Stochastic finite automata: a mathematical model for sequential decision making under uncertainty

Kulandai Therese.A 1,* and James Albert 2

1Karpagam University and Mathematics, Nirmala College for Women, Coimbatore, Tamil Nadu, India.
2Faculty of Arts, Science and Humanities, Karpagam University, Coimbatore, Tamil Nadu, India.

ARTICLE INFO
Article history:
Received: 14 April 2014;
Received in revised form: 21 August 2014;
Accepted: 29 August 2014;

Keywords
Sequential Decision Making,
Uncertainty,
Probabilistic Distribution,
Utility,
Probabilistic Transition System,
Stochastic Finite Automata.

ABSTRACT
Decisions are complex in nature. Decisions may include uncertainty to some extent. Some decisions may be sequential in nature. Decision once made and executed is irrevocable. And so it is imperative that good decision be made at all point of time. Many decisions are made in situations of uncertainty too. Every decision is made within a decision environment, which involves the collection of information, alternatives, values, and preferences of the decision maker at the time of decision.

Sequential Decision Making
In general, Decision Makers do not make decisions in the dark without observing something about the world, nor do they make just a single decision. Decision making involves sequential decision making as and when necessary. The essence of sequential decision making is that decisions made at one point of time may have both immediate and long-term effects[1]. A more typical scenario is that the decision maker (DM) makes an observation, decides on an action, carries out that action, makes observations in the resulting outcomes, then makes another decision conditioned on the observations and so on. Subsequent actions can depend on what is observed on outcomes, and what is observed can depend on previous actions. At this point in time, it is often the case that the sole reason for carrying out an action is to provide information for future actions. Hence it is logical to mention that sequence learning is an important component of learning various tasks, domains of intelligent systems, decision, planning, reasoning, time series prediction and so on[2]. In sequential decision making one needs to take sequence of decisions one after the other. On taking one action, the decision maker attains one state. In general, a sequential decision consists of \( n \) sequential states, independent or interdependent, where decision made at a state is passed on to the next state and the overall decision depends on the decision made at each state.

Uncertainty
Representation and reasoning about actions is a basic component in decision making process. In reasoning about actions of a DM operating in real world environments, one of the most crucial problems that the DM has to face is uncertainty, both at the initial situation of the DM’s world and about the results of the actions taken by the DM. Uncertainty touches most aspects of one’s life, especially when one makes decisions that have consequences which cannot be predicted. Although much research has been done on uncertainty, there is a lack of consensus about the definition. Uncertainty is a general concept that reflects over lack of sureness about something or some one or lack of conviction about an outcome. It can be best described as the gap between the information available and the information a DM wants[3]. It is important to know that mere gathering of information cannot always minimize uncertainty. Uncertainty can also exist when there is excess of information too. Often the factors (such as interests, objective and procedure) that determine an individual’s decision would be different from the others and are uncertain to the others.

An ongoing debate on uncertainty is about objective versus perceived uncertainty. The objective view on uncertainty defines uncertainty as a characteristic of the environment that can be measured objectively. The perceptive view on uncertainty argues that uncertainty is dependent on the individual and cannot be measured objectively[4]. The term ‘perception’ refers to the process by which individuals organize and evaluate stimuli from the environment. What is certain to one person need not be certain to another. The existence of information...
Sequential Decision Problem under Uncertainty consists of a state space and probability distribution governing possible state transitions indicating how the next state of system is related on past states. It is important to understand that when an action is taken at a state, based on the outcome, the decision maker moves on to the next state \( (S) \). It is assumed that the finite set of actions are available to the DM. When an action is performed at one state, the state changes stochastically due to the action[7]. A DM chooses an action on the basis of outcomes that the chosen action produces. Other factors may interact with an action (state of the world) to produce a particular outcome.

\[ A = \{a\} \text{, set of possible well defined actions} \]

\[ S = \{s\} \text{, set of possible states of the world} \]

\[ O = \{o\} \text{, set of outcomes} \]

A combination of an action \( a \in A \) and a state \( s \in S \) will produce a particular outcome \( o \in O \) such that \( f : S \times A \times O \rightarrow S \).

