity Center
Search
More / Speaker Series / FBI Cybersecurity Regional Collegiate Academy
FBI Cybersecurity Regional Collegiate Academy
Learn how the FBI catches CYBER CRIMINALS
Apply to the
FBI CYBERSECURITY REGIONAL COLLEGIATE ACADEMY
4-WEEK SESSION: OCT 4, OCT 11, OCT 18, OCT 25
6PM - 8PM | B1127
ATTEND ALL SESSIONS TO EARN A CYBER CERTIFICATION FROM THE FBI!
VCU Cybersecurity Center
Scan this QR code to apply
Hosted by Milos Manic
4- Week Session: Oct. 4, Oct. 11, Oct. 18 and Oct. 25
Snead Hall, Room B1127
https://cybersecurity.vcu
Apply today
13 & 14. Updated on 10/3/23 9:08:49 AM

5


Milos Manic • You
Director, VCU Cybersecurity Center, IEEE Fellow, Senior NAI, IEEE IES P...
1w •
Join me at the Women Building Bio Conference as we discuss the transformative effects of #AI on Biotechnology Research. Let's discover what AI & cybersecurity means for any business today and how to motivate women to specialize in those fields essential for any organization today.
Join me this Thursday, September 28, in Manassas at Hylton Performing Arts Center (George Mason University)!
Register here: https://lnkd.in/eav8d6fB.
Virginia Commonwealth University
Virginia Commonwealth University - College of Engineering
Virginia Bio
IEEE IEEE Industrial Electronics Society
#VABioWBB2023 #biotech #womeninbio #DEI #Diversity
SPEAKER
2023 Women Building Bio
September 28, 2023
MILOS MANIC
Professor
VCU School of Engineering
Session:
BUILDING BETTER ADVANCEMENTS
HYLTON PERFORMING ARTS CENTER
MANASSAS, VA
Milos Manic • You
Director, VCU Cybersecurity Center, IEEE Fellow, Senior NAI, IEEE IES P...
18h • Edited •
Great pleasure discussing #AI and #Cybersecurity in #Biotech to empower women at the Virginia Bio's Women Building Bio conference who demonstrated their interest, desire, and readiness to responsibly learn and implement #AI in biotech.
My two cents may not change the world, but should prompt you to think about the importance of empowering women to not only participate more in the #AI and #CS fields, but also to achieve greater presence in leadership roles. The progress is there, but it's undoubtedly slower than technology development.
Virginia Commonwealth University Virginia Commonwealth University - College of Engineering Virginia Bio IEEE Industrial Electronics Society
#VABioWBB2023 #biotech #womeninbio #DEI #Diversity
Virginia Bio
3,059 followers
4d • Edited •
Today's final session on #AI helped us find some answers, opened new questions, and provided some guidance on the use of this technology, utilized throughout industry, government and science. Special note from our all-male panel: More women are needed in this field and should be supported and encouraged to obtain #CS degree. Thank you to our panelists Milos Manic, Virginia Commonwealth University, Brandon Frost, Altria, and Sanat Mohanty, Pienomial to their informative, yet witty discussion.

6

Robotics

7

A little bit of history…

Karel Capek (1890-1938)

- The acclaimed Czech writer and playwright, reportedly several times a candidate for the Nobel prize.
- Introduced the word into his play *R.U.R. (Rossum's Universal Robots)* which opened in Prague in January 1921.
- *Robota* (labor, Czech), figuratively "drudgery" or "hard work"), *rossum* (an allusion to the Czech word rozum, meaning "reason", "wisdom", or "intellect")

(Taken from http://www.robotics.utexas.edu/rrg/learn_more/history/#definition http://karelcapek.net/word_robot.htm, http://capek.misto.cz/, http://en.wikipedia.org/wiki/R.U.R._(Rossum%27s_Universal_Robots))

8

Etymology of the word "*Robot*"

Karel Čapek in the Czech daily Lidove Noviny on the Christmas Eve 1933 writes about it thus:

……In an unguarded moment an idea about a play came to the author. While still hot, he took it to his brother Josef, the painter, who was standing at his easel merrily painting away.
"Look Josef", the author said, "I got an idea for a play."
"What is it about?", growled the painter (literally growled, because he held the brush in his mouth).
The author explained in a nutshell what it was about.
"Well, write it down", said the painter without taking the brush out of his mouth nor stopping to paint on his canvass. It was almost insultingly indolent.
*"But I don't know", said the author, "**how to call those artificial workers**. I thought about calling them labors **(labori)**, but it doesn't sound right to me."*
*"Well, call them **robots**," muttered the painter with the brush in his mouth and continued painting. And that was it. This is how the word robot was born; let it be accredited to its true begetter.*
Lidove noviny, 24.12.1933
(taken from http://karelcapek.net/word_robot.htm)

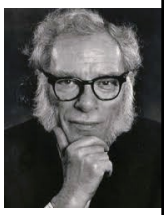
The Čapek brothers - Josef (right), Karel (left)

9

Etymology of the word "*robotics*"

Isaac Asimov (1920-1992)

- Born **Isaak Yudovick Ozimov** in Russia, Jan. 2, 1920
  - *Taught himself to read by the age of 5, graduated high school at 15, and soon became a biochemistry professor with aspirations of one day becoming a writer*
  - *Parents immigrated to the U.S. when Asimov was a toddler, settling into Brooklyn*
  - *1955 became a professor of biochemistry at Boston Univ.*
  - *Authored nearly 500 books*
  - *Influential sci-fi works like I, Robot and the Foundation trilogy, as well as books in a variety of other genres*
  - *Asimov died in New York City on April 6, 1992 at the age of 72.*



(Taken from https://en.wikipedia.org/wiki/Isaac_Asimov
https://www.used-robots.com/education/isaac-asimov-and-his-three-laws-of-robotics
http://www.biography.com/people/isaac-asimov-9190737)

10

Etymology of the word *"robotics"*

Isaac Asimov (1920-1992)

- Used in a short story *Runaround* (1942)
  - *Used the word "robotics" to describe the technology, predicted the rise of a robot industry, presented the "Three Laws of Robotics".*
- Story collection *I, Robot* (1950)
  - *I, Robot looked at human/construct relationships and featured the* ***Three Laws of Robotics****.*
  - *The narrative would be adapted for a blockbuster starring Will Smith decades later. Asimov would later be credited with coming up with the term "robotics.*

(Taken from http://www.robotics.utexas.edu/rrg/learn_more/history/#definition
http://www.biography.com/people/isaac-asimov-9190737#early-life-and-education

11

Etymology of the word *"robotics"*

Isaac Asimov (1920-1992)

- *The year after after "I, Robot", he published…*
- *Foundation (1951)*
  - *Another seminal work, Foundation, a novel that looked at the end of the Galactic Empire and a statistical method of predicting outcomes known as "****psychohistory****."*
- *Robotherapist* - one of the first robots Asimov wrote about
- *Eliza* - a modern counterpart to Asimov's fictional character
  - *MIT Professor Joseph Weizenbaum (1966) wrote a computer program for the study of natural language communication between man and machine.*
  - *Initially programmed with 240 lines of code to simulate a psychotherapist by answering questions with questions.*

(Taken from http://www.robotics.utexas.edu/rrg/learn_more/history/#definition
http://www.biography.com/people/isaac-asimov-9190737#early-life-and-education)

12

Three laws of *robotics* (Asimov)

(Lower numbered laws supersede the higher numbered laws;
a 'zeroth law' added later)

Law Zero:

- *(4th, added later) A robot may not injure humanity, or, through inaction, allow humanity to come to harm.*

Law One:

- *A robot may not injure a human being, or, through inaction, allow a human being to come to harm (unless this would violate a higher order law).*

Law Two:

- *A robot must obey orders given it by human beings, except where such orders would conflict with a higher order law (first law).*

Law Three:

- *A robot must protect its own existence as long as such protection does not conflict with a higher order law (first or second law).*

(Taken from http://www.robotics.utexas.edu/rrg/learn_more/history/#definition; https://en.wikipedia.org/wiki/Three_Laws_of_Robotics#First_Law_modified)

(remember the discussion from the beginning of the course and the "Humans"?)

