INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

Comparative study of Symbolic Reasoning and Statistical Reasoning

Shweta Agarwal

Assistant Professor
MVN University
Palwal, 121105
gud22aga@gmail.com

***Abstract**: This paper examines reasoning under certainty by symbolic and statistical methods and the comparison between these two methods. Basically reasoning is a process of logically arguing and drawing inference. For reasoning, the system must find out what it needs to know from what it already knows. In this paper we review symbolic reasoning and different methods of statistical reasoning like probability, Bayes theorem and Bayesian network.*

***Keywords:** Statistical Reasoning, Symbolic Reasoning, Bayes Network, Bayesian network, Truth Maintenance System.*

1. INTRODUCTION

Artificial intelligence (AI) is the branch of computer science and it is the intelligence of machines. In text books it is defined as "the study and design of intelligent agents" where an intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success. AI is defined as the science and engineering of making intelligent machines.

1.1 Reasoning

The reasoning is a mental ability found in humans. Reasoning means the ability to generate conclusions from assumptions and premises.

2. REASONING

Reasoning is the act of deriving a conclusion from certain premises using methodology. Reasoning is process of thinking. It is a logical arguing and drawing inference. Any knowledge system must reason, if is required to do something which has not been told explicitly. For reasoning the system must find out what it needs to know from what it already knows. The information about anything may be incomplete or volatile or may be some important facts and details about the problem could be missing or the facts available may be fuzzy. Human being can deal with uncertainties on the daily basis and found the some reasonable solutions. Any AI system that reasoning in such a world must be deal with the:
Incompleteness: Compensate for lack of knowledge.
Inconsistencies: Resolve ambiguities and contradictions.
Change: It must be able to update its world knowledge base over time.
There are three basic methods to deal with uncertainty are:

- Symbolic methods
- Statistical methods
- Fuzzy logic methods

2.1 How reasoning can be done

Human reasoning capabilities are categorized into three:
Mathematical reasoning: theorems, proofs.
Logical reasoning: deductive, inductive.
Non logical reasoning: linguistic, language.

These all are reside in every human being but the ability level depends on the education and environment.
The IQ is the summation of mathematical reasoning skills and the logical reasoning.
The EQ is dependent on non- logical reasoning.

2.2 Various methods for reason

1) Logical Reasoning: Logical reasoning is the process of drawing conclusions from premises using rules of inference. The logic may be formal or informal.
Formal logic is the study of symbolic abstractions that captures the formal features. Logic also needs semantics which defines how to assign meaning to expressions.
Informal logic is the study of natural language arguments. The focus lies in distinguishing good arguments from bad arguments.
2) Procedural Reasoning: Procedural reasoning uses procedures that define how to solve a problem.
This includes reasoning by analogy, generalization and abstraction, Meta level reasoning.
3) Monotonic Reasoning: In monotonic reasoning if we enlarge set of axioms we can not retract any existing assertions or axioms. Axioms are sentences that are true with the system. Logic is monotonic if the truth of a proposition does not change when new information are added to it.
4) Non Monotonic Reasoning: The non monotonic reasoning is caused by the fact that our knowledge about the world is always incomplete. In this new facts which are contradicting and invalidating the old knowledge came into picture.
A non monotonic logic is a formal logic whose consequence relation is not monotonic. A logic is non monotonic if the truth of a proposition may change when new information are added. Human reasoning is nonmonotonic.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

3. SYMBOLIC REASONING

Symbolic reasoning is the science of reasoning by the aid of representing symbols. The intelligent mathematical software is the integration of the power of symbolic mathematical tools with the suitable proof technology.
Mathematical reasoning has the property called monotonicity that says that if a conclusion follows from given premises then it is also follows from larger set of premises.

3.1 Nonmonotonic Reasoning

It is another form of reasoning. As new facts are added old knowledge came into picture are invalidating and contradicting. By adding new facts old knowledge in knowledge base became invalid thereby requiring further retractions. This retractions lead to nonmonotonic growth in knowledge at times. If we are go on adding new axioms in the knowledge base, it only increase amount of knowledge base rather than decreasing.
The nonmonotonic reasoning is of many types:

- Default reasoning
- Circumscription
- Truth Maintenance System.

