ARTIFICIAL INTELLIGENCE

Pushpak Bhattacharyya
Department of Computer Science and Engineering, Indian Institute of Technology Bombay, India.

Keywords: strong and weak AI, concentric circles of AI areas, search, A* algorithm, categories of reasoning, propositional calculus, predicate calculus, resolution, knowledge representation, description logic

Contents

1. Introduction
1.1. Definition of AI
1.2. Subfields of AI
1.3. Theoretical Underpinnings
2. Search
2.1. Algorithmics of Search
2.1.1 The 8-puzzle problem
2.1.2. Missionaries and the Cannibals Problem [Rich and Knight 1991]
2.1.3. The General Graph Search Algorithm [Nilsson 1986]
2.1.4. A* algorithm
2.1.5. Properties of *A**
3. Logic and inferencing
3.1. Deduction, Induction and Abduction
3.2. Propositional Calculus
3.2.1. Truth Table
3.2.2. Functional completeness of $\Rightarrow$ and $\neg$
3.2.3. Laws of Propositional Calculus
3.2.4. Tautologies
3.2.5. Semantic Tableau Method of Proving a *Tautology*
3.3. Formal System
3.3.1. Formalization of Propositional Calculus
3.3.2. Proofs and Theorems in a Formal System
3.3.3. The Deduction Theorem
3.4. Theorems, Meta-Theorems And Self Reference
3.6. Tautologies and Theorems
3.6. Soundness, completeness and consistency
3.7. An example to illustrate the application of propositional calculus
3.8. Predicate calculus
3.8.1. Limitations of propositional calculus
3.8.2. The powerful constructs of predicate calculus: *predicates, variables* and *quantifiers*
3.9. Inferencing in Predicate Calculus: Resolution and Unification
3.10. An example to illustrate the application of predicate calculus
4. Knowledge Representation
5. Description Logic and Semantic Web

5.1. Description Logic and Predicate Calculus
5.2. Constituents of DL
5.3. Inheritance and Subsumption
6. Conclusion
6.1. The Role of Human Intelligence in the Formal System
6.2. The Role of Human Intelligence in Puzzle Solving by Propositional Calculus
Acknowledgements
Glossary
Bibliography
Biographical Sketch

Summary

A deep question challenging humanity has always been, "*what is intelligence and what are the structures and processes underlying it*"? Simultaneously, there has been the quest for determining *how much of intelligence can be mechanized*. These efforts have given rise to the field of Artificial Intelligence (AI). We describe here this exciting field, introducing its sub areas, with detailed discussion on two core topics- *search* and *logic*. Fundamental concepts are presented with numerous illustrations. A few insightful exercises assist further grasp of concepts.

1. Introduction

Very few intellectual pursuits have captured so much of imagination, generated so much excitement and undergone so many upheavals as the field of *Artificial Intelligence* (*AI*)- also called *Machine Intelligence* (*MI*). AI has been the theme of science fictions, futuristic movies and extensive speculations. In this matter AI can be likened to alchemy which pursued the goal of turning ordinary metal to gold. Though nobody believes today in alchemy, the pursuit of the impossible led to numerous insights into chemistry and material science. Likewise, AI has motivated many branches of Computer Science (CS), so much so that it has been called the *forcing function* for Computer Science. Some examples of this fact are:

(i) *Compilers* which translate descriptions of a task in a high level programming language to computer executable instructions, had their genesis in *automatic programming*- an area aspiring to automatically create computer programs from commands in natural languages.
(ii) *Time sharing systems* originated in the effort to create an illusion of the computer talking to and serving a number of people at the same time.

Unlike alchemy, however, AI is not an abandoned field. Current times are witnessing a resurgence of interest in AI, caused by advancements in (a) hardware, (b) systems programming, (c) programming languages, (d) design and analysis of algorithms, and last but not the least (e) the internet. Simultaneously, progress in neurosciences, quantum mechanics, linguistics, cognitive science and consciousness studies has impacted and enriched AI, leading the field into hitherto uncharted territories.

Intelligence on silicon is, thus, a live question- often termed as the last frontier for the human mind to conquer.