13

Three laws of *robotics* (Asimov)

(later added a 'zeroth law')

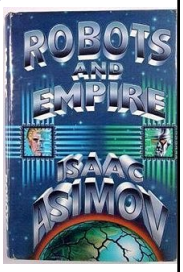
- *From early session of this course (on "Humans")….*
  - *How to replicate something we do not understand?*
  - *human emotions, love, dream, conciseness, fear, anger, violence, memory…*

(abstractions…)

14

Three laws of *robotics* (Asimov)

(later added a 'zeroth law')



Robots and Empire, novel by Asimov (1985)

- R. Daneel Olivaw
  - *a fictional robot (the "R" initial stands for "Robot," a naming convention in Asimov's future society)*
- The plot (robots discovers danger to humanity)…
  - *R. Daneel discovers the plan of roboticist Amadiro to destroy the population of the Earth, but are hampered by the First Law of Robotics "A robot may not injure a human being, or, through inaction, allow a human being to come to harm"*
- R. Daneel (Asimov) tries to resolve...
  - *R. Daneel meannwhile formulates an additional Zeroth Law: "A robot may not injure humanity, or, through inaction, allow humanity to come to harm"*
- But then "*the humanity*" is an "abstraction"

From: https://en.wikipedia.org/wiki/Robots_and_Empire
https://en.wikipedia.org/wiki/Three_Laws_of_Robotics#First_Law_modified

15

Definitions of robot

International Organization for Standardization (ISO 8373):

• "An automatically controlled, reprogrammable, multipurpose, manipulator programmable in three or more axes, which may be either fixed in place or mobile for use in industrial automation applications."

Robotics Institute of America (1979):

• A re-programmable multi-functional manipulator designed to move materials, parts, tools, or specialized devices through variable programmed motions for the performance of a variety of tasks.

Encyclopedia Britannica:

• A "Robot" is any automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner. By extension, robotics is the engineering discipline dealing with the design, construction, and operation of robots.

16

Definitions of robot

Merriam-Webster:

Function: *noun*
Etymology: Czech, from *robota* compulsory labor; akin to Old High German *arabeit* trouble, Latin *orbus* orphaned
Date: 1923
1 a: a machine that looks like a human being and performs various complex acts (as walking or talking) of a human being; *also* : a similar but fictional machine whose lack of capacity for human emotions is often emphasized **b:** an efficient insensitive person who functions automatically
2: a device that automatically performs complicated often repetitive tasks
3: a mechanism guided by automatic controls

George Bekey in "Autonomous Robots"

*A machine that **senses**, **thinks**, and **acts**.*
A robot must have 1) sensors (to obtain info from environment); 2) ability to process (from reflex to cognitive); 3) actuators (to exert forces upon environ.); Can be industrial, mobile, flying, submersible, humanoid, or household.

17

A robot

*A machine that **senses**, thinks, and acts.*

Sensors:

- *Internal/external environment; imitate biological sensors*
- **Exteroceptive sensors**
  - *For obtaining info from the external environment*
    - *Vision, hearing, olfaction, touch, and taste*
  - *Biological*
    - *Coupling of sensory apparatus with the brain*
    - *Often highly tuned – frog & insects*
  - *Robot*
    - *light sensor with 'visual cortex', camera w/ software for edge detection*
    - *Can be specialized, too.*

18

A robot (cont.)

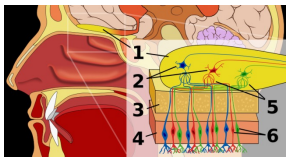
*A machine that **senses**, thinks, and acts.*

Sensors:

- *Internal/external environment; imitate biological sensors*
- **Exteroceptive sensors**

- **Limitations**
  - *superiority of biological (insect's olfactory sensors – pheromones; hawk's eye)*
  - *Eagles can see eight times as far as the sharpest human eyes can*
  - *superiority of robot sensors (Geiger counter, ultraviolet/microwave radiation)*

Taken from http://en.wikipedia.org/wiki/Olfactory_system



1: Olfactory bulb 2: Mitral cells 3: Bone 4: Nasal Epithelium 5: Glomerulus 6: Olfactory receptor cells

19

A robot (cont.)

*A machine that **senses**, thinks, and acts.*

Sensors:

- *Internal/external environment; imitate biological sensors*

- **Proprioceptive sensors**
  - *Sensors relating to stimuli produced/perceived within an organism, especially those connected with the position and movement of the body*
  - *For obtaining info from the internal environment*
  - *No analogy with biological*
    - *battery voltage, wheel rotation*
  - *Yes analogy with biological*
    - *joint angles, leg motor currents – tendons, muscle spindles, artificial muscles*

20

A robot (cont.)

