Intelligence implies….

- Reasoning (plan)
  - Modelling the world: objects and interactions
  - Inferring implicit relationships
  - Problem solving, search for an answer, planning
- Interaction with the outside world (sense & act)
  - Perception: the inference of objects and relationships from what sensors deliver.
    - Sensors deliver "arrays of numbers"
  - Intelligent behavior
- Learning
  - Systems that acquire and **incorporate** new data
  - Systems that take instructions (from *us*)
  - Knowing what we are doing **wrong**

Early Chronology

- George Boole, Gottlob Frege, Alfred Tarski: human thought
- Alan Turing, John von Neumann, Claude Shannon:
  - Cybernetics
  - Equivalence/analogy between **computation** and **thought !!!**
- AI: The 40s and 50s
  - McCulloch and Pitts: Describe neural networks that could compute any computable function
  - Samuels: Checker playing machine that learned to play better.
  - "Dartmouth Conference" (1956) : McCarthy: coined term "Artificial Intelligence"
  - McCarthy:Defined LISP.
  - Newell and Simon: The Logic Theorist. It was able to prove most of the theorems in Russell and Whiteheadís Principia Mathematica. Bounded Rationality, Logic Theorist becomes General Problem Solver.

- Early Successes
- Minksy: microworlds
- Evan's ANALOGY solved geometric analogy problems that appear on IQ tests
- Bobrow's STUDENT solved algebra world problems
- Gelernter: Geometry Theorem Prover used axioms plus diagram information.
- Early success with neural networks.

Expert Systems and the Commercialization of AI

- Buchanan and Feigenbaum: DENDRAL (1969)
- MYCIN (1976): diagnose infections.
- LUNAR (1973): First natural language question/answer system used in real life
- Rejuvenation of neural nets
  - In theory, they can learn almost any function.
    - In practice, it might take a millenium.
  - Neural nets, while having obvious limitations, have surpassed hand-crafted systems in some key domains.

State-of-the-Art

- Almost grandmaster chess (ask me about checkers).
- Real-time speech recognition
- Expert systems "aren't really AI anymore", but many exist

The Bad News

- Heavily oversold, with ensuing backlash.
- Almost every AI problem in NP-complete.
  – Lighthill report (1973).
- Perceptrons (a kind of neural network) shown to have extremely limited representation ability (Minsky and Papert).
- Some of AI seen as poorly formalized hackery or an mathematical self-indulgence..

An early intelligent system: the brain

- "Intelligent" processing in the brain is carried out by neurons, mainly in the cerebral cortex.
  - Roughly 10^{12} neurons (10^{11} if you participated in frosh week) and thousands of *connections* per neuron.
  - "Clock speed" (refractory period): 1 to 10 milliseconds
  - Processing involves massive parallelism and distributed representation.
- 5 to 7 rule
  - You can simultaneously conceptualize only 7 objects at once.

Comparison

- Computers
  - Roughly 10 million transistors per chip
  - Parallel machines: hundreds of CPU elements, 1010 bits of RAM
  - Clock speed: roughly 1 nanosecond
  - Recall rate (for stored data) appears *much* faster.
- Does the different hardware imply that fundamentally different approaches must be used?
  - Neural nets people suggest “no”.
  - Some suggest “yes”.

What is intelligence?

- Stock answer: "the ability to learn and to solve problems" [Webster's]
  - The ability to adapt to new situations.
- Your answers (paraphrased):
  - "The ability to laugh at humorous situations."
  - "Understanding how other agents behave…"
  - "The ability to analyze and solve a problem" *****
  - "The ability to work with abstract concepts" *****
  - "The ability to recognize patterns…"
  - "The ability to use new information/experience to make decisions"
  - "Something that defines humanity."
  - "The ability to solve an **under-specified** problem."

What does the future hold?

- Many of you thought artificially intelligent systems were a long way off.
  - “Never”
  - “In some timescale comparable to the evolution of intelligence in animals”
  - “”In 50 years”
  - “Not very soon”
- Some of the same people said things like:
  - “Intelligence is the ability to solve problems that would be complex for a human being.”

Playing Chess

- In 1997 the computer "Deep Blue" played the human world chess champion "Garry Kasparov"
  - (whom some have claimed is the best chess player in history!)
  - DB: 200 board positions per move.
    - 11 ply
  - GK: 7 ply?

Monty Newborne at SOCS/McGill has been a pioneer in computer chess. Moderated the DB/GK match.

“We’ll never really have artificial intelligence”

- Garry Kasparov:

  “I could feel -- I could smell -- a new kind of intelligence across the table.”

- Drew McDermott

“Saying Deep Blue doesn’t really think about chess is like saying an airplane doesn’t really fly because it doesn’t flap it’s wings.”

[Yale CS, quoted in NY Times, May 1997]

- Robbins’ problem:
  - In 1932 E. V. Huntington presented a basis for Boolean algebra: commutativity, associativity and the Huntington equation.
  - Herbert Robbins conjectured it could be replaced by one simpler equation (the Robbins equation), leading (later) to Robbins algebras. Are all Robbins algebras Boolean algebras?
  - Despite work by Robbins and Huntington and Tarski and other, no solution was found.

In November 1997, a computer solves the Robbins conjecture.

- First “creative” proof by computer.
  - Qualitative difference from prior results based more heavily of exhaustive search such as the four-color theorem:
    - Any planar map can be colored in using 4 colors so that no two edge-adjacent regions have the same color
    - Proven in 1976 with a combination of human effort and “sophisticated computing” that enumerated many different special cases.