1.1. Definition of AI

AI is the study of creating computational systems for solving problems which currently *human beings are proficient at*. It is observed that there is a dichotomy in the set of problems that can be solved by computers and that by humans. Humans' set of problems involves pattern recognition, *e.g.*, face recognition, sentence comprehension, reasoning *etc*. Computers' set of problems, on the other hand, involves *number crunching*, *e.g.*, scientific computation, very fast look up of data items, efficiently arranging a set of items in some order and so on.

To be considered successful, therefore, AI must solve problems from human task domain. AI systems that have only functional similarity with humans form *Weak AI*, while AI systems that have similarity with humans at every level belong to *strong AI*. The latter propose to mimic humans in every small detail including the architecture. Skeptics of AI maintain that strong AI is impossible. *Intelligence can emerge and reside only on carbon.*

The goals of AI are three fold: The *science goal* enquires into the nature of intelligence, its conditions and determinants and the possibility of replicating it on computers modeled as *Turing Machines*. The engineering goal aims to create *computer programs* that display close to human level performance in specific tasks, *e.g.*, robot path planning. The goal involving *both science and engineering* emphasizes in-depth investigation into specific cognitive phenomena, building models of intelligence and the use of these models to create computer programs.

1.2. Subfields of AI

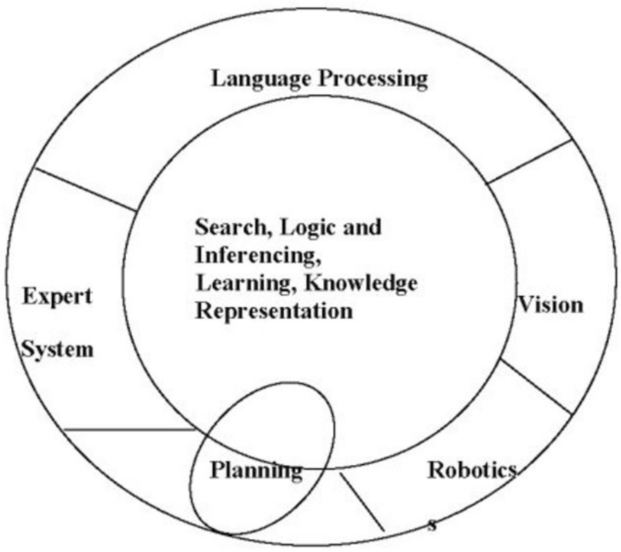

Figure 1. Different subfields of AI

AI has evolved over time into many subfields, driven by intellectual and utilitarian goals. Figure 1 shows two concentric circles, depicting the core areas and the application areas. *Search, Logic and Inferencing, Knowledge Representation* and *Learning* form the core of AI. *Computer Vision, Natural Language Processing, Robotics* and *Expert Systems* are application areas. *Planning* is a subfield which is both core and application, because it uses the core topics and is in turn core for other areas. The subfields are briefly described (Figure 1):

1. The desire for human like information processing has motivated investigation into automation of linguistic and visual information processing, giving rise to the fields of Natural Language Processing (NLP) and Computer Vision (CV). NLP and CV are shown in the outer circle.
2. Humans have aspired to construct their mechanical replica in *intelligent robots* that could plan and perform motor actions. Robotics too is placed in the outer circle.
3. *Experts* in any field- besides being highly skillful at their tasks- employ a large body of specialized knowledge which is stored at many levels in their minds and is brought to bear upon specific problem situations. However, expert tasks often involve large components of routine activity which can be *mechanized*. For example, in medical diagnosis preliminary routine investigations can be performed by a machine (MYCIN is a famous example of this.). This has given rise to the field of Expert Systems (ES), also known as Knowledge Based Systems (KBS). Most commercial successes of AI have come from expert systems. ES is shown in the outer circle.
4. While NLP, CV and ES are areas of visible impact, there are core AI areas which form the foundation for these areas. For example, in ES, inferencing is of crucial importance. NLP requires inferencing at various stages. Logic and inferencing has come to be regarded as core AI and is placed in the inner circle.
5. Human problem solving is characterized by searching amongst and choosing from a large number of options. Section 2 gives many examples of this. The area of search forms core AI and is shown in the inner circle.
6. AI systems are knowledge intensive. In computer vision, the physics of the scenario needs to be captured efficiently. In natural language processing, the knowledge of the language as well as the knowledge of the world has to be represented in the computer. Knowledge Representation (KR) too is core AI and appears in the inner circle.
7. No system can be called intelligent if it cannot learn from experience. Presented with examples, a machine should be able to generalize and deal with previously unseen situations. Machine Learning, therefore, is another core activity for AI and is given a place in the inner circle.
8. Given a goal, deciding on a set of actions to achieve it and sequencing the actions is fundamental to intelligence. This is called *planning* and is again core AI.

