THE NEW BAYESIAN SURPRISE RARENESS (BSR) MEASURE FOR COGNITIVE AGENTS' REASONING AND DECISION MAKING

DAVINNA JEREMIAH

UNIVERSITI SAINS MALAYSIA

2008

THE NEW BAYESIAN SURPRISE RARENESS (BSR) MEASURE FOR COGNITIVE AGENTS' REASONING AND DECISION MAKING

by

DAVINNA JEREMIAH

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

December 2008

ACKNOWLEDGEMENTS

I wish to express my gratitude firstly to God my Heavenly Father who led me through this entire process of writing this thesis and also throughout my research study. I would like to thank Him especially for His faithfulness in enabling me and equipping me in all ways to begin this research and go through each and every phase of it. Truly I could have never begun or completed this research without Him. I truly appreciate Him for blessing me in so many ways which words would be inadequate to describe.

A special thanks also to my supervisor Dr. Chan Huah Yong who indeed in every way was very helpful. I would like to thank him for his support and understanding. Besides that I would like to specially thank him for patiently guiding me and for giving valuable insights which had contributed much to the betterment of my work. His constructive criticsms especially have developed much of my analytical skills and honed my writing ability. I would also like to sincerely thank Ms. Norliza Hani and Mr. Tan Ewe Hoe of UTMK who were both very obliging in sacrificing their time in helping me with the Malay language translation.

I also would like to express my appreciation to my dear husband who has sacrificed in so many ways so to that I could concentrate on my studies. I thank him for accompanying me through the many nights which I had spent writing this thesis. His constant moral support, encouragement and patience gave me much confidence from day to day. Not forgetting my wonderful parents, I would like to specially thank them for being the best parents anyone could have. I thank God for their lives and for their unceasing prayers, comfort, support and love. Truly they had played a major role in moulding me from young which has produced the necessary qualities in me today to take me through every challenge I faced as a student.

Last but not least, I would also like to thank all those who have helped in one way or another and those who have contributed in making my journey as a student a memorable one.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

B2B Business-to-Business

BDI Belief, Desire, Intention

BSR Bayesian Surprise Rareness

BSRIID Bayesian Surprise Rareness Incorporated Influence Diagram

CPD Conditional Probability Distribution

DDM Data Distance Measure

EU Expected Utility

MA Mildly Altruistic

MAS Multi-Agent System

MB mildly Bold

MC Mildly Cautious

MEU Maximized Expected Utility

MS Mildly Selfish

VA Very Altruistic

VB Very Bold

VC Very Cautious

VS Very Selfish

LIST OF SYMBOLS

$\max a$ maximizing operator to find the maximum value of variable a

$\arg\max_{i} a$ argument maximizing operator to find the parameter i which gives the maximum value of variable a

$\max_{i} a$ argument maximizing operator to find the parameter i which gives the maximum value of variable a

$\min a$ minimizing operator to find the minimum value of variable a

P(M) the prior probability of a model M

P(D|M) the likelihood probability of observing D if model M holds

P(M|D) the posterior probability of M after observing data D

P(D) the probability of observing data D

$\mathbf{D}_r$ the data received from the respondent agent at inference round r

μ the expected data of the likelihood distribution

σ the standard deviation of the likelihood distribution

$\mathbf{P}_{cp}$ the cumulative maximum posterior probability

$\mathbf{P}_{cpi}$ the cumulative maximum posterior probability for a particular agent behaviour subtype i

$\mathbf{P}_{cpf}$ the cumulative posterior frequency

$\mathbf{P}_{cl}$ the cumulative maximum normalized likelihood

$\mathbf{P}_{cli}$ the cumulative maximum normalized likelihood probability for a particular agent behaviour subtype i

$\mathbf{P}_{clf}$ the cumulative likelihood frequency

$\mathbf{P}_{avg}$ the average probability of the mean data for every likelihood distribution

$\mathbf{D}_{\mu}$ the mean data of the likelihood distribution

$\mathbf{P}_{min}$ the minimum probability of all the likelihood distributions for all the agent behaviour subtypes

$\mathbf{P}_{th}$ the threshold probability which is the probability value that determines whether a given probability is small enough to be considered as anomalous

γ_p the level of partial anomaly

γ_t the level of total anomaly

ref the reference model

$\mathbf{D}_{KL}(U||L)$ The Kullback-Leibler divergence of U from L

$\mathbf{D}_{KL}(L||U)$ The Kullback-Leibler divergence of L from U

β_{ref} The normalized likelihood for data D_r conditioned on the reference model

β_{high} The normalized likelihood probability for data D_r conditioned on the rare model

