Intelligent Systems: Reasoning and Recognition

James L. Crowley

ENSIMAG 2 and MoSIG M1 — Winter Semester 2016

Lecture 1 — 3 February 2016

Intelligence: Reasoning, Understanding and Knowledge

Outline

Class notes on the web :

http://www-prima.inrialpes.fr/Prima/Homepages/jlc/Courses/2015/ENSI2.SIRR/ENSI2.SIRR.html

Intelligence, Knowledge and Reasoning

What do we mean by Intelligence?

INTELLIGENCE :

> (Petit Robert) "La faculté de connaître et comprendre,
> incluant la perception, l'apprentissage, l'intuition, le jugement et la conception."
>
> (Dictionnaire American Heritage) "The ability to know and to reason"

In this course we are concerned with technologies for Knowledge, Reason and Understanding.

The term "Artificial Intelligence" emerged from a pioneering workshop at Dartmouth University in 1956. Pioneers attending this workshop included Alan Newell, Herb Simon, John McCarthy, Marvin Minsky, Nils Nilsson, and Ed Feigenbaum.

Intelligence as the Ability to Solve Problems

A. Newell and H. Simon defined Intelligence as the Application of Knowledge to Problem Solving"

Newell, A.; Shaw, J.C.; Simon, H.A. (1959). Report on a general problem-solving program. *Proceedings of the International Conference on Information Processing*. pp. 256-264.

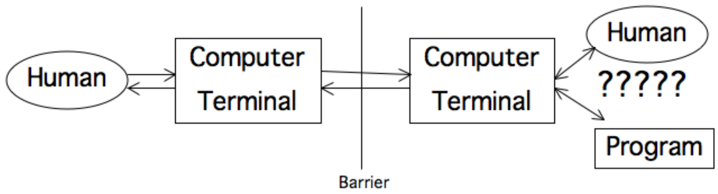
Nilsson: STRIPS, A* GraphSearch
R. Fikes and N. Nilsson (1971), "STRIPS: A new approach to the application of theorem proving to problem solving", Artificial Intelligence 2: 189–208

This view allows us to define knowledge in terms of the ability to solve problems, and reasoning as the ability to generate knowledge.

Intelligence as a Description of Behaviour

But what do we mean by Intelligence? Alain Turing asked this question in 1936

The Turing Test: an imitation game



In 1936, Alan Turing claimed that a machine would exhibit intelligence if it exhibited behaviour that could not be distinguished from a person.

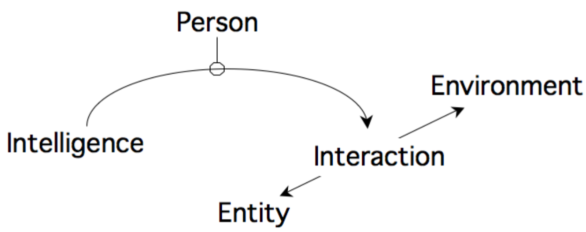
Limits: 1) Assumes that only humans are intelligent
2) Reduced intelligence to human linguistic interaction

Turing posed the problem in terms of linguistic and social interaction, ignoring many other forms of intelligence.

However, Turing gave an important insight: Intelligence is NOT an intrinsic property of an agent. Intelligence is a "DESCRIPTION".

A Modern View of Intelligence

In the 1990's, research in robotics and perception, combined with insights from Cognitive science to bring about a new view of intelligence as a description of interaction. The key idea is that Intelligence is a descriptive label not an intrinsic property.



Intelligence describes the interaction of an entity with its environment.*
Intelligence is a description (an ascribed property)
Intelligence describes an entity that interacts.

To be considered "intelligent", a system must be embodied, autonomous, and situated [Breazeal 02], [Brooks 94].

Embodied:	Possessing a body (sensory/motor components)
Autonomous:	Self-governing; Have independent existence
Situated:	Behaviour determined by the environment

[Breazeal 02] C. Breazeal, Designing Sociable Robots, MIT Press, 2002.
[Steels and Brooks 94] L. Steels, and R. Brooks, The artificial life route to artificial intelligence: Building Situated Embodied Agents. New Haven: Lawrence Erlbaum Ass., 1994.

Embodied: Incarnated. Possessing a body.

Body: A sensori-motor system for tightly coupled interaction with an environment.