To summarize, the areas in the inner circle of Figure 1 are core ones which feed into the areas in the outer circle, which are closer to the user. The area of planning is unique in the sense that it is both a core and an application area, since logic and search feed it, and it in turn it feeds into robotics and language generation.

1.3. Theoretical Underpinnings

AI has deep connection to *computability*. Figure 2 shows a Turing Machine which is the celebrated model of computation proposed by Alan Turing in 1936. The machine has a *head* (which serves as the *Central Processing Unit*) and an *infinite tape* (which serves as the *memory*). The head does only 4 operations: *reading* the symbol on the cell, *writing* on to the cell and *moving right* or *moving left*. Depending on the symbol read, the *finite state control* changes state. This simple machine is said to capture the whole of *computation* as stated by the *Church Turing Hypothesis*:

Any function that is computable is computable by a Turing Machine; conversely, functions computable by Turing Machines are the only computable functions.

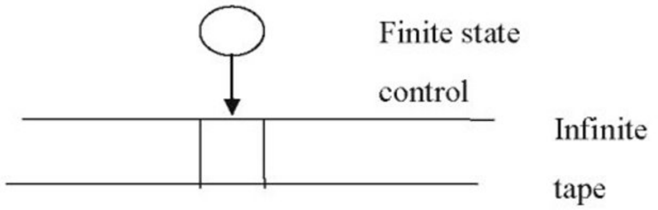

Figure 2. Turing Machine

The next hypothesis is called the *Physical Symbol System Hypothesis* and is due to Newel and Simon. It states

Intelligence emerges from manipulating symbols and nothing else is needed

Opposed to this is the *Animacy Hypothesis* which takes the position that

Intelligence is producible only on carbon base

Another hypothesis comes from *Goedel's theorem* which essentially says that

In any formal, rule governed system of sufficient representational or expressive power there will always be statements which are neither provable nor disprovable by the formal machinery of the system.

Many of the ultimate limits to the mechanization of intelligence can be traced to *self reference* which is further explicated in Section 3.

2. Search

We begin discussing two core areas of AI. The first topic is *search* which is ubiquitous in AI. We take a few examples.

1. Planning: planning is concerned with the selection and sequencing of a set of actions to achieve a particular goal. Refer to Figure 3. A robot has to arrange the blocks *A*, *B* and *C* in a vertical tower in the order *C*, *B*, *A* from top to bottom.

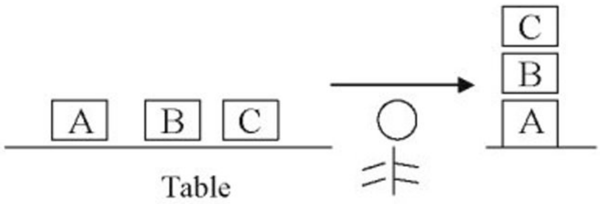

Figure 3. Search in Robot action planning

At every step of the operation many possibilities exist: (a) which block to *pick*, (b) which to *stack*, (c) which to *unstack*, (d) whether to *stack* a block or (e) whether to *unstack* an already stacked block. These options have to be searched in order to arrive at the right sequence of actions.

Vision: interpretation and understanding of real scenes is the concern of this area. Refer to Figure 4. Two robotic eyes *R* and *L* receive the image of the outside world. It is required to solve what is known as the *correspondence problem* in stereo vision. A search needs to be carried out to find which point in the image of *L* corresponds to which point in *R*. Naively carried out, this can become an $O(n^2)$ process where n is the number of points in the retinal images.

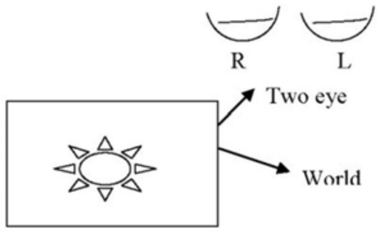

Figure 4. Search in scene deciphering

Robot path planning: navigation of a robot on a terrain by avoiding obstacles is the concern here. Refer to Figure 5.