*A machine that senses, thinks, and **acts**.*

Actuators:

The key difference between a robot and a softbot

- *Interaction with the environment via actuators*
- *Artificial muscles*
- *Electric (motors that provide locomotion – wheels/legs/arms, precision steppers)*
- *Pneumatic and hydraulic (industry, manipulation)*

A robot (cont.)

*A machine that senses, **thinks**, and acts.*

Intelligence in robots:

Coming from software rather than hardware:

(well…there are smart materials like piezoelectric, shape memory, chromo/photo active…material science…)

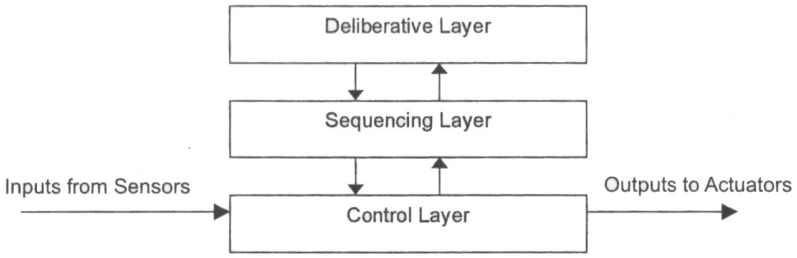
- *Sensor processing (preprocessing at site)*
- *Reflex behaviors (rapid reaction path w/o involving higher CNS)*
  - *touching a hot object, knee-jerk, legged insect touching an obstacle)*
- *Special purpose programs (navigation, localization, obstacle avoidance)*
- *Cognitive functions (reasoning, learning, planning)*

Intelligence in robots (cont.)

Coming from software rather than hardware:

Software Architecture

- *Hierarchical*
  - *Deliberative (human coupled)*
  - *Sequencing (intermediate, supervisory)*
  - *Control layer (I/O, sensor/actuator)*



23

Biological inspiration

Control and regulation in biological systems

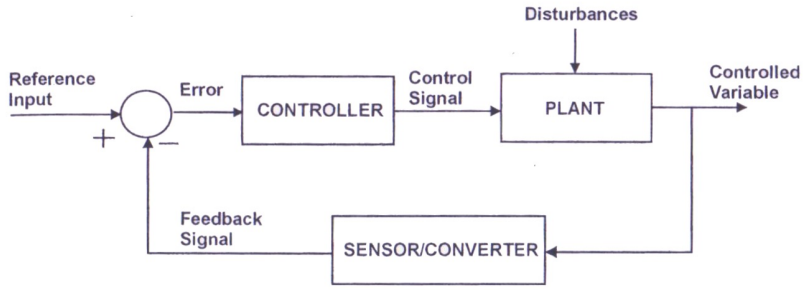
- *Important for robot design*
- *The Wisdom of the Body 1932, "Homeostasis" 1939*
  - *physiologist at Harvard medical school, military physician in World War I*
  - *Theorized that the sympathetic nervous system and the adrenal gland work together as a unit to maintain homeostasis in emergencies.*
  - *Known by body's fight-or-flight response, also known as the acute stress response*
  - *(HOH-mee-oh-STAY-sis) A state of balance among all the body systems needed for the body to survive and function correctly*
- *Mammals' body internal environment (core temp, BP, biochemical substance concentration)*
- ***Regulation** - "normal" and other ranges of values (depending on state)*
- *Example?*
  - *involving sensing, neural control, chemical control, actuation*

24

Engineering counterpart

Feedback control system

- *Reference value, sensor, comparator, controller (actuators/effectors)*



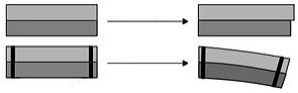
25

Engineering vs. biological

Feedback control system

(reference values, comparator, sensors, actuators, adaptivity, control redundancy, multipurpose control)

- ***Reference value***
  - ***Fixed***
    - *atomic clock, high-quality quartz crystal*
  - ***Relative***
    - *absolute ref. point does not exist, rather dynamical changes*
    - *BP (**change** in blood volume through kidneys, water passed to bladder for excretion)*
    - *thermostat (fixed reference or bi-metal)*



Taken from http://en.wikipedia.org/wiki/Bi-metallic_strip

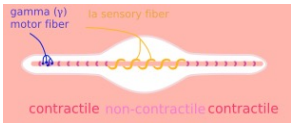
26

Engineering vs. biological (cont.)

