Chapter 2. Clues

Lecture Notes on Artificial Intelligence, Spring 2012

Summarized by Seok, Ho-Sik

Biointelligence Laboratory
School of Computer Science and Engineering
Seoul National University

http://bi.snu.ac.kr
Contents

2.1 From Philosophy to Logic
   Historical attempts for mechanization of reasoning

2.2 From Life Itself
   Insights from neuronal architectures and nature
   2.2.1 Neurons and the Brain
   2.2.2 Psychology and Cognitive Science
   2.2.3 Evolution
   2.2.4 Development and Maturation
   2.2.5 Bionics

2.3 From Engineering
   Primitive computers and foundation of modern computers
   2.3.1 Automata, Sensing, and Feedback
   2.3.2 Statistics and Probability
   2.3.3 The Computer

Appendix
Overview of Chapter 2

- Clues about what might be needed to make machines intelligent are scattered abundantly throughout philosophy, logic, biology, psychology, statistics, and engineering.

- With gradually increasing intensity, people set about to exploit clues from these areas in their separate quests to automate some aspects of intelligence.
Chapter 2. Clues

2.1 From Philosophy and Logic
From Philosophy and Logic (1/3)

- **Syllogism**
  1. All humans are mortal. (stated)
  2. All Greeks are humans. (stated)
  3. All Greeks are mortal. (result)

- **How one might automate reasoning**
  1) Patterns of reasoning can be economically represented as *forms* or *templates* (employing generic symbols).
  2) After the general symbols are replaced by ones pertaining to a specific problem, one only has to “turn the crank” to get an answer.
Gottfried Wilhelm Leibniz

- Lingua characteristica or universal language
  - Knowledge could be built from a smaller number of primitive ones.
  - Alphabet for human thoughts: if the items in the alphabet were represented by numbers, then a complex proposition could be obtained from its primitive constituents → Calculus ratiocinator (calculus of reasoning)
  - Question: how to discover the component of the primitive “alphabet”?

Charles Stanhope

- Stanhope’s box
  - An example:
    
    \[
    \text{Eight of ten A’s are B’s; Four of ten A’s are C’s; Therefore, at least two B’s are C’s.}
    \]

  - Stanhope’s box → push the red slide (representing B) eight units across the window (representing A) and the gray slide (representing C) four units from the opposite direction. The two units that the slides overlapped represented the minimum number of B’s that were also C’s.
George Boole
- Some kinds of logical reasoning could be performed by manipulating equations representing logical propositions → a clue about the mechanization of reasoning.
- Shortcoming: propositions p, q, and so on were “atomic” they don’t reveal any entities internal to propositions.

Friedrich Ludwig Gottlob Frege
- Begriffsschrift or concept writing: propositions along with their internal components could be written in a kind of graphical form.
- The forerunner of “predicate calculus” and “semantic network”.
- Sentences expressing information to be reasoned about could be written in unambiguous, symbolic form.
Chapter 2. Clues

2.2 From Life Itself
2.2.1 Neurons and the Brain (1/3)

- "Neuron doctrine"
  - Living cells called "neurons" together with their interconnections were fundamental to what the brain does.
  - Neuron
    - $10^{10}$ neurons in the human brain.
    - Soma: a central part.

- Dendrite: incoming fibers.
- Axon: outgoing fibers.
- Terminal button: the axon of a neuron has projections that come very close to one or more of the dendrites of other neurons.
- Synapse: a gap between the terminal button of one neuron and a dendrite of another (20nm).
2.2.1 Neurons and the Brain (2/3)

- **Electrochemical action**
  - A neuron may send out a stream of pulses down its axon.
  - When a pulse arrives at the synapse adjacent to a dendrite of another neuron, it may act to excite or to inhibit electrochemical activity of the other neuron across the synapse.

- **Firing factors**
  - How many and what kinds of pulses (excitatory or inhibitory) arrive at the synapses of its various incoming dendrites.
  - The efficiency of those synapses in transmitting electrochemical activity.

- **A neuron as a “logic unit” (McCulloch-Pitt model)**
  - Networks of simple models of neurons could perform all possible operations.
  - McCulloch-Pitt “neuron”
    - A mathematical abstraction with inputs and outputs.
    - The neural elements can be connected together into networks such that the output of one neural element is an input to others and so on.
Networks of McCulloch-Pitts neural elements
2.2.1 Neurons and the Brain (3/3)

- Donald O. Hebb
  - when an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.
  - Cell assembly: the phenomenon of “firing together” tended to persist in the brain and was the brain’s way of representing the perceptual event that led to a cell-assembly’s formation.
2.2.2 Psychology and Cognitive Science (1/2)

- **Behaviorists**
  - Psychology should be a science of behavior, not of the mind
  - B. F. Skinner
    - What could be objectively measured? → specific behavior in reaction to specific stimuli.
    - Reinforcing stimuli: one that rewards recent behavior and tends to make it more likely to occur (under similar circumstances) in the future.