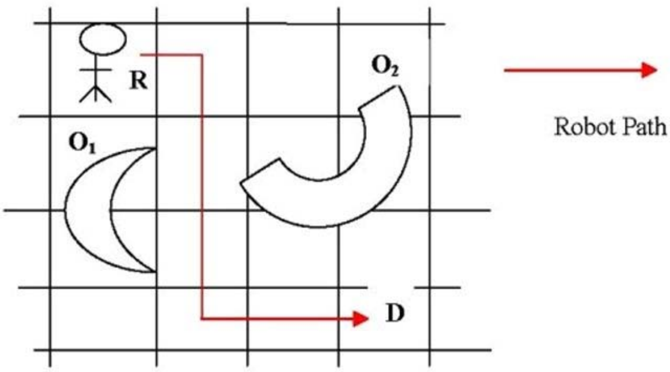

Figure 5. Robot Path Planning

The robot **R** has to reach the destination **D** by avoiding obstacles $\mathbf{O_1}$ and $\mathbf{O_2}$. This needs searching amongst the options of moving ***L***eft, ***R***ight, ***U***p or ***D***own. Additionally, each movement has an associated cost representing the relative difficulty of each movement. The search then will have to find the *optimal*, *i.e.*, the *least cost* path.

Language processing: natural language processing is concerned with the analysis and generation of human languages. Consider the sentence shown in Figure 6. The parts of speech of ambiguous words are listed.

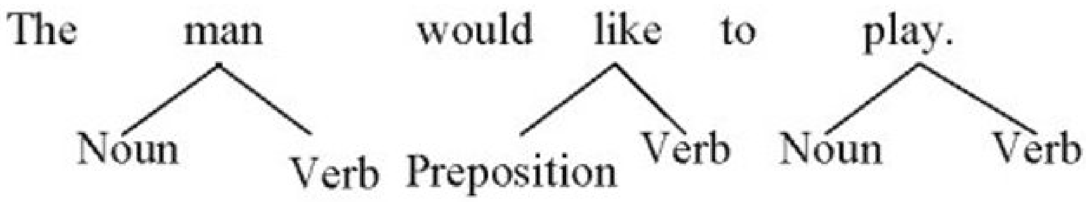

Figure 6. Sentence with multiple possibilities on parts of speech tag on words

A language understanding system has to search among many combinations of parts of speech on the way to deciphering the meaning. This applies to every level of processing- *syntax*, *semantics*, *pragmatics* and *discourse*.

Expert Systems: In expert systems, a program has to make use of large amounts of *domain specific knowledge* to produce *inferences*. Consider the problem of inferring diseases based on symptoms from a rule base of symptom-to-disease mappings. A typical rule is:

$$
\begin{array}{l}
IF \\
\left\{ \begin{array}{l} the\ infection\ is\ primary\text{-}bacteremia \\ AND\ the\ site\ of\ the\ culture\ is\ one\ of\ the\ sterile\ sites \\ AND\ the\ suspected\ portal\ of\ entry\ is\ the\ gastrointestinal\ tract \end{array} \right\} \text{If conditions} \\
THEN \\
there\ is\ suggestive\ evidence\ (0.7)\ that\ infection\ is\ bacteroid.
\end{array}
$$

This rule is from the famous expert system MYCIN which diagnoses bacterial infections with an accuracy of about 65%- a figure comparable to a doctor's performance. The set of *if-conditions* are conjunctions of clauses each of which can be conditions in other rules. Diagnosis of diseases then is a staged search through the set of rules, prompted by more and more information obtained from new symptoms.

2.1. Algorithmics of Search

There are four basic elements in search:

1. *The State Space*: *States* represent the constituents and constraints of a problem. After

a thorough understanding and careful analysis of the problem at hand, *states* are defined and their dependencies specified. This gives rise to the *state space graph*, navigation in which produces the solution to the problem.
2. *Start and Goal states*: These are special states in the state space graph. The start state, denoted *S*, specifies the state from which the search process begins. The goal *G* is the state in which the search process terminates. It is possible to have multiple goal states, but the start state is unique.
3. *Operators*: They operate on states to produce new states. Each operator tells the search algorithm what states to visit/produce from a state. Each operator has a *cost* associated with it.
4. *Optimal path*: The search process finds a path from *S* to *G*. This path should be optimal, *i.e.*, least cost. The cost of the path is the sum of costs of all the arcs constituting the path.