Choosing an action ‘\( a\)’ determines an outcome over state-dependent consequences associated with action ‘\( a\)’ if \( f \) is constant to state of the world, then the decision is taken under certainty. More must be mentioned about the state of nature. With respect to any decision problem, the set of ‘state of nature’ is assumed to form a mutually exclusive and exhaustive listing of those aspects of nature which are relevant to a particular choice of problem about which the decision maker is uncertain. Although the characterization is quite vague, often there is a natural enumeration of possible, pertinent, state of the world in particular contexts. Assume that there is a ‘true’ state of the world which is unknown to the decision maker at the time of making a choice.

The decision problem model for the study is formulated as follows. Each action ‘\( a\)’ when applied from the current state \( s \), produces a new state \( s' \) based on outcome \( o \) as specified by a state transition function \( f \) to express a state transition equation \( s' = f(s,a,o) \). Let \( A(s) \) denote the action space for each state \( s \), which represents the set of all actions that could be applied from \( s \). For distinct \( s, s' \in S \), \( A(s) \) and \( A(s') \) are not necessarily disjoint; the same action may be applicable to the set \( A \) of all possible actions all over the states.

\[ A = \cup A(s) \quad s \in S \]

A DM needs to achieve certain goals. A goal is a description of a set of desirable states of the world Sequential decision problems can be represented as a set of states and a set of rules of how one state is transformed to another. And in these problems the DM must choose a sequence of actions in order to attain his goal state. So the DM can start from the initial state and explore the state space and its objective is to reach one of the goal states \( S_G \)(say) which is a subset of \( S \) i.e. \( S_G \subseteq S \). The task of optimal sequential decision making is to find a finite sequence of actions that when applied, transforms the initial state \( s_i \) to the final state or goal state \( S_G \). The model is summarized as a set of six elements as follows:

\[ \text{Probabilistic Distribution} \]

Uncertainty over events will be modeled through the DM's degree of belief. Probability theory is the rational way to think about uncertainty. This means that a probability distribution must be assigned to every set of uncertain states of nature in the decision problem. Probability Theory is devoted to measuring quantitatively the likelihood that a given event occurs. Two definitions are derived from two different approaches to the concept of probability: subjective and objective[6]. The objective probability viewpoint posits the likelihood that a particular event will occur is a property of the system under study, which is ultimately grounded on the physical laws bearing on the given system. The subjective probability viewpoint argues that the likelihood of the occurrence of a particular event is a measure of the belief of the observer of the system given his/her state of information at the time. It is meaningless to talk about "the actual probability of occurrence" of an event because such a conception is unknowable and impossible to define outside the observer's mental space. It has been successfully viewed by the researchers that subjective beliefs can be modelled as probabilities called subjective probabilities. This allows us to treat uncertainty due to stochasticity and due to partial information in a unified framework for decision making. To extend the discussion from a deterministic world to a stochastic one, it is necessary to specify a probability distribution in place of the deterministic prediction or generation.

Sequential Decision Problem under Uncertainty defined Sequential decision problem involves sequence of decisions.
\[ P = (S, A, O, f, s_1, S_G) \]

Where

- \( S \): A non-empty state space which is finite
- \( A \): For each state \( s \in S \), a finite action space
- \( O \): Set of consequences / outcomes
- \( f \): A state transition function that produces a state \( f(s, a, o) \in S \) for every \( s \in S \), \( a \in A \) and \( o \in O \).

The state transition equation is derived from \( f \) as \( s' = f(s, a, o) \) i.e., \( f : S \times A \times O \to S \)
- \( s_1 \): Initial state
- \( S_G \): A set of final states or goal states and \( S_G \subseteq S \).