UKURAN BARU KEHAIRANAN DAN KETIDAKLAZIMAN BAYESIAN UNTUK PENAAKULAN DAN PROSES PEMBUATAN KEPUTUSAN AGEN KOGNITIF

ABSTRAK

Kehairanan Bayesian (Bayesian Surprise) ialah satu kajian tentang kejadian kehairanan dan sebagaimana yang pernah dikaji oleh pengkaji-pengkaji lain sebelum ini, kejadian-kejadian yang jarang berlaku dianggap sebagai suatu yang menghairankan juga. Namun, daripada kajian ini, kami dapati bahawa kehairanan dan ketidaklaziman adalah berbeza dari segi konsep, maka itu kedua-dua tanggapan ini telah dipisahkan. Memandangkan tanggapan kehairanan dan ketidaklaziman kini tidak lagi dianggap sama, kami telah mereka bentuk kaedah pengiraan yang baru dan berasingan bagi setiap tanggapan dan kami namai kaedah kami Ketidaklaziman Kehairanan Bayesian (Bayesian Surprise Rareness - BSR). Pada dasarnya, pengiraan kami adalah berasaskan taburan kemungkinan, tidak seperti kajian-kajian lalu yang mengukur kehairanan daripada taburan sebelum dan taburan selepas. Pengiraan berasaskan kemungkinan ini dilakukan oleh sebab taburan sebelum dan selepas tidak dapat menggambarkan setiap satu kejadian kes-kes yang jarang berlaku. Pengiraan kehairanan juga adalah berasaskan taburan kemungkinan memandangkan kaedah taburan mampu untuk benar-benar menggambarkan nilai kebarangkalian suatu data yang diperhatikan. Konsep serta pengiraan ketidaklaziman dan kehairanan kemudiannya diaplikasikan pada sistem multiagen yang akan meramalkan perilaku agen. Suatu agen yang sedang diramalkan mungkin mempamerkan perilaku yang tidak lazim dan mungkin memberikan tindak balas yang menghairankan yang perlu dikenal pasti serta dinyatakan kuantitinya. Selain daripada itu, kami juga menunjukkan kepentingan uku-

ran BSR dalam proses pembuatan keputusan suatu agen dengan menunjukkan kebergunaannya dengan menggunakan hasil pengiraan tersebut untuk menentukan strategi optimum agar keputusan yang tepat boleh dibuat. Kami menggunakan gambar rajah pengaruh untuk memodelkan ramalan kami dan proses pembuatan keputusan. Namun, gambar rajah pengaruh yang sedia ada tidak boleh menggambarkan kaedah-kaedah pengiraan kami dan dengan itu kami telah meluaskannya agar merangkumi pengiraan kehairanan dan ketidaklaziman. Representasi grafik yang baru ini dikenali sebagai Gambar rajah Pengaruh Merangkumi Kehairanan dan Ketidaklaziman Bayesian (Bayesian Surprise Rareness Incorporated Influence Diagram - BSRIID). Kajian-kajian telah dijalankan untuk mengesahkan ketepatan kaedah-kaedah pengiraan kami dan sebagaimana yang dilihat dalam suatu situasi khas, proses pembuatan keputusan kami yang menggunakan BSRIID untuk memilih sejenis cadangan tidak langsung telah menunjukkan ketepatan setinggi 14.48% berbanding dengan cara yang tradisional. Kemudian apabila taburan selepas digunakan sebagai kaedah pengiraan untuk mengesan kejadian kehairanan dan ketidaklaziman, hanya sebanyak 9% dapat dikesan tetapi apabila kaedah BSR digunakan, setiap satu kejadian mampu dikesan.

THE NEW BAYESIAN SURPRISE RARENESS (BSR) MEASURE FOR COGNITIVE AGENTS' REASONING AND DECISION MAKING

ABSTRACT

Bayesian Surprise is the study of surprise occurrence and as previously studied by others, occurrences that were rare are considered surprising as well. However, through our study we have found surprise and rareness to be conceptually different, thus both these notions have been separated. Since the notions of surprise and rareness are now no longer considered the same, we therefore have designed separate and new methods of computation for each notion and we have named our method as Bayesian Surprise Rareness (BSR). Basically our computation is based on the likelihood distribution unlike past works which measure surprise from the prior and posterior distributions. The computation is based on the likelihood is due to the prior and posterior distributions being not able to reflect every single occurrence of surprise and rare cases. The concepts and computation of rareness and surprise are then applied to a multi-agent system where the behaviours of agents are predicted. An agent that is being predicted may exhibit rare behaviour and may give surprising responses which need to be identified and quantified. Apart from that we also show the importance of the BSR measure in an agent's decision making process through demonstrating its usefulness by using the result of the computation to determine the optimal strategy so that accurate decisions can be made. We use influence diagram to model our prediction and decision making process. However the existing influence diagram is not able to reflect our computation methods and therefore we have extended it to accommodate the surprise and rareness computations. The new graphical representation is

known as the Bayesian Surprise Rareness Incorporated Influence Diagram (BSRIID). We have performed experiments to verify our computation methods and in one instant, decision making for selecting a hint type using BSRIID is 14.48% more accurate then the traditional or standard influence diagram. Then when the posterior is used to detect events that are surprising or rare, only 9% are detected but as for the BSR measure, all events are detected.

CHAPTER 1

INTRODUCTION

In the field of artificial intelligence, researches that are related to probabilistic reasoning such as those that are based on Bayesian statistics have been gaining momentum in the recent years and it has been viewed as an area of much potential.

Another field in artificial intelligence that has been gaining attention is the study on intelligent agents which has been garnering a wide interest among researchers and it is an area that is extensively researched upon thus bringing many new developments.

In the research that we have performed, our scope of study is mainly within the area of Bayesian reasoning and decision making under uncertainty with minor inclusion of other areas that are also related to Bayesian statistics. It is important to note that Bayesian statistics is a wide area of research which has accorded numerous studies and extensions even from areas that are of different domain.

In the following sections, we describe our research background and this is followed by the problem definition, objective, research methodology and proposed solution. This chapter will end with summary of research contributions and a description on the thesis outline.

1.1 Research Background

The agent environment of our research is a multi-agent system where every agent belongs to a particular organization structure (Kollingbaum and Norman, 2005; Hoogendoorn et al., 2004;

Jonker et al., 2003). An organization would have its own organizational beliefs, desires and intentions or goals which we called them as the external beliefs, desires and intentions. In our environment all the agents are socially committed to adhere to the beliefs, desires and intentions of their organization.