Examples of Bodies:
  Natural: Human, mammal, insects, bacteria, plants,
  Artificial: Humanoid Robot, AIBO, mobile robots, roomba?

Environment: A system composed of multiple interacting entities.

Examples of Environments:
  Natural: Jungle, desert, sea floor….
  Artificial: Office, home, family, social network, computer games…
  Abstract: Chess, mathematics, any academic discipline…

What is Knowledge?

What is knowledge? - Competence
Whatever enables the solution of problems.

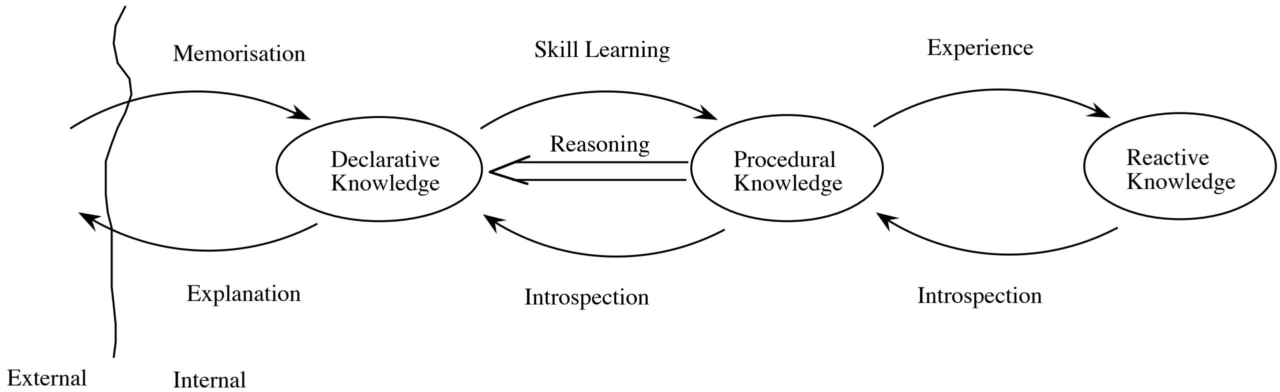
Knowledge is defined by function and not by representation.

Kinds of Knowledge
Cognitive Psychologists identify different categories of knowledge representation.

Declarative: A symbolic expression of competence.
Declarative knowledge is abstract
Declarative knowledge is used to communicate and to reason.
Declarative knowledge must be interpreted to be used.

Procedural: A series of steps to solve a problem.
A compiled expression of knowledge

Reactive: stimulus - response.



Newell proposes the distinction between "superficial" knowledge and "deep" knowledge.

Superficial knowledge provides reasoning without understanding. A common example of **superficial** reasoning is reasoning by symbol manipulation, without regard to the meaning of the symbols.

Deep knowledge requires the ability to predict and explain, and requires some form of model.

What is Reasoning?
Generation of new knowledge by inference.

Examples of types of inference using Symbolic Logic

Deduction : $(p \wedge (p \rightarrow q)) \Rightarrow (q)$
Abduction : $(q \wedge (p \rightarrow q)) \Rightarrow \text{Maybe}(p)$
Induction: $p(A) \rightarrow q, p(B) \rightarrow q, \ldots \Rightarrow \forall x\, (p(x) \rightarrow q)$

Recognition

Recognition (noun) The identification of something as having been previously perceived, experienced or known.

Perception: (noun) 1: the ability become aware of something through the senses.
2: The process of recognizing and interpreting sensory stimulus.

Perception is more than sensing. Perception requires (1) Attention, (2) Sensing, (3) Recognition and (4) Assimilation.

Phenomena: Any pattern or event that can be perceived.

Attention: The human senses can be modeled as a hierarchy of filters. Attention is the process of tuning the filters to allow perception.

Sensing: Transformation of an external stimulus into an internal sensation. In biological systems, sensory signals are expressed as activations of neurons.
In digital systems, sensory signals are expressed as symbols.

Recognition: Association of an external stimulus with a previously known or experienced phenomena.

Humans learn to recognize and reason through experience.

Understanding

Understanding can be described as the ability to predict and explain.
Understanding typically relies on some form of model that can be used to predict the outcome of a process or phenomena.
Decomposing the model into components and interactions between components provides a means to explain a process or phenomena.

The ability to predict phenomena is fundamental to survival.
The ability to explain phenomena is key to learning.

Our goal in this course is to learn the theories and models for building systems that can recognize and reason.