**Measuring of Outcomes**

When we make decisions, or choose between outcomes, we try to obtain a good outcome as possible according to some standard of what is good or bad. Outcomes or consequences can be measured in a variety of ways, depending on the stakeholders. There is no consensus to date on the best approach to defining and measuring outcomes and outcome means different thing to different people. Outcomes measurement is defined as the systematic quantitative observation, at a point in time, of outcomes indicators. Although outcomes measures are continually evolving, they provide valid information that can assist DM in determining the optimal decision. Various types of outcomes measures include physical / performance / function / economic and humanistic measurements. The outcomes of an action include only those states of affairs that would be realized if the action was performed and they are probabilistically related to the action.

**Utility**

Utility is the key factor which enables one to make the decision considering all rewards like costs, values and so on that is summed for each state on the way to the goal state[8]. In short, utility is something that satisfies the needs of the DM. Utility theory is a prescriptive theory and it specifies the course of action of a DM. It further allows him to be consistent with his preferences and judgments. In short, it states that if a DM accepts the following axioms of the utility theory, then the course of action which maximizes his expected utility must be used because utility being a numerical scale over the decision maker’s preferences.

**Axiom 1.**

For any two outcomes say \( o_1 \) and \( o_2 \) of an action \( a \) can be ranked as \( o_1 \leq o_2 \) or \( o_2 \leq o_1 \).

**Axiom 2.**

For any three outcomes of an action \( a \) say \( o_1, o_2 \) and \( o_3 \) if \( o_1 \leq o_2 \) and \( o_2 \leq o_3 \) then \( o_1 \leq o_3 \).

Definition: A lottery is defined as the set \( \{(o_1, p_1), (o_2, p_2), \ldots, (o_n, p_n)\} \) such that \( \sum_{i=1}^{n} p_i = 1 \) and \( 0 \leq p_i \leq 1 \). In ordinary lottery the outcome \( o_i \) occurs with probability \( p_i \).

Definition: The utility function of an outcome of an action can be defined as a mapping from \( S \) to \( S \) such that \( U : S \times A \times O \times S \to R \) where \( R \) is a real number.

In the study, utility is a measure to determine the value of an outcome. Here \( u_i \) is the utility associated to the outcome \( o_i \) of an action \( a \). That is, an outcome \( o_i \) of an action \( a \) having utility \( u_i \) leads to a state \( s_1 \), an outcome \( o_j \) of an action \( a \) having utility \( u_j \) leads to the state \( s_2 \) and so on. We can usually assign a subjective preference for each outcome \( o \) of an action \( a \), thus making the function \( u \) well-defined. Value of an outcome can be an arithmetic product of the probability of the occurrence of that outcome and its utility. Values of the same kinds of utility can be compared on the basis of their probability distribution. Thus the value of an outcome is a single numeric value. i.e., \( v_j = u(o_j) \cdot p(o_j) \).

Definition: In any decision making process in the environment, the outcomes of an action \( a \in A \) say \( o_1, o_2 \in O \) and \( p_1 \) and \( p_2 \) are the corresponding probabilities for the occurrence of the outcomes such that \( p_1 + p_2 = 1 \). One shall prefer \( o_1 \) to \( o_2 \) and write \( o_1 > o_2 \) if and only if \( u(o_1) \cdot p_1 > u(o_2) \cdot p_2 \) and \( o_1 = o_2 \) if and only if \( u(o_1) \cdot p_1 = u(o_2) \cdot p_2 \).

**Policy**

Given a finite horizon of size \( n \) a DM executes \( n \) actions at stages \( 0 \) through \( n-1 \) of the process, ending up in a terminal state at stage \( n \). The DM receives value for each state \( s \) passed through at stages \( 0 \) through \( n \) (its trajectory). In choosing the action to perform at stage \( k \) of the process, the DM can rely only on its knowledge of the initial state \( s_0 \) and the history of actions it performed and observations it received prior to stage \( k \). Different observation leads a DM to choose different actions. Thus, a policy can be represented as a mapping from any initial state estimate, and \( k \)-stage history, to the action for stage \( k+1 \). That is, a policy is a function that determines the choice of action at any stage of the system’s evolution. The value of a policy is the expected sum of values accumulated. A policy is optimal if no other policy has larger value. In general, human attitude towards decision making depends on beliefs and desires[9]