- **Reinforcement learning of an “artificial animal” in a game**
  - The more the opponent repeats the same move after any given stimulus, the more the animal model becomes ‘conditioned’ to that move.

- **Noam Chomsky**
  - How can a person produce a potentially infinite variety of previously unheard and unspoken sentences having arbitrarily complex structure (as indeed they can do) through experience alone?
  - Linguistic abilities must be innate – not learned.
  - Universal grammar accounts for much of the ability to learn and use languages.
The magical number seven

George A. Miller

The “immediate memory” capacity of humans was approximately seven “chunks” of information.

The span of immediate memory seems to be almost independent of the number of bits per chunk.

TOTE (Test – Operate – Test – Exit) unit: a specific internal mechanism responsible for behavior.
2.2.3 Evolution (1/2)

- **Clues from evolution**
  - The processes of evolution itself – namely, random generation and selective survival – might be simulated on computers to produce the machines.
  - Those paths that evolution followed in producing increasingly intelligent animals can be used as a guide for creating increasingly intelligent artifacts.

- **Genetic algorithms**
  - John Holland
  - Evolving chromosomes (strings of binary symbols) that were better and better at solving the problem.
  - Individual chromosomes are subjected to “mutations” and “crossover (pairs of chromosomes were combined to make a new chromosome)”.
2.2.3 Evolution (2/2)

- **Machina speculatrix**
  - W. Grey Walter
  - Controlled by “brains” consisting of very simple vacuum-tube circuits.
  - Sensing environments with photocells and controlling wheel motors.

Grey Walter (top left), *Machina speculatrix* (top right), and its circuit diagram (bottom)
2.2.4 Development and Maturation

- John Piaget
  - There exists a set of stages in young children in the maturation of thinking abilities.

- Ontogenetic
  - Start with a machine that is able to do what an infant to do and then design machines that can mimic the abilities of later developmental stages.

- Phylogenetic
  - Use of simulated evolution.

2.2.5 Bionics

- The field that learns lessons from nature to apply to technology.

- Bionics aims to take advantage of millions of years of evolution of living systems during which they adapted themselves for optimum survival. One of the outstanding successes of evolution is the information processing capability of living systems [the study of which is] one of the principal areas of Bionics research.
Chapter 2. Clues

2.3 From Engineering
2.3.1 Automata, Sensing, and Feedback (1/2)

- **Feedback control**
  - Feeding some aspect of a machine’s behavior back into the internals of the machine.
  - Negative feedback: if the aspect of behavior that is fed back acts to diminish or reverse that aspect of behavior.
  - Positive feedback: If the aspect of behavior that is fed back acts to increase or accentuate that aspect of behavior.

A verge-and-foliot mechanism (left) and automata at the Munich Glockenspiel (right)
- Their behavior was fully automatic, requiring no human guidance.
- They did not perceive their environments.

Jacquard loom
Watt’s flyball governor
As the speed of the engine increases, the balls fly outward, which causes a linking mechanism to decrease air flow, which causes the speed to decrease, which causes the balls to fall back inward, which causes the speed to increase, and so on, resulting in an equilibrium speed.

W. Ross Ashby, Warren McCulloch, Grey Walter, and Norbert Wiener
2.3.1 Automata, Sensing, and Feedback (2/2)

- **Self-organizing systems**
  - Many unorganized combinations of simple parts, including combinations of atoms and molecules, respond to energetic “jostling” by falling into stable states in which the parts are organized in more complex assemblies.
  - Self-organization is not a property of an organism itself in response to its environment and experience.
  - A property of an organism and its environment *taken together.*
2.3.2 Statistics and Probability

- Nearly all reasoning and decision making take place in the presence of uncertainty.

- Bayes’s rule

\[ p(x \mid y) = \frac{p(y \mid x)p(x)}{p(y)} \]

- Example: Determine a tone (A or B) given the signal \( y \) and actual tone \( x \).

  - Decides in favor of tone A if \( p(y \mid x = A)p(x = A) / p(y) \) is greater than \( p(y \mid x = B)p(x = B) / p(y) \). Otherwise, decide in favor of tone B.

  - Decides in favor of tone A if \( p(y \mid x = A)p(x = A) \) is greater than \( p(y \mid x = B)p(x = B) \). Otherwise, decide in favor of tone B.