Feedback control system

(reference values, comparator, sensors, actuators, adaptivity, control redundancy, multipurpose control); Biology – homeostasis; Mechanics – equilibrium

- ***Comparator***
  - *often negative feedback (AC, glucose level, VSA, ABS)*
  - *instability (Nyquist stability criterion)*
- ***Sensors***
  - *or receptors (biol.)*
  - *examples*
    - *baroreceptors (BP sensors)*
    - *osmoceptors (hypothalamus, osmotic BP)*
    - *olfactory, tactile*
    - *muscle spindle – convey length info to the CNS via sensory neurons (efferent & afferent fibers, neurons)*
    - *example (stretch reflex - skiing)*
    - *velocity and degree of stretching -> neuron firing frequency*
    - *other stretch receptors (blood volume sensors in large veins)*



Taken from http://en.wikipedia.org/wiki/Muscle_spindle

27

Engineering vs. biological (cont.)

Feedback control system

(reference values, comparator, sensors, actuators, adaptivity, control redundancy, multipurpose control); Biology – homeostasis; Mechanics – equilibrium

- ***Actuators (effectors)***
  - *muscles (biol.)*
  - *contraction of arterial vessels (arterioroles) -> increase in upstream BP*
- ***Adaptivity***
  - *change of the feedback control system characteristics*
- ***Control redundancy***
  - *multiple control loops*
  - *BP example*
    - *osmotic pressure blood volume control*
    - *baroceptors in arteries (antidiuretic hormone ADH secretion)*
    - *stretch receptors*
    - *kidney sodium concentration control (aldosterone hormonal adrenal cortex secretion)*

28

Engineering vs. biological (cont.)

Feedback control system

(reference values, comparator, sensors, actuators, adaptivity, control redundancy, multipurpose control); Biology – homeostasis; Mechanics – equilibrium

- ***Multipurpose controllers***
  - *control of one var. affects another one (desired/undesired)*
  - *examples (perspiration – body temp., fluid volume; respirations – O2, heat, water)*
- ***Robot control systems***
  - *typically much simpler than homeostatic systems*
  - *typically of reference value-controller-sensor-comparator type*
  - *a key element in achieving living system resemblance*
  - *a key part of more sophisticated robots (humanoid?)*

Robot Learning

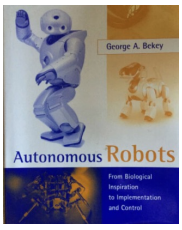
Robot learning

- *Types of learning*
- *Learning controllers*
- *Adaptive learning*

Robot learning (G. Bekey)

Types

- *Reinforcement learning*
- *Neural network learning*
- *Evolutionary learning*



Intelligence?

- *Reasoning, thinking, solving, learning*
- *Improve (speed, accuracy, ability)*
- *Reduce errors*

Bekey's Textbook (Connel & Mahadevan 1993)

What should robots learn?

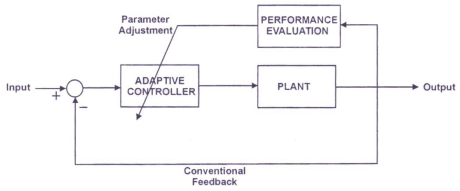
1. *Hard-to-program knowledge (ex. imitation)*
2. *Unknown information (ex. search & rescue robots)*
3. *Changing environments (dynamic, adaptive, in time, in future):* ***learning***

Adaptive Learning

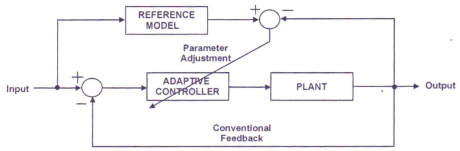
Adaptive learning (learning controllers)

Nonlinear and learning in new scenarios (3rd one of C-M types)

1. *performance evaluation (towards system performance optimum)*
2. *reference model (receives the same input but produces desired response)*



(a)



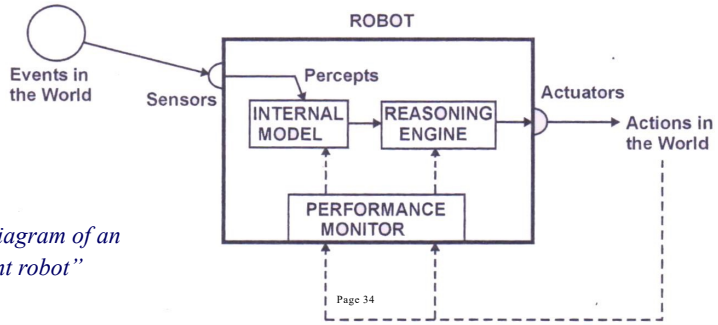
(b)