3.1.1 Default Reasoning: This type of reasoning is very simple. By this the decisions are made on what is most likely to be true. Nonmonotonic logic and default logic are two approaches for made decisions. Both approaches are logic type.
Nonmonotonic logic: A nonmonotonic logic is a formal logic whose consequence relation is not monotonic. A logic nonmonotonic logic is if the truth of a proposition may change when new information are added. This allows a statement to be retracted and used to formalize plausible reasoning.
Default logic: This logic initiates new inference rule A:B/C
Where, A=prerequisite, B=justification and C= consequent
The inference rule will be read as:
"if A, and if it is consistent with the rest of assume that it is B, then conclude C".
The idea of nonmonotonic reasoning is to reason with first order logic, and if an inference cannot be obtained then use the set of default rules available within the first order formulation.

While applying to default rule it is necessary to check their justification for consistency, not only with initial data, but also consequents of any other default rules that may be applied. The application one rule may block the other application. To solve this problem default theory is used.
4. Circumscription: Circumscription is a nonmonotonic logic to formalize the common sense assumption. It is a formalized rule of conjecture that can be used along with the rules of inference of first order logic. This involves formulating rules of thumb with abnormality predicates and then restricting the extensions of these predicates, circumscribing them, so that they apply to only those things to which they are currently known.
5. Truth Maintenance System: The purpose of truth maintenance system is to assure that inferences made by the RS are valid. The RS provides the Reasoning maintenance system with information about each inference it performs, and in return the RMS provides the RS with the information about the whole set of inference.
The Truth Maintenance System (TMS) is a knowledge representation method for representing beliefs and their dependencies. It has the ability to restore consistency.

The TMS maintains the consistency of a knowledge base as soon as new knowledge is added. It considers only one state at a time so it is difficult to manipulate environment.
The functions of TMS are:
Provide justification for conclusions: when a problem solving system gives an answer to a user's query an explanation is required to that answer. An explanation can be constructed by an inference engine.
Recognize inconsistencies: The inference engine may tell the TMS about the contradictory sentences. Then TMS finds all those that are believed to be true and reports to the inference engine to eliminate inconsistency.
Support default reasoning: In the absence of firmer knowledge we want to reason from default assumptions.
Support dependency driven backtracking: The justification of a sentence can be maintained by TMS.

4. STATISTICAL REASONING

In logic based approaches everything is either believed true or false. It is often useful to represent that something is probably true.
This is useful for dealing with problems where there is randomness and unpredictability. To do all this in a particular manner requires techniques for probabilistic reasoning.

1) Basic Statistical Methods Probability: Probability is a way of expressing knowledge or belief that an event will occur or has occurred.
Probabilities are real numbers in the range from 0 to 1.
A probability of $P(A)=0$ indicates total uncertainty in A and $P(A)=1$ indicates total certainty and values in between shows some degree of uncertainty.

2) Classical Probability: This is also known as priori theory of probability. The probability of an event A = Number of outcomes favorable to the occurrence of an event (f) divided by the total number of possible outcomes (n).

$$P(A) = f/n$$

All possible outcomes are equally likely.

Empirical Probability: This is also known as experimental probability. It determined analytically the knowledge about the nature of experiment rather than actual experiment.
Conditional Probability: The probability of some event A given the probability of some other event B. It is written as $P(A/B)$ and read as the probability of A given B.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

Joint Probability: It is the probability of two events together. The joint probability of two events written as P(A∩B) or P(A,B).

Marginal Probability: The probability of one event regardless of other event.
Probabilistic reasoning came from probability concept. It is used when outcomes are unpredictable.
A prior probability is an initial probability value originally obtained before any additional information is obtained.
A posterior probability is a probability value that has been revised by using additional information that is later obtained.

5. BAYES THEOREM:

This theorem relates the conditional and marginal probabilities of two random events. This theorem is given by Thomas Bayes' and called Bayes' rules. It is also called Bayesian reasoning. Bayesian view of probability is related to degree of belief. It is a measure of plausibility of an event. The fundamental notion of Bayesian statistics is that of conditional probability. Bayes theorem allows new information to be used to update the conditional probability of the event.
This theorem addressed both discrete probability distributions of data and more complicated case of continuous probability distributions. In discrete case, it relates conditional and marginal probabilities of events A and B, provided that probability of b does not equal to zero.

$$P(A/B)=P(B/A)P(A)/P(B)$$

Each probability has a conventional name:
is the prior probability of A. It is prior as it does not take into account the information about B.
is the conditional probability of A given by B. It is also called posterior probability because it is derived from or depends upon the value of B.
is the conditional probability of B given by A. It is also called likelihood.
is the marginal or prior probability.
Bayes theorem is used to calculate conditional probabilities.
Another formula:

$$P(B/B\cap C) = \frac{p(A)p(B/A)\,p(C/A\cap B)}{p(B)p(C/B)}$$

beyes theorem in terms of Odds (O) and Likelihood (Λ) ratio:

$$O(A/B)=O(A)\wedge(A/B)$$

WhereO(A/B) are the posterior odds of A given B,
$O(A/B)=P(A/B)/P(\tilde{A}/B)$

$$O(A)=\frac{P(A)}{P(\tilde{A})}$$

And $\wedge(A/B)= \frac{p(B/A)}{p(B/Ac)}$ is the likelihood ratio.