Symbolic Reasoning

In 1980, Alan Newell proposed that Intelligence REQUIRED the manipulation of symbols by a physical system. Newell's was based on a linguistic view of intelligence dating back to the 19th century.

What is a symbol?
A symbol is a 3rd order relation between
A sign
A thing
An interpreter

Power of Symbolic Reasoning: Generalisation!
Weakness: Closed system - very limited possibilities for learning.

Problems:
1) 1st order logic is restricted to closed domains. The set of symbols are predefined and not extensible. The real world is open

2) 1st order logic is a purely syntactic (shallow) reasoning system.

3) Knowledge acquisition (learning) is very expensive and difficult.

Limitations of Physical Symbol Systems:

There are several problems with Newell's hypothesis
1) It restricts intelligence to symbol manipulation. Intelligence is more general.

Newell claimed that symbol manipulation was necessary and sufficient for intelligence. We now know that this is not show.

We can show limited forms of intelligence without symbols (or generalisation).
Thus symbol systems are not necessary.

2) Intelligence requires an open universe. Current techniques in symbolic reasoning require a pre-defined (closed) system of symbols.

More broadly, I would argue that Newel confused "What intelligence is" with "How intelligence is achieved".

Machine Learning

Machine Learning: the construction of algorithms that can make predictions from data.

Machine learning is a domain of Informatics concerned with pattern recognition, computational learning and artificial intelligence.

Definition proposed by Tom Mitchell (CMU): A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.

In short: Learn(E, T, P) : P(T) improves with E.

Most popular machine learning technique concerns the task of Recognition (also called Classification).

Recognition is a process that labels a sensory signal X as an example of a phenomena of class C.

However, Machine Learning concerns other tasks such as Function Learning, Skill Acquisition, Clustering, Data mining, Concept Formation and Reasoning.

Categories of Machine Learning algorithms:

Supervised Learning: Learning a function from a labeled data set
Unsupervised Learning: Learning a function from unlabeled data.
Clustering: Discovering groups (categories) of data based on similarity of some property or measure.
Reinforcement learning: Learning to perform actions from experience

In the early 1960's, Frank Rosenblatt demonstrated a numerical learning algorithm named the "perceptron", [Rosenblat 62], using an analog computer.

The dominant paradigm was Probabilistic Pattern Recognition, summarized in textbooks by [Nilsson 65] and [Duda-Hart 73], using methods based on Probability, Statistics and Bayes rule.

In the 1970s and 1980s AI researchers rejected Probabilistic Pattern Recognition in favor of hand coded symbolic reasoning systems. The dominant paradigm was "Expert systems", by programming problem solving techniques used by domain experts.

In the 1990's when it was shown that many intractable AI problems could be easily solved with a two-layer perceptron, also called an Artificial Neural Networks (Hinton 86).

Within a few years, Bayesian machine learning techniques were shown to outperform Artificial Neural Networks for learning functions to detect and recognize patterns.

In the early 2000's discriminative techniques such as boosted learning, support Vector Machines and Kernel Methods provided improved performance for recognition.

Since 2010 a new revolution in recognition has occurred as massive amounts of data and computing power, made possible by the Internet and grid computing, have enabled techniques for learning deep Multi-layer Neural Networks with up to 20 or so layers (Deep Learning). Performance gains of 20% to 30% on classic problems in Computer Vision and Speech recognition were observed.

It is hard to say where the next revolution will come from.

A current challenge is to combine the learning power and open universe of Machine Learning with the expressive power and generalization of symbolic logic.

Rather than learning this year's hot technique, in this class we will concentrate on the fundamentals of intelligent systems.

Course Overview

Part 1 – Recognition

1) Regression Analysis
2) Classification
3) Clustering
4) Artificial Neural Networks
5) Graphical Models

Part 2 – Reasoning

1) Rule Based Systems
2) Temporal Reasoning
3) Structured Knowledge Representations
4) Situation Models
5) Planning

Exercises will NOT be graded.
Feedback and corrections will be provided for COMPLETED exercises.
Completed exercises should be COPIED INTO AN EMAIL including the names of all persons who contributed to the solution. Feedback will be returned by email. Please allow at least 2 weeks for feedback.

Programming exercise in part 2 will use CLIPS. – C Language Integrated Production Systems.
Note: DO NOT Send a file named file.clips or Exercise.clips, etc.