33

Adaptive Learning

Adaptive learning – model

- *Reasoning engine*
- *Performance evaluation*
- *Training vs. testing data, i.e. task set*
  - *no circular reasoning, i.e. testing on training data*
- *Supervised vs. unsupervised learning*
- *Percepts subject to noise (sensor imperfection)*



Generic diagram of an "intelligent robot"

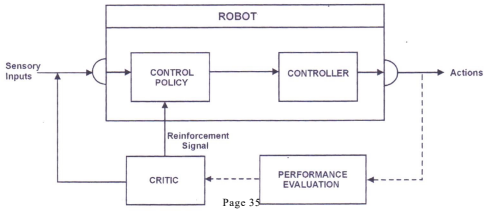
34

Reinforcement learning

Learning through reinforcement

Good robot – bad robot ☺

- *Trial and error (no correct behavior indication given)*
  - *lengthy (large number of trials)*
- *Supervised (immediate feedback)*
  - *set of tasks needs to be prepared upfront*
  - *set of tasks needs to cover extensive list of possible scenarios*
- ***Robot needs to construct a strategy for maximizing the reward***
  => control policy for maximizing the reward
- *Example (Bekey): negative-reinforcement policy (simple critic, +1 or -1)*



35

Reinforcement learning

- *Very **different** from intelligent agents (**model**)!*
- *Box pushing example (Bekey textbook)*
- *Obelix*
  - *sensors equipped*
  - *Credit assignment problem*
  - *Behavior based software architecture*
- Q-Learning
- Strategy learning examples
  - *Tower of Hanoi puzzle*
  - *Switch the frogs*

36

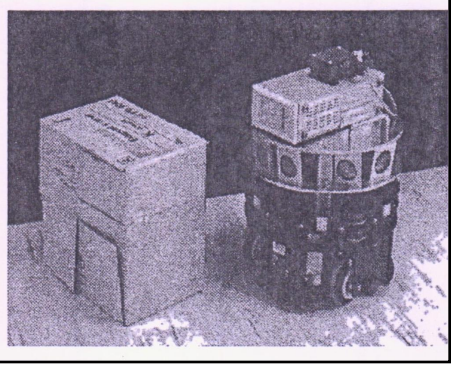
Reinforcement learning

Example (Bekey): negative-reinforcement policy

- Two types sensors (sonar & bumper)
  - *robot decides which input is more important in avoiding negative reinforcements*
- Credit assignment problem
  - ***Structural**: which sensory data associated with which reward (+, -)*
  - ***Temporal**: the sequence of actions associated with the reward (+, -)*

*Obelix – a **box pushing** problem*

(Connel & Mahadevan 1993)



37

Reinforcement learning

Example (Bekey): negative-reinforcement policy

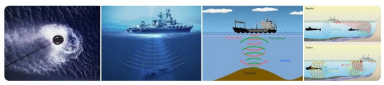
- Obelix – a box pushing problem
  - *Available:*
    - *8 sonars (4 forward, 2 on each side), NEAR or FAR bit (16 bit signals)*
    - *IR sensor (facing forward, BUMP bit)*
    - *Motor current monitor (threshold T, STUCK bit)*

      => Total=18 bits
  - *if IR activated -> bit=BUMP*
  - *if motor current>T, bit=STUCK*
  - *possible: 18 bits (>256,000 perceptual states, 2^18 = 256k)*
  - *Goal: choose one of 5 actions (out of 256k): move forward; turn in place right or left (22° or 45°)*
  - *reward=1, if box pushed (BUMP in 2 successive times while moving forward & not STUCK); 0 otherwise;*

An infrared sensor (IR sensor) is a **radiation-sensitive optoelectronic component with a spectral sensitivity in the infrared wavelength range 780 nm** ... 50 µm. IR sensors are now widely used in motion detectors, which are used in building services to switch on lamps or in alarm systems to detect unwelcome guests.

https://www.infratec-infrared.com › glossary › infrared-se...

Infrared sensor - IR sensor I Sensor division knowledge



Sonar (**sound navigation and ranging**) is a technique that uses sound propagation (usually underwater, as in submarine navigation) to navigate, measure distances (ranging), communicate with or detect objects on or under the surface of the water, such as other vessels.