These concepts are illustrated through two well known examples.

-
-
-

Bibliography

D. Nardi, R. J. Brachman, (2002). An Introduction to Description Logics. In the *Description Logic Handbook*, (ed.) F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi, P.F. Patel-Schneider, Cambridge University Press, pages 5-44.[One of the most authentic treatises on DL]

F. Donini. Complexity of Reasoning. In the *Description Logic Handbook*, edited by F. Baader, D. Calvanese, D.L McGuinness, D. Nardi, P.F. Patel-Schneider, (2002). Cambridge University Press, pages 101-141.[Discusses why it is nontrivial to automate human level reasoning]

Manna, Z. (1974), *Mathematical Theory of Computation*, Computer Science Series, McGraw-Hill Inc [A classic on theoretical computer science].

Mitchell, T., (1997). "Machine Learning", McGraw-Hill . [Standard text book on machine learning].

Nilsson N.J. (1986). *Principles of Artificial Intelligence*, Morgan Kaufmann Publishers. [A classic introduction to Artificial Intelligence; especially covers *search*, *predicate calculus* and *planning* really well].

Ramsay Alan, (1988), *Formal Methods in Artificial Intelligence*, Cambridge University Press [Discusses well the theoretical foundations of AI].

Rich E, and Knight K., (1991) *Artificial Intelligence*, McGraw-Hill College [A very lucid introduction to AI with broad coverage]

Russel S. and Norvig P., (1998) *Artificial Intelligence: a Modern Approach*, Prentice Hall [A modern, comprehensive coverage of AI fundamentals].

Papers: selected topic based papers from the

(i) Journals: Artificial Intelligence, Artificial Intelligence Programming, Machine Learning, IEEE Expert,

Data and

Knowledge Engineering, Pattern Recognition *etc.*

(ii) Conferences: American Association for AI, International Joint Conference on AI, Uncertainty in AI, International

Conference of Machine Learning, Association of Computational Linguistics, SIGAI, SIGIR *etc.*

Course and tutorials (*description logic part*)*:*

Enrico Franconi, Course on Description Logic *http://www.inf.unibz.it/~franconi/dl/course/*

Ian Horrocks and Ulrike Sattler Tutorial titled *"Description Logics- Basics, Applications and more"* in *ECAI-2002, Lyon, France http://www.cs.man.ac.uk/~horrocks/Slides/index.html*

Biographical Sketch

Dr. Pushpak Bhattacharyya is a Professor of Computer Science and Engineering, at IIT Bombay. He did his Bachelor of Technology in Electrical Engineering from Indian Institute of Technology, Kharagpur, India in 1984, Master of Technology in Computer Science and Engineering from Indian Institute of Technology, Kanpur, India in 1986 and Ph.D. in Computer Science and Engineering from Indian Institute of Technology, Bombay, India in 1989-1994. He was a Visiting Research Fellow at Massachusetts Institute of Technology, USA in 1990.

Dr. Pushpak Bhattacharyya was a Professeur Invite' at Groupe D'Etude en Traduction Automatique-Communication Langagiere et. Interation Personne-Systeme (GETA-CLIPS) and Universite Joseph-Fourier, Grenoble, France in 2005. He was also a Visiting Professor at Stanford University in 2004. In 1986-87 he was a Research and Development Engineer at Wipro Information Technology Ltd a top IT industry of India.

Dr. Bhattacharyya's research areas include Natural Language Understanding, Machine Translation, Information Extraction, Machine Learning and Neural Nets. He has published close to 100 research papers in these area in leading journals and conferences (*Journal of Machine Translation, Pattern Recognition Journal, ACL, Machine Translation Summit, IJCAI, WWW etc.*)

Dr. Bhattacharyya has served as program committee member of leading conferences (*COLING, ICON, MT Summit etc.*) *and* as referee for leading journals (*IEEE and ACM journals related to AI*) and conferences (*ACL, WWW, COLING, ICON, KBCS etc.*). He is a founding member of the Natural Language Processing Association of India and the South Association for Language Processing.

Dr. Bhattacharyya has been leading large international and national projects on natural language processing and machine translation sponsored by United Nations, Ministry of Information Technology of India and prominent IT industries (*Tata Consultancy Services*). The research team led by includes close to 30 research staff and students.