In this research we study how an agent of an organization could predict another agent's behaviour through its communication with it. The goal or intention of the predictor agent is to find and select another agent which does not belong to its organization but has the right behaviour that is acceptable by the organization. However in the multi-agent domain not all the agents behave in the exact straight-forward manner. There may be several agents occurring with rare and surprising behaviours every now and then and the main purpose of this study generally is to identify and quantify occurrences of surprise and rareness in a consistent manner. The prediction process of the agents is performed through inference or update using Bayes' theorem which is an increasingly popular graphical framework for Bayesian reasoning, a probabilistic approach to inference based on combining prior knowledge with observed data.

Other than prediction of behaviour along with identification of surprise and rareness, the predictor agent in our application also has the task of making decisions. The decisions that it makes need to also consider the occurrences of surprise and rareness so that only the right decisions will be made. For this purpose, the influence diagram is used to represent the decision making of the agent. However we have discovered that the influence diagram is unable to handle sequential inference process with decision making performed at the last inference iteration or inference round. Hence to overcome this problem, the Bayesian Surprise Rareness Incorporated Influence Diagram (BSRIID) is introduced.

1.2 Problem Definition

The following are the limitations that we have encountered in the existing methods:

i. In order to perform prediction, the Bayesian inference is used where the final maximum posterior value is regarded as the final prediction result. However the final maximum posterior value may not accurately reflect the actual behaviour of the agent being predicted. This is because it is usual to find that every now and then between few inference rounds, other models are instead represented as the maximum posterior. The inconsistencies such as this where different models are involved at every inference round show that it is necessary to have a method to track the occurrence of these inconsistencies.

ii. If any of the previous methods were to be used, surprise and rareness that should have been identified in the agent's response would not have been consistently identified due to their method of computation being different from ours as it is based on concepts that are also theoretically different from ours.

iii. The work by Bayesian Surprise (Itti and Baldi, 2005a) considers rare occurrence to be surprising. However we find that rareness and surprise are two different notions and either one might occur by itself or it could also occur together. The cause of rareness is found to be different from the cause of surprise and hence both these notions should not be combined and since it is based on different causes it then should be computed separately.

iv. The most common form of graphical representation for decision making process which is the influence diagram (Howard and Matheson, 2005) is found not to be suitable for the decision making process that we have implemented as it cannot represent a decision making process which occurs once at the end of a sequential prediction.

1.3 Objective

The objective of this research generally is to overcome the limitations that have been encountered and listed in the previous section:

i. Devise a method to obtain accurate inference result. To achieve this we base our prediction result on a model which we called as the reference model that takes consideration of all the posterior and likelihood probabilities involved in the inference.

ii. Devise a method that is able to consistently identify and quantify all surprising and rare events in order to contribute to more accurate prediction result. The method of measuring is known as the Bayesian Surprise Rareness(BSR) measure.

iii. Devise a new method of measuring which considers surprise and rareness as different notions.

iv. Devise a new decision making graphical representation which is able to handle all the agents' application requirement which is known as Bayesian Surprise and Rareness Incorporated Influence Diagram(BSRIID).

1.4 Research Methodology

To achieve the overall objective of accurate prediction and decision making result, we have designed a new computation called the BSR measure which is carried out during prediction process that is performed through Bayesian inference. The method of our design is then applied on an agent application where agents have roles such as predictor and respondent and where prediction of behaviour will be carried through a sequential question and answer session. The result of the BSR measure will also be used in the decision making process which is carried out by the predictor agent. To verify the performance of our measure, we have performed experiments and also perform comparison between our work and the related works.

1.5 Proposed Solution

In this section we firstly describe what is actually meant by surprise and rareness which are the two concepts that contribute to our propose measure. Our description is within the context of a multi-agent system where our BSR measure has been implemented.

So to illustrate what is meant by rareness in an agent environment, let's say that the respondent agent for some time had been consistently predicted as an agent that is very altruistic. We need to recall that the prediction is performed through a series of questions and answers by the predictor agent. That means the respondent agent had been giving a series of answers that each time consistently categorizes it to be of the particular behaviour subtype called very altruistic. But let's say there was once or twice during the question and answer session, the respondent agent was found to be giving answers that had categorized it to be of the behaviour subtype that belongs to another kind such as the very selfish behaviour subtype, then in this case to the predictor agent, a rare occurrence is considered to have happened and that which is based on its belief model. This is because the predictor agent had been updating its belief based on a series of incoming data that had consistently deduced the respondent agent to be of the very altruistic behaviour subtype. However there were several times where the data from the respondent agent's answers had caused the occurrence of another behaviour subtype which so far had not occurred or seldom occurred to be inferred. So the predictor agent regards this as rare occurrences as it now has to consider the case where the respondent agent may have infrequent occurrence of some other kind of behaviour subtypes. Therefore it has to take note of these rare occurrences and update its belief accordingly.

As for surprise let's say that the respondent agent has a belief model for every answer that it receives with reference to the behaviour subtype. And from its belief model it has the knowledge that the respondent agent will always answer according to its behaviour subtype. Therefore a certain type of answer may refer more to a certain particular kind of behaviour and for some other answers it may reflect other kinds of behavior subtype. However from its experience, the predictor agent also believes that there are some types of answers given by the respondent agent as a response to a certain type of questions which are extremely unlikely for any respondent agent belonging to any kind of behaviour subtype or at least the behaviour subtype that is expected to be predicted. So such occurences are considered as surprising to the predictor agent. In short the predictor agent considers surprise has happened when there is an occurrence of an event which is extremely unlikely to happen and which is due to the answer given by the respondent agent of any kind of behaviour subtype or at least the behaviour subtype that is expected to be predicted.