38

Reinforcement learning

Example (Bekey): negative-reinforcement policy (cont.)

- Problems:
  - *Sensory input space too large*
  - *Simple reward function -> infrequent rewards*
  - *If more complex reward function attempted:*
    - *if emphasis on avoiding getting stuck -> Obelix avoids all contacts*
    - *if emphasis on contact -> Obelix pushing walls (local minima problem!)*
  - No history included (*if it gets stuck, it continues to stay stuck, forgets that it has just pushed wall*)
  - Sensors too simple (*sonar sensors see the wall and the box the same*)
- Solution
  - *Decomposition of the problem (**behavior based software architecture – credited to Rodney Brooks et. al.**)*
  - *Find a box*
  - *Push the box*
  - *Avoid getting stuck*

39

Reinforcement learning

Example (Bekey): negative-reinforcement policy (cont.)

Solution (decomposition of the problem)

Behavior (subsumption) based software architecture

Three behaviors:

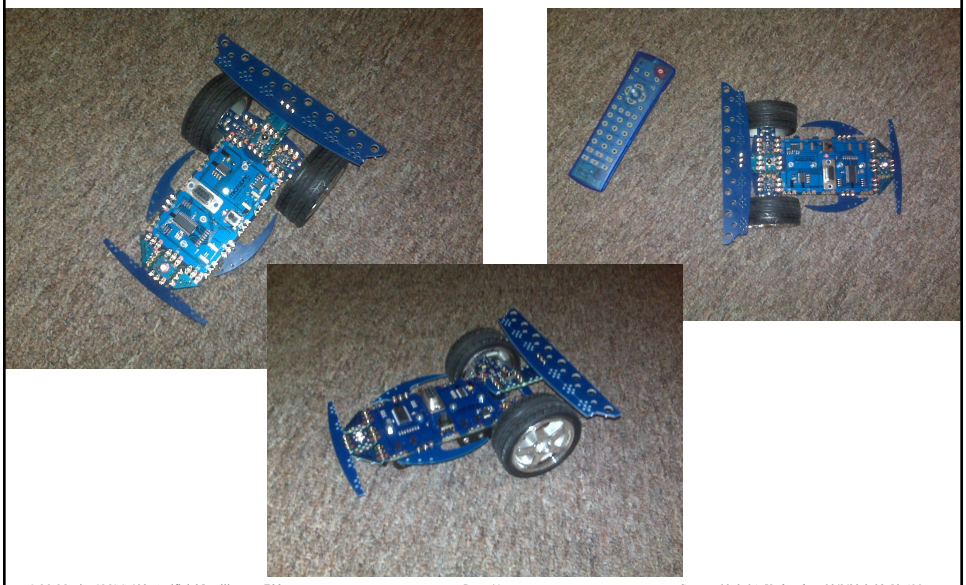
1. *Find a box – reward whenever NEAR on (negative if off)*
2. *Push the box – reward whenever BUMP on (negative if off, or if pushing a wall, or pushing a box against the wall)*
3. *Avoid getting stuck – reward if goes out of STUCK (negative if in STUCK)*

Results in:

- *Sensory input set for each behavior, each behavior has its own rewards*

40

Viper



41

Reinforcement learning

42

Reinforcement learning

- *Box pushing example (Bekey textbook)*
- *Obelix*
  - *sensors equipped*
  - *Credit assignment problem*
  - *Behavior based software architecture*
- Q-Learning
- Strategy learning examples
  - *Tower of Hanoi puzzle*
  - *Switch the frogs*

Q-Learning

Q-learning

- (Environment) ***model-free*** reinforcement learning techniques
- Learning of the most appropriate behaviors
  - ***Highest possible combination of immediate & expected future reward***
  - *S - set of all states, A – set of all actions;*
  - *Q-learning is a control policy of reward for a given action in a given state by the function Q(s,a)*
  - *For a robot to **learn complete control policy**, the goal is to **learn the values Q over the whole state-action space (SxA):***

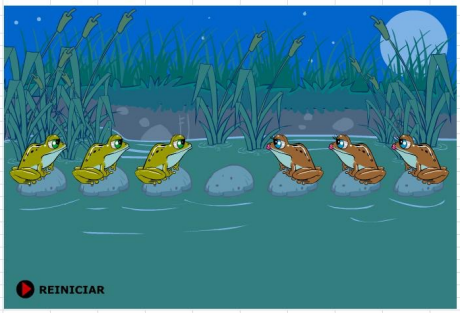
$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha(r + \lambda E(y) - Q_t(s,a))$$