Learning

- Backgammon:
  TD-gammon [Tesauro]
  - Plays world-champion level backgammon.
  - Learns suitable strategies by playing games against itself.
  - Plays millions of games.
  - Based on a neural network trained using "backpropagation": incremental changes based on observed errors.
- Method as not generalized too well to other games
  - Eg. Neuro-chess loses to gnu-chess most of the time, and gnu-chess isn't tops.

Important challenges

- Domain specificity:
  - Successful systems are restricted to a narrow domains and specific tasks.
- Coping with noisy data
  - Most successes have been in domains where the objectives and the “rules” were closely specified and formalized.
- Incorporation of commonsense knowledge
  - Does every little thing have to be encoded or derived explicitly?
  - Compare deep-blue to human performance re. pruning.

Domain specificity

- Natural language systems work "well" only when the "domain of discourse" is restricted.
- If not, things get very hard very fast.
- Consider these alternative meanings of "give":
  - John have Pete a book. [tangible object delivered]
  - John gave Pete a hard time. [mode of behavior]
  - John gave Pete a black eye. [specific action]
  - John gave up.
  - John gave in.
  - John doesn't give a hoot about his courses.
  - I give him a week before he quits.
  - He'll quit next month, give or take a week.

Problem progression in Speech

- "Word spotting" is good today. Key is to ignore all of an utterance except keywords of interest.
- Speaker dependent continuous speech: quite good.
- Speaker *independent* continuous speech is getting good. Works well with a limited vocabulary.
- Speech *understanding* has proved to be very difficult, and is still not very good.
  - Especially is the speech is unrestricted,
- Apocyrphal example: "The spirit is willing but the flesh is weak" *translated into Russian as* "The vodka is good, but the meat is rotten."

Topic 2: Propositional logic

How to we **explicitly** represent our knowledge about the world?

References:

Dean, Allen, Aloimonos, Chapter 3
Russell and Norvig: Chapter 6

One of two or three logical languages we will consider.

Logical languages are analogous to programming languages: systems for describing knowledge that have a rigid syntax.

Logical languages (unlike programming languages) emphasize syntax. In principle, the semantics is irrelevant (in a narrow sense).

Knowledge Representation

- Most programs are a set of procedures that accomplish something using rules and “knowledge” embedded in the program itself.
- This is an example of

  implicitly encoded information

  – If you want to change the way Microsoft Word implements variables in macros, you have to hack the code.

  – When my tax program needs to be upgraded for a new tax rule, the code needs to be rewritten.

  – In contrast, when my accountant incorporates the same new rule, little of no brain surgery is required.

Explicit knowledge

When we encode rules in a separate rule book or

Knowledge Base (KB)

we have

explicitly encoded

(some of) the information of interest.

i.e. the rules are separate from the procedures for interpreting them.

- Explicit knowledge encoding, in general, makes it easier to update and manipulate (assuming the encoding is good).

Q. Is a "plug-in" implicit or explicit knowledge?

Knowledge and reasoning

Objective: to explicitly represent knowledge about the world.

- So that a computer can use it efficiently….
  - Simply to use the facts we have encoded
  - To make **inferences** about things it doesn't know yet
- So that we can easily enter facts and modify our knowledge base.

- The combination of a formal language and a reasoning mechanism is a **logic**.
- Each fact: encoded as a sentence.

Wff's

- In practice, with logical languages we combine symbols to express truths, or relationships, about the world.
- If we put the symbols together in a permitted way, we get a **well-formed formula** or **wff**
- A proposition is another term for an allowed formula.
- A **propositional variable** is a proposition that is atomic: that it, it cannot be subdivided into other (smaller) propositions.
- We can combine propositional variables into compound statements (wffs) using truth-functional connectives.

AND, OR, NOT, IMPLIES, EQUIVALENCE

- Formuli are made from propositional variables and the connectives.

Terminology

- A set of wffs connected by AND's is a **<u>conjunction</u>**.
- A set of wffs connected by OR's is a **<u>disjunction</u>**.
- **<u>Literals</u>** plain propositional variables, or their negations: P and ¬ P.

<u>Semantics</u>

- We attach meaning to wffs in 2 steps: 1. By assigning truth values to the propositional variables 2. By associating real-world concepts with symbols
- Step 1, assigning truth values, is called an **interpretation**.
- Step 2 is called **symbol grounding**, and is not related to the logical consistency or mathematical soundness of the logical system.

Discovering “new” truths

- Want to be able to generate new sentences that must be true, given the facts in the KB.
- Generation of new true sentences from the KB is called

entailment.

- We do this with an **inference procedure.**
- If the **inference procedure** works “right”: only get entailed sentences. Then the procedure is **sound** or **truth-preserving**.

  Q. Why would we ever consider any other kind of inference?

Knowing about knowing

- We would like to have knowledge both about the world, as well as the s*tate of our own knowledge* (i.e. **meta-knowledge**).

- **Ontological commitments** refer to the guarantees given by our logic and KB regarding the real world.
- **Epistemological commitments** relate to the states of knowledge, or kinds of knowledge, that a system can represent.

A particular set of truth assignments associated with propositional variables is a **model** IF THE ASSOCIATED FORMULA (or formuli) come out with the value true.

e.g. For the formula

```
(A and B) implies ( C and D)
```

the assignment

```
A=true B=true C=true D=true
```

is a model.

The assignment

```
A=false B=true C=true D=true
```

is another model, but the assignment

```
A=true B=true C=true D=false
```

is not a model.

Satisfiability

- If *no model is possible * for a formula, then the formula is NOT **SATISFIABLE**, otherwise it is satisfiable.
- A **<u>Theory</u>** is a set of formuli (in the context of propositional logic).
- If no model is possible for the negation of a formula, then we say the original formula is **<u>valid</u>** (also a formula is always true, it is a ***tautology)***.
- An axiom is a wff that states a priori information.
- Proper axioms state facts.