So the BSR measure that we have designed is carried out during a prediction process performed by the predictor agent where a series of questions will be posed to the respondent agent. We make the assumption that the other agent called as the respondent agent will answer every question asked, so therefore no questions will be unanswered. Every answer by the respondent agent will be used as evidence to update the belief of the predictor concerning the behaviour of the respondent agent. It is assumed that the answers will always reflect the respondent agent's behaviour. By the end of the question and answer session, the predictor will generally have a belief that had been updated numerous times through inference and hence the final inference result is due to have some degree of accuracy. This agent behaviour prediction process is illustrated in Figure 1.1 and we can see that surprise and rareness are identified and quantified for every response received from the respondent agent which otherwise might had been overlooked if the BSR measure was not part of the prediction process.

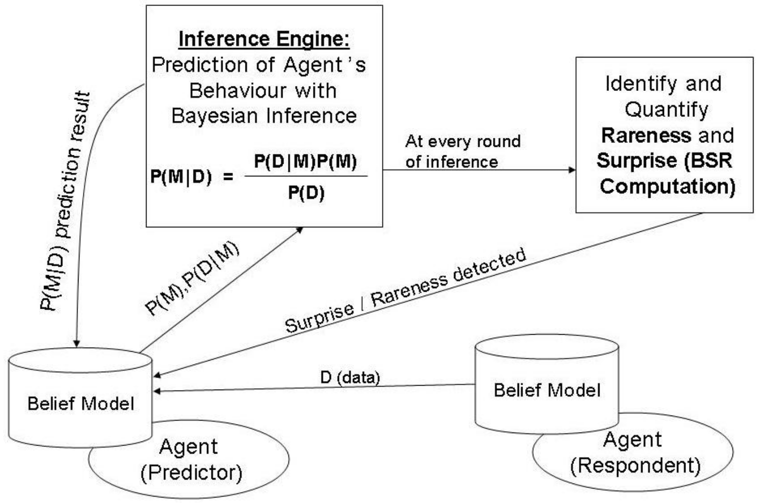

Figure 1.1: Identifying and Quantifying Rareness and Surprise During Agent Behaviour Prediction

The inference process is specifically based on Bayesian theorem and the prediction result will be derived from the result of final inference round. Then appended to this result would be the additional information derived from the BSR computation which will contribute the complete picture concerning the hypothesis being inferred which in this case is the respondent agent's behaviour. With such an implementation, the predictor agent will be alerted of both surprise and rare occurrences of the respondent agent's behaviour that it is predicting. And as a result, the predictor agent is also enabled to perform accurate decision making as the results derived from the BSR computation are used for the decision making process performed by the predictor agent to determine the suitability of the respondent agent to be part of the organization.

As mentioned, the decision making process is graphically represented by the adapted influence diagram called the BSRIID that is able to incorporate the results of the BSR measure. In fact this is part of the effort in achieving our objective of introducing a better decision making approach. Through the decision making process, the importance and the value of the BSR measure are demonstrated and we have shown in our experimentation that without the BSR computation, the decision making process might be erroneous. So with these improvement measures that we have introduced such as the computation methods, the concept re-definition and also the adapted decision making graphical model, our propose method is expected to overcome the limitations that we have mentioned and which has actually been verified through the results derived from our experimentations.

1.6 Summary of Research Contribution

In this research, we have contributed to a new method of computation to identify and quantify surprise and rare occurrence called as the BSR measure which has overcome the limitations found in the Bayesian inference technique and also the previous computation methods.

This new method of computation is based on proposed method where the notions of rareness and surprise are considered as two separate notions and by doing so our work differs from previous work which considers both this notions to be the same. Now with the notions being separated has lead to the computation of rareness and surprise to be separated as well.

However as mentioned, since the current graphical representation for decision making known as the influence diagram (Howard and Matheson, 2005) is unable to accommodate the BSR concept and computation, we therefore have introduced a novel graphical representation model for decision making called the BSRIID which is an adaptation of the influence diagram. The BSRIID is able to make use of the result of the BSR quantification in the decision making process which in turn prevent erroneous decision making from occurring.

1.7 Thesis Outline

In this chapter, we have presented the scope of our research, the objective, problems encountered and the proposed solution. We also have briefly described the contributions of this research.

In the next chapter, we give a comprehensive overview of our research domain and also describe the extended areas. We discuss the application environment that would be used in this study's implementation. We also include discussion on past works and contributions given by others and highlight the problems faced. Apart from that, we also look into the specific area that we had focused upon and had made our contributions.

In Chapter 3 the BSR computation method is described with specific details on the computational steps and justification of our design.

Chapter 4 presents the experimental results, complexity analysis and thethe implementation of our BSRIID in an agent's decision making process. The chapter ends with describing the recommendations for how the BSR computation can be implemented in different scenarios and conditions.

Chapter 5 concludes the thesis by discussing general issues of our research work which we have done and also the future work. In this chapter we will also revisit the contributions.

CHAPTER 2

LITERATURE REVIEW

This chapter firstly gives readers an understanding of our research domain that is within the field of artificial intelligence, where we give an overview of where our research domain is generally situated besides including some description on its extended areas. We had also included discussions on past works and the contributions made by others besides highlighting the problems that exist.

2.1 A Brief Description of our Research Domain

Before we proceed describing our research area and its extended areas we would like to briefly highlight that this research is conducted within a combination of different fields that are found within artificial intelligence. Our problem definition and proposed solution lie within the area of probabilistic reasoning which is known as Bayesian reasoning and also within the area of decision making under uncertainty. Our solution is then applied to an environment of a multi agent system. The subjects of Bayesian reasoning, decision making and agents are within the area of artificial intelligence (Russell and Norvig, 2003).

The probabilistic reasoning and decision making involved in this study are based on the Bayesian approach that based on statistical methods. Therefore in this thesis, the scope of our study is sometimes generally referred to as Bayesian statistics.

2.1.1 Reasoning with Uncertainty

Bayesian reasoning is one of the methods found under the field of reasoning under uncertainty (Russell and Norvig, 2003). Though initially Bayesian looked promising as a method of uncertainty reasoning that is probabilistic based, it fell out of favour because of the intractable calculations involved and these are due to the exponential number of probabilities needed in the full joint distribution. But later it gained acceptance and then increased in its popularity when efficient inference algorithms were introduced.

During the time when Bayesian lost its popularity other alternative approaches under reasoning under uncertainty were introduced and that includes both probabilistic and logical reasoning approaches. The approaches that were introduced are rule-based methods, Dempster-Shafer theory and fuzzy logic (Russell and Norvig, 2003). However, it is a known fact that among all the approaches the Bayesian method of reasoning is currently found to be more favourable when compared to the other alternatives. In the following a description for each of the alternative approach is given along with the description concerning the limitations found in them which we also use as the basis to support our reason for choosing Bayesian reasoning as the reasoning approach for our research.

i. **Rule-based system:**

   Uses if-then-else rule statements in representing problem-solving knowledge where ‘if’ means ‘when the condition is true’ whereas ‘then’ means ‘actions are to be taken’ and ‘else’ means ‘to take an alternative action if condition is found to be false’ . This method is considered as logical reasoning method (Luger, 1989).

   Limitation: If the environment calls for two rules, it is found that these two rules form a feedback loop and it is unable to keep track of the paths where evidence is propagated (Luger, 1989).

ii. **Dempster-Shafer Theory:**

It is a probabilistic framework that is based on lower and upper bound on probabilities. It is designed to consider the difference between uncertainty and ignorance. This is a probabilistic reasoning method (Russell and Norvig, 2003).

Limitations: Computational complexity grows exponentially with the number of hypothesis and there is a problem of connecting belief to action. And it also faces difficulty in deciding when to acquire evidence if more evidences are needed (Russell and Norvig, 2003).

iii. **Fuzzy logic**

It is a method for 'reasoning with logical expressions describing membership in fuzzy sets' and it can be considered as a truth-functional system. It uses probabilities in its reasoning (Russell and Norvig, 2003).

Limitations: Since it is considered as a truth-functional system, it therefore has the inability to take into account the correlations or anti-correlations among its components propositions (Russell and Norvig, 2003).

We chose our research domain in an area related to probabilistic reasoning because besides being a field that is gaining attention among researchers, it is known and claimed since the 1970s that strict logical reasoning, is impractical for most real-world domain as it has a very limited scope. It is limited because it infers from a proposition that needs to be true but the usual case is we are never able to be certain about a proposition and yet we still need to perform inference based on uncertain and incomplete knowledge. This is especially true in human reasoning and due to this reason probabilistic reasoning is widely used especially in the medical domain.

Probabilistic reasoning means reasoning under uncertainty through representations and inference algorithms with accordance to the laws of probability theory (Russell and Norvig, 2003). Probabilistic reasoning could be based on the current state of the environment alone or it could be based on both the current state and the previous state. When the previous state is considered along with the current state, then this method is known as probabilistic reasoning over time. The reason that the previous state is considered is because the previous state has an influence on the current state and this is actually the Markov assumption (Russell and Norvig, 2003). There are several methods of probabilistic reasoning over time and those widely used are Hidden Markov Models, Kalman Filter and Dynamic Bayesian Network. The Hidden Markov Model(HMM) is a statistical tool for modeling generative sequences that can be characterized by an underlying process generating an observable sequence. HMMs have been applied in many areas such as signal processing, and in particular speech processing (Blunsom, 2004). Kalman Filters is a method of representation and inference algorithms that is able to estimate the state of a physical system from noisy observation over time whereas Dynamic Bayesian Network is actually a Bayesian network that has been specially designed to represent temporal probability model where states are dynamically and rapidly changing over time (Russell and Norvig, 2003).

The main focus of this study is to introduce methods of computation for surprise and rareness. The method of computation that we introduce could actually be applied in either type of probabilistic reasoning that is with time involvement or without. However we have chosen an application that although makes use of time slice, each of the time slices is independent from one another. Therefore the time slice in our work is not based on the Markov assumption. Therefore any of the probabilistic reasoning over time methods mentioned above is not suitable for our application. Hence we are using an object-oriented method (Bangso and Wuillemin, 2000) to graphically represent our application. And among the various types of probabilistic reasoning methods which we have discussed in this section, we had chosen the Bayesian approach because the other probabilistic reasoning approaches had limitations which we have described above.

2.1.2 The Bayesian Approach

In this section, we will describe all the various sub-areas found under the area of Bayesian statistical reasoning and through the description given, one would be able to understand the scope which has many different sub-areas existing which had been researched upon. However before describing each of the sub-areas we would first like to give a general description on Bayesian reasoning.

Bayesian Reasoning

The probabilistic reasoning of Bayesian has a qualitative aspect and also a quantitative aspect to it. The qualitative aspect is the probability distribution and the quantitative aspect is the graphical model which is also known as the Bayesian network (Kjaerulff and Madsen, 2005).

Bayesian reasoning is based on probability theory that refers to subjective probability instead of the frequentist probability. Briefly, subjective probability is belief based probability and frequentist probability is based on empirical results where its value is based on the number of event occurrences (Leonard and Hsu, 1999).

Bayesian Network

i. Definition (Jensen, 2001): A Bayesian network consists of the following:

   - A set of variables and a set of directed edges between variables.
   - Every variable has a finite set of mutually exclusive states.

- The variables and the directed edges form a directed acyclic graph(DAG). (A directed graph is acyclic if there is no directed path).
- Attached to each variable A with parents B_1,..., B_n is a potential table $P(A| B_1,..., B_n)$.

ii. Reasons to choose Bayesian network as the modeling framework:

- Intuitive and compact representation of cause-effect relations and conditional dependence and independence relations.
- Efficient solution to the queries given the occurrence of evidence.
- Ability to support different types of analysis for the result produced. These analysis are such as conflict analysis and sensitivity analysis.
- Coherent and mathematically sound handling of uncertainty.

iii. Causal independence in Bayesian network

With causal independence the number of variables involved in the joint probability computation are reduced, thus making computation more tractable.

Object-oriented Bayesian Network

In the recent years there had been extensive research on various adaptations to Bayesian network. Such adaptations were introduced mainly with the purpose of having a network that is suitable for the application to be implemented. It is important to note that Bayesian network can at times be found limited and therefore it is not able to properly represent the application environment. So the object-oriented Bayesian network (Bangso and Wuillemin, 2000) has proven itself to be useful especially for applications that not only evolve over time but are complex with several different occurrences of instantiations at a given time. Take note that these instantiations would not be well represented by a normal Bayesian network.

An object-oriented Bayesian network is a framework to support systems that are composed of similar or identical components (Kjaerulff and Madsen, 2005). So it allows a very compact specification of knowledge especially if it contains repetitive structures (Bangso and Wuillemin, 2000). In another words it allows the Bayesian network to be implemented piece by piece and to use a piece any number of times in any step of the construction process, meaning the pieces are reusable (Bangso and Wuillemin, 2000).

The object-oriented Bayesian network actually has several variations to it as it had been studied by many and mostly contributed by (Bangso and Wuillemin, 2000; Koller and Pfeffer, 1997).

Decision Making

Usually the probabilities provided by the Bayesian network are used to support some kind of decision making. There are several types of graphical representations for decision making and those commonly used are decision trees (Jensen, 2001) and influence diagram (Howard and Matheson, 2005). Lately influence diagram is beginning to be widely used. In a way it is similar to Bayesian network as it is also represented as a causal model plus it describes the dependence relations between entities of a domain. The difference of the influence diagram from the Bayesian network is the involvement of precedence ordering which specifies the order on the decisions, the observations and the preference of the decision maker. Besides that it also has additional nodes and these are the decision nodes and utility function nodes. The decision nodes represents the various actions which the decision maker would take whereas the utility function represents the preferences of the decision maker. Basically the steps to solve a decision problem are: determine the optimal strategy that had maximized the expected utility and compute the maximal expected utility adhering to this strategy (Kjaerulff and Madsen, 2005).

Modeling Techniques

Many modeling techniques have been introduced to suit the different types of environment that the network is representing as some environments have large set of cases and are complex. There could be techniques to cater for cases where there are a large number of parent nodes in the network and where all of these nodes point to a single child node. Other than that there are also techniques to cater for cases where the parent node configuration is too small or where the domain that is involved evolves over time.

To specifically cater for these varied requirements, there are many different modeling techniques and we name a few such as the divorcing technique, noisy-or, experts disagreement, etc. (Jensen, 2001)

Inference

There are several inference algorithms to perform reasoning under uncertainty and we describe them as follows:

i. **Bayes theorem**

   Bayes theorem bases its inferred result on the posterior probability. The posterior probability states the probability of the model hypothesis given the occurrence of data or event. The result of the posterior probability $P(M|D)$ which is the probabability of the hypothesis model or model given data is derived from the multiplication of the likelihood probability and the prior probability. The likelihood probability $P(D|M)$is the probability of the event given the occurrence of a model whereas the prior probability $P(M)$ is the probability of the hypothesis model based on the prior belief one had concerning the hypothesis. The result of the multiplication is then divided by the normalized value of the likelihood probability of the observed data of all the hypothesis models. And the variable that is normalized or summed is the variable referring to the model.

ii. **Exact Inference**

The joint probability in the computation of Bayes theorem increases exponentially with the number of variables. A joint probability is needed to answer all possible inference by summing or marginalizing irrelevant variables. However because it increases exponentially due to the fact that a joint probability distribution is of the size $O(2^n)$, where n is the number of nodes and assumption is made that each nodes may have 2 states, more efficient exact inference methods are needed. Examples of these methods are junction trees (Pearl, 1982) and variable elimination (Zhang and Poole, 1994). However no methods guarantee a tractable calculation task.

iii. **Approximate Inference**

Exact inference may be intractable for multiple connected, repetitive structured Bayesian networks. Therefore in such cases approximate inference is needed. Examples of approximate inference are likelihood weighting, Gibbs sampling and loopy belief propagation (Murphy, 1998.).