  - If we know the a priori probabilities \( p(x = A) \) and \( p(x = B) \), it is sufficient to calculate \( p(y \mid x) \) (likelihood) for A and B (Maximum-likelihood).
2.3.3 The Computer – Early Computational Devices

- Physical device
  - Wilhelm Schickard – addition, subtraction of six-digit numbers for use in calculating astronomical tables.
  - Blaise Pascal – An adding machine capable of performing automatic carries from one position to the next.
  - Gottfried Leibniz – “Step Reckoner” capable of addition, subtraction, and multiplication.
  - Charles Babbage
    - “Difference Engine” calculating mathematical tables using the method of finite differences.
    - “Analytical Engine” for general computation.
  - Ada Lovelace: “world’s first programmer” due to her role in devising programs for the Analytical engine.
2.3.3 The Computer – Computation Theory

- Alan Turing
  - “Logical computing machine (LCM)” → “Turing machine”.
  - A hypothetical computational device
  - An infinite tape (cells 0 or 1 and a tape drive) + a read-write head + a logic unit (changing state, commanding the write function on a cell, moving the tape, terminating operation)
  - The input is written on the tape and the output ends up being printed on the tape.
  - Turing showed that one could encode on the tape itself a prescription for any logic unit specialized for a particular problem and then use a general-purpose logic unit for all problems.
  - The encoding for the special-purpose logic unit can be thought of as the “program” for the machine.
2.3.3 The Computer – Digital Computers

- Claude Shannon
  - 1937 master’s thesis
    - Boolean algebra and binary arithmetic could be used to simplify telephone switching circuits.
    - Switching circuits could be used to implement operations in Boolean logic.

- Konrad Zuse, J. Presper Eckert, J. W. Mauchly
  - The idea of storing a computer’s program along with its data in the computer’s memory allows the program to change itself by changing appropriate parts of the memory where the program is stored.
  - Konrad Zuse’s “Z3” (2,400 electromechanical relays)

- EDVAC
  - Von Neumann architecture.
  - Separating the (task-specific) stored program (software) from the general purpose hardware circuitry (hardware).

- The stored-program digital computer
  - It can be used for any computational purpose.
2.3.3 The Computer – “Thinking” Computers (1/2)

- Alan Turing’s 1950 paper
  - Importance 1) Turing thought that the question “can a machine think?” was too ambiguous. → Turing proposed that the matter of machine intelligence be settled by “the Turing test”
    - Turing test – attempting to have a computer “in the other room” fool an interrogator into believing the computer is a human.
  - Importance 2) Handling of arguments that people might raise against the possibility of achieving intelligent computers (see appendix).
  - Importance 3) Suggestion about how we might go about producing programs with human-level intellectual abilities.
    - Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain.

2.3.3 The Computer – “Thinking” Computers (2/2)

- **Allen Newell and Herb Simon**
  - Digital computer’s universality meant that it could be used to mechanize intelligence in all its manifestations – provided it had the right software.
  - “Physical Symbol System Hypothesis”
    - A physical symbol system has the necessary and sufficient means for intelligent action
      - Appropriately programmed digital computers would be capable of intelligent action.
      - Because humans are capable of intelligent action, hypothesis, physical symbol systems.

- The imagined Turing machine and the very real digital computer are symbol systems in the sense Newell and Simon meant the phrase.

- “The symbol system hypothesis implies that the symbolic behavior of man arises because he has the characteristics of a physical symbol system. Hence, the results of efforts to model human behavior with symbol systems become an important part of the evidence for the hypothesis, and research in artificial intelligence goes on in close collaboration with research in information processing psychology, as it is usually called.”
Appendix - Turing mentions (1/2)

(1) The Theological Objection
- Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any other animal or to machines. Hence no animal or machine can think.

(2) The 'Heads in the Sand' Objection
- “The consequences of machines thinking would be too dreadful. Let us hope and believe that they cannot do so”.

(3) The Mathematical Objection
- There are a number of results of mathematical logic that can be used to show that there are limitations to the powers of discrete-state machines.

(4) The Argument from Consciousness
- These arguments take the form, “I grant you that you can make machines do all the things you have mentioned but you will never be able to make one to do X.”

(5) Arguments from Various Disabilities
- These arguments take the form, “I grant you that you can make machines do all the things you have mentioned but you will never be able to make one to do X.”
(6) Lady Lovelace's Objection

- Our most detailed information of Babbage's Analytical Engine comes from a memoir by Lady Lovelace. In it she states, “The Analytical Engine has no pretensions to originate anything. It can do whatever we know how to order it to perform”.

(7) Argument from Continuity in the Nervous System

- The nervous system is certainly not a discrete-state machine. A small error in the information about the size of a nervous impulse impinging on a neuron may make a large difference to the size of the outgoing impulse. It may be argued that, this being so, one cannot expect to be able to mimic the behavior of the nervous system with a discrete-state system.

(8) The Argument from Informality of Behavior

- It is not possible to produce a set of rules purporting to describe what a man should do in every conceivable set of circumstances.

(9) The Argument from Extra-Sensory Perception.