$0 < \lambda \leq 1$

  - *new value = the current value + α* error*
  - *until current and future rewards = Q, error exists*
  - *where r – current reward, E(y) – expected reward (utility), y – state resulting from action a, α – learning rate, λ – discount factor*
- ***Method results in a state-action (SxA) table that robot can use to maximize Q(s, a)***

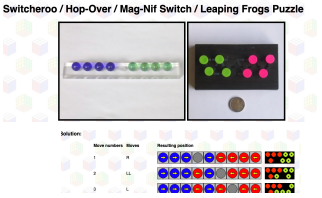
Q-Learning

Example - *Switch the frogs to the opposite side within shortest time possible*



https://data.bangtech.com/algorithm/switch_frogs_to_the_opposite_side.htm

https://nrich.maths.org/content/00/12/game1/frogs/index.html#/student/2/2



https://www.jaapsch.net/puzzles/froghop.htm

45

Q-Learning

Example - *Tower of Hanoi puzzle*

- Also known as *The Towers of Benares*
- *French mathematician Edouard Lucas in 1883 (*Prof. N. Claus)
- *Lucas became known for his work with the Fibonacci number sequence*
  - *The Fibonacci-related Lucas number series*
    - *Each number is the sum of the two preceding numbers*
    - *Break down into smaller problems until cannot reduce anymore*
    - *Recursive theory*
- *Question: Given optimal conditions, how many pairs of rabbits can be produced from a single pair of rabbits in one year?*
  - *This thought experiment dictates that the female rabbits always give birth to pairs, and each pair consists of one male and one female.*
  - *1, 1, 2, 3, 5, 8, 13, 21….*
  - *Golden ratio (golden number): 1.618034*

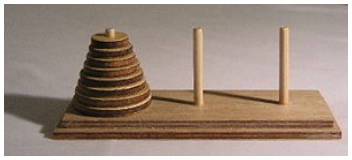
Taken from: http://science.howstuffworks.com/math-concepts/fibonacci-nature.htm

46

Q-Learning

Example - *Tower of Hanoi puzzle*

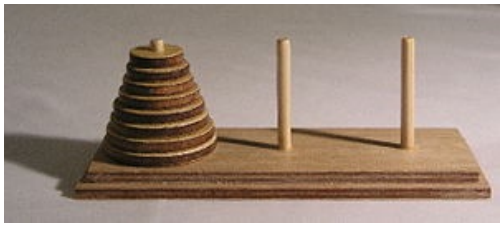
- Also known as *The Towers of Benares*
- *French mathematician Edouard Lucas in 1883 (*Prof. N. Claus)
- *The "legend" which accompanied the game stated that in Benares, during the reign of the Emperor Fo Hi, there was a temple with a dome which marked the center of the world. Within the dome, priests moved golden disks between diamond needlepoints, a cubit high and as thick as the body of a bee. God placed 64 gold disks on one needle at the time of creation. It was said that when they completed their task, the universe would come to an end.*
  - *Since it would take at least 264 -1 moves to complete the task, we're safe for now. Assuming one move per second, and no wrong moves, it would take almost 585,000,000,000 years to complete.*



Q-Learning

Example - *Tower of Hanoi puzzle*

- The objective is to move all disks to another rod (no disk on smaller disk)
- *Play here:*
  - *http://www.superkids.com/aweb/tools/logic/towers/*
  - *http://www.dynamicdrive.com/dynamicindex12/towerhanoi.htm*
  - *http://people.revoledu.com/kardi/tutorial/ReinforcementLearning/Tower-of-Hanoi.htm*



Min number of moves is 2^n - 1, n- number of disks.

Extra Games

49

Another AI game…

Solve and explain the **cannibals and missionaries problem**.

- Note: At most two riders can ride in a boat. Also, boat cannot move on its own (hence the oars)…
- *Play, read, watch:*
  - *https://javalab.org/en/boat_puzzle_en/*
  - *https://en.wikipedia.org/wiki/Missionaries_and_cannibals_problem*
  - *https://www.youtube.com/watch?v=enCyv5UbDhA*

50