Extensions:
Theorem analogous to Bayes' theorem covers more than two events.

$$P(B/B\cap C) = \frac{p(A)p(B/A)\,p(C/A\cap B)}{p(B)p(C/B)}$$

Example: John's cricket match is tomorrow.
In recent years, each year it has rained only 5 days.
The weatherman has predicted rain for tomorrow.
When it actually rains, the weatherman correctly forecasts rain 90% of the time.
When it doesn't rains, the weatherman incorrectly forecasts rain 10% of the time.

Question is what is the probability that it will rain on the day of John's cricket match.

Solution:
There are two mutually exclusive events "it rains" or "it doesn't rain" and the third event occurs when the weatherman predicts the rain.
EventA1: Rain on the day of John's cricket match
EventA2: Rain doesn't on John's cricket match.
Event B: Weatherman predicts rain.

P(A1)=5/365=0.0136985(rains 5 days in year)

P(A2)=360/365=0.9863014(doesn't rain 360 days in a year)

P(B/A1) 0.9(when it rains, weatherman predicts 90%)of time

P(B/A2)0.1(when it doesn't rains, weatherman predicts 10%)of time

We want to know P(A1/B), the probability that it will rain on the day of John's cricket match given a forecast of rain given by weatherman.
The answer can be determined by Bayes' theorem:

$$P(A1/B) = \frac{P(A1)P(B/A1)}{P(A1)P(B/A1)+P(A2)P(B/A2)}$$

= (0.014) (0.9)/[(0.014)(0.9)+(0.986)(0.1)
= 0.111

So despite the weatherman's prediction, there is a good chance that John's will not get rain on at his cricket match.
Thus Bayes' theorem is used to calculate conditional probability.

5.1 Bayesian Network

A Bayesian network is a belief network or directed acyclic graphical model is a classical graphical model that represents a set of variables and their probabilistic independencies. As example Bayesian network represents the probabilistic relationship between diseases and symptoms.

Bayesian network consist a set of nodes and a set of directed edges between them. The effects are not completely deterministic. The strength of an effect is modeled as a probability.
Bayesian network consist a set of nodes and a set of directed edges between them. The effects are not completely deterministic. The strength of an effect is modeled as a probability.

5.2 Certainty and Factors

This approach is given by Shortliffe and Buchanan and used in their famous medical diagnosis MYCIN system. MYCIN represents knowledge as a set of rules and a certainty factor is associated with each rule. A certainty factor is based on measures of belief B and disbelief D of a hypothesis Hi, evidence E is given as:

B(Hi/E)=1 if P(Hi)=1
= max [p(Hi/E),P(Hi)]-P(Hi)/((1-P(Hi))P(Hi/E)

Otherwise
D(Hi/E)=0 if P(Hi)=0

$$= \frac{P(Hi)-\min[P(Hi/E),P(Hi)]}{P(Hi)P(Hi/E)}$$

Otherwise
The certainty factor C of some hypothesis Hi given evidence E is given by:
C(Hi/E)=B(Hi/E)-D(Hi/E)

This overcome the Bayes' rule

5.3 Dempsgtor- Shafter Model

This is more general approach to representing uncertainty than Bayesian. The basic idea in representing uncertainty in this model is: Set up a confidence interval : an interval of probabilities within which the true probability lies with a certain confidence based on the Belief B and plausibility PL provided by some evidence E for a proposition P. The belief brings together all the evidence that would lead us to believe in P with some certainty. The plausibility brings together the evidence that is compatible with P and is not inconsistent with it. This method allows for further additions to the set of knowledge and does not assume disjoint outcomes within which the true probability lies with a certain confidence based on the Belief B and plausibility PL provided by some evidence E for a proposition P.

6. CONCLUSION

This paper examined reasoning under uncertainty. Symbolic reasoning basically represents uncertainty as true, false, neither true nor false while statistical reasoning provide a method for representing belief which are uncertain and they may have some contradictory evidence. Statistical reasoning represents belief as true or false but not both. Symbolic methods is generally deal the problems which have randomness and unpredictability. Symbolic methods had problem with incomplete knowledge and contradictions in the knowledge. In symbolic methods if the number of exception is large then these systems are break down while statistical techniques can solve large exceptions.