Statistical Learning

Statistical learning is learning the probability theories of a domain from experience. The Bayesian view of learning is able to give general solutions to the problem of noise, overfitting and optimal decision (Russell and Norvig, 2003). There are basically four types of learning Bayesian networks from data:

- Known structure and observable variables
- Unknown structure and observable variables
- Known structure and unobservable variables
- Unknown structure and unobservable variables

Learning can also be understood as parameter learning and structure learning where the former is the estimation of the conditional probabilities and the latter is the estimation of the links or topology of the network (Hertzman, 2004).

Analysis Tools

Analysis tools are used to analyse whether the evidence entered into the Bayesian network is coherent as there may be occurrence of flawed data that needs to be traced. Besides that analysis tools are also needed to identify evidence that are in favour or against a hypothesis and the tool for this is called conflict analysis. Analysis tools are also used to analyse which parts of the evidence have an impact on the hypothesis and it is also used to identify the parameters that are most influential on the posterior probability of a hypothesis given the evidence. The tool for such purpose is called sensitivity analysis. There is also a tool called value of information analysis that is used for analysing the potential usefulness of additional information before the information source is consulted (Kjaerulff and Madsen, 2005).

Bayesian Surprise

The function of this analysis is to identify and quantify the occurrence of surprise with the computation based on the prior, posterior or likelihood distributions (Itti and Baldi, 2005; Box, 1980; Rubin, 1984; Bayarri and Berger, 1997).

Bayesian statistics distribtution : The Beta distribution

Beta distribution is commonly used in Bayesian parameter learning (Russell and Norvig, 2003) and it is defined by parameters **a** and **b** such that

$$beta[a,b](K) = \alpha K^{a-1}(1-K)^{b-1}$$

For K is in the range [0,1] (Russell and Norvig, 2003).

Beta distribution is widely used in Bayesian statistics since the beta distribution is the conjugate prior distribution to the binomial distribution. ‘In Bayesian probability theory, a conjugate prior is a family of prior probability distributions which has the property which specifies that the posterior probability distribution will also be belonged to the family of prior probability’ (ML Pedia) So the beta family has this wonderful property where if a *K* has a prior *beta(a,b)*, then after a data point is observed the posterior distribution for *K* is also a beta distribution’ (Russell and Norvig, 2003).

Before we illustrate how a beta distribution is used in a Bayesian parameter learning, we would like to show examples of beta distributions with different parameters for *’a’* and *’b’* and it is as follows:

i. The diagram below, Figure 2.1 is an example of beta distribution with a uniform distribution. A uniform distribution will have the parameters *’a’* and *’b’* of *’beta(a,b)’* as *’beta(1,1)′*.

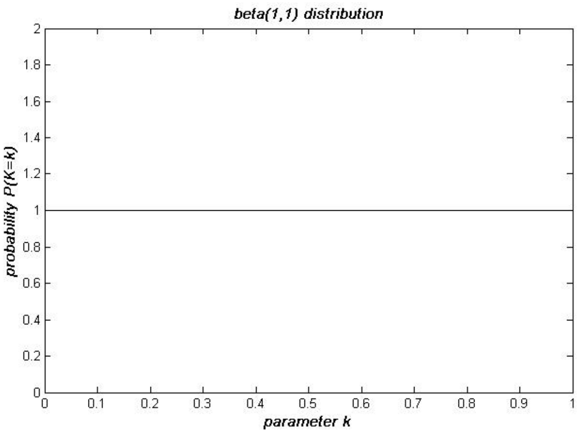

Figure 2.1: The graph of Beta(1,1)

ii. The distribution of *'beta(1,5)'*, *'beta(5,1)'*, *'beta(3,3)'* and *'beta(8,10)'* are respectively shown in Figures (2.2(a),(b),(c) and (d)).

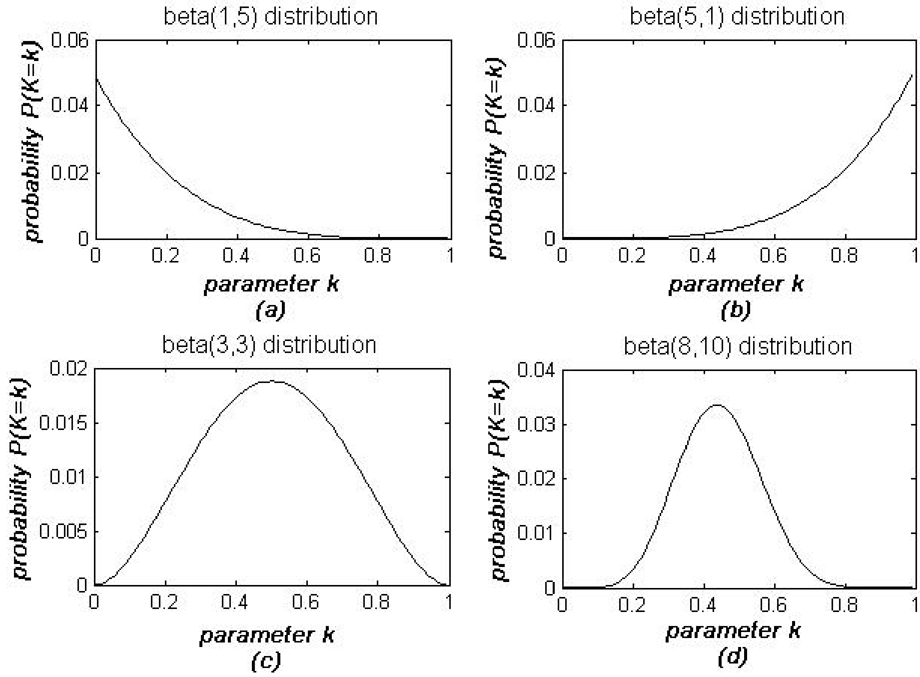

Figure 2.2: Beta distributions representing various types of parameter values,

'Let's say the x-axis represents the parameter K which is the probability that a randomly selected piece of candy is cherry flavoured and now suppose we observe a cherry candy, then:

$$P(K|data1 = cherry) = P(data1 = cherry|K)P(K)$$

$$= \alpha' K \cdot beta[a,b](K)$$

$$= \alpha' K \cdot K^{a-1}(1-K)^{b-1}$$

$$= \alpha' K^{a}(1-K)^{b-1}$$

$$= beta[a+1,b](K)$$

Thus, after seeing a cherry candy, we can simply increment the *''a'* parameter to get the posterior; similarly, after seeing a lime candy, we increment the *'b'* parameter. By examining a sequence of beta distributions for increasing values of *'a'* and *'b'*, keeping the proportions fixed, we can see how posterior distribution over the parameter *K* changes as data arrive' (Russell and Norvig, 2003). So to put it simply, in order to obtain the posterior by updating the prior, all we do is add to the parameters of *'a'* and *'b'*.

2.2 The Chosen Application Domain

The purpose of designing an agent scenario as described in the following is to show the usefulness of the surprise and rareness computation. So the agent environment is specifically designed to depict a real-world environment as much as possible so that it will highlight plus demonstrate the usefulness, capability and effectiveness of the method of our computation.

We chose an agent based environment based on concepts such as cognition, social influence, social reasoning, organization structure, etc. as these are well studied areas which are still undergoing extensive research. Besides that, it also has the capability to contribute to other types of domain such as the field of science and education.

In this section we give an overview of the agent that has been studied by others and as it is a large area of study with rapid, extensive development, it is therefore impossible to cover every area. We therefore give a general introduction on the agents and then describe the features and capabilities found in them. We have focused our discussion not only on the state-of-the-art features but those features that we have adapted in our design of agents.

Once again we stress on the point that the agent environment described here is not the main focus of this research and it is not the area of contribution as well instead it is used to describe the usefulness of our contributions in this research. Therefore in our agent design, we have selected and incorporated those features that are able to highlight our contributions as a solution to the problems which we had found and which we had mentioned in Chapter 1.

2.2.1 Software Agents

An agent can be defined as a 'computer system that is situated in an environment, and that is capable of autonomous action in this environment in order to meet its design objectives' (Weiss, 1999). We consider an agent to be intelligent when it is equipped with the capability to flexibly perform autonomous actions that would meet its design objectives. The flexibility of an agent can be understood as the agent's reaction which is its capability of responding as it perceives the environment and it does it in a timely fashion to satisfy its objectives. Flexibility can also mean that the agent's proactiveness in taking initiatives which demonstrates its goal directed behaviour. Besides that flexibility it is also the agent's social ability which is the agent's capability of interaction with other agents or humans.

There are several types of agent architectures such as logic-based agents, reactive agents and Belief-Desire-Intention (BDI) agents. Logic-based agents make decisions through logic based deductions whereas reactive agents' decision making process is through some form of direct mapping from situation to action. An agent with Belief-Desire-Intention (BDI) architecture makes decision based on the manipulation of data that are based on the beliefs, desires and intentions of the agent.

In our work we design our agent with reference to the BDI architecture which is based on the cognitive concepts of human behaviours and actions. BDI architecture has been known to be a popular architecture for agent design. Huhns and Singh (1998) states that the basis of ascribing belief, desires and intentions to an agent is due to the fact that the agent design is specified and commanded by human. Since humans actions use cognitive terms such as belief, desire and intention, therefore it is found to be natural for agents to use the same cognitive terms. For an agent to perform proper actions in an ever changing environment, the agent needs to have information of its environment. Due to the fact that all information pertaining to the state of the environment is according to how the agent perceives, this information is therefore termed as the belief of the agent. As for the agent's desire, it describes the preferences of the agent towards the environment state. Finally an agent's intention is pertaining to the state of environment that the agent is attempting to achieve. Agents intentions should be consistently a subset of agent's desires and it is also should be directly related to agent's actions.

The applicability regarding the cognitive basis for an agent is evident when the agent is commissioned to serve as a personal assistant in a user interface. Besides that, it also useful in simulated environment of complex social phenomena such as the evolution of roles and organization structures for the purpose of investigation of social aspects of intelligence (Weiss, 1999). Examples of agent applications based on this simulated environment are such as training agents or virtual tutors for simulation based on military training or a human performance modeling agents in a dynamic and adaptive human behavorial simulation for the study of human psychology in a tense environment (*CHI System: Cognitive Agent Development*, 2008). And other examples of systems that are based on agent technology are such as automated scheduling coordination, electronic procurement activities by manufacturing supplier and etc.

2.2.2 The Multi-agent System(MAS)

During the 1990s, multi-agent system (MAS) emerges as one of the most important research and development area that had seen a rapid growth. Multi-agent system is where we find many intelligent agents interacting with one another. The interaction amongst them can either be cooperative or selfish. Cooperative means the agents are willing to share common goal whereas selfish means agent are interested only to pursue their desires. The study of MAS is not just confined within artificial intelligence but ideas are also drawn from other disciplines