REFERENCES

[1] Anderson, J.R., Acquisition of proof skills in geometry, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell (Eds.), Machine Learning: An Artificial Intelligence Approach (Tioga, Palo Alto, CA, 1983) 191-220.
[2] Anderson, J.R., The Architecture of Cognition (Harvard University Press, Cambridge, MA, 1983).
[3] Anderson, J.R., Farrell, R. and Sauers, R., Learning to program in LISP, Cognitive Sci. 8 (1984) 87-129.
[4] Anderson, J.R., Boyle, C.F. and Reiser, B.J., Intelligent tutoring systems, Science 228 (1985) 456-462.
[5] Anderson, J.R., Knowledge composition: The general learning mechanism, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell (Eds.), Machine Learning: An Artificial Intelligence Approach (Morgan Kaufmann, Los Altos, CA, 1986) 289-310.
[6] August, S.E. and Dyer, M.G. Understanding analogies in editorials, in: Proceedings Seventh Annual Conference of the Cognitive Science Society, Irvine, CA (1985) 845-847.
[7] Bayman, P. and Mayer, R.E., A diagnosis of beginning programmers' misconceptions of BASIC programming statements, Cominun. ACM 26 (9) (1983) 677-679.
[8] Becker, J.D., The modeling of simple analogic and inductive processes in a semantic memory system, in: Proceedings UCA1-69, Washington, DC (1969) 655-668.
[9] Becker, J.D., A model for the encoding of experiential information, in: R.C. Schank and K.M. Colby (Eds.), Computer Models of Thought and Language (Freeman, San Francisco, CA, 1973) 396-435.
[10] Brown, A.L. and Campione, J.C., Three faces of transfer: Implications for early competence, individual differences, and instruction, in: M. Lamb, A.L. Brown and B. Rogoff (Eds.), Advances in Developmental Psychology (Erlbaum, Hillsdale, NJ, 1985).
[11] Brown, R., Use of analogy to achieve new expertise, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA (1977). J2. Brown, R., Words and Things (Free Press, Glencoe, IL, 1958).
[12] Burstein, M.H., Concept formation through the interaction of multiple models, in: Proceedings Third Annual Conference of the Cognitive Science Society, Berkeley, CA (1981) 271-273.
[13] Burstein, M.H., Concept formation by incremental analogical reasoning and debugging, in: Proceedings International Machine Learning Workshop, Monticello, IL (1983) 19-25.
[14] Burstein, M.H., Concept formation by incremental analogical reasoning and debugging, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell (Eds.), Machine

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS…..

Learning: An Artificial Intelligence Approach (Morgan Kaufmann, Los Altos, CA, 1986) 351-370.

[15] Carbonell, J.G., Invariancc hierarchies in metaphor interpretation, Proceedings Third Annual Meeting of the Cognitive Science Society, Berkeley, CA (1981) 292-295.

[16] Carbonell, J.G., Metaphor: An inescapable phenomenon in natural-language comprehension, in: W.G. Lehnert and M.H. Ringle (Eds.), Siategies for Natural Language Processing (Erlbaum, Hilldale, NJ, 1982) 415-435.Applegate, D. L.; Bixby, R. M.; Chvátal, V.; Cook, W. J. (2006), The Traveling Salesman Problem, ISBN 0691129932 .

[17] Carbonell, J.G., Learning by analogy: Formulating and generalizing plans from past experience, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell (Eds.), Machine Learning: An Artificial Intelligence Approach (Tioga, Palo Alto, CA, 1983) 137-162.

[18] Carbonell, J.G., Derivational analogy in problem solving and knowledge acquisition, in: Proceedings International Machine Learning Workshop, Monticello, IL (1983) 12-18. COMPUTATIONAL APPROACHES TO ANALOGY 117

[19] Carbonell, J.G., Derivational analogy: A theory of reconstructive problem solving and expertise acquisition, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell (Eds.), Machine Learning: An Artificial Intelligence Approach (Morgan Kaufmann, Los Altos, CA, 1986)
371-392.

[20] Carbonell, J.G., Michalski, R.S. and Mitchell, T.M., An overview of machine learning, in: R.S. Michalski, J.G. Carbonell and T.M. Mitchell (Eds.), Machine Learning: An Artificial Intelligence Approach (Tioga, Palo Alto, CA, 1983) 3-24.

[21] Carrol, .I.M. and Mack, R.L., Metaphor, computing systems, and active learning, RC-9636(42545), IBM Watson Research Center, Yorktown Heights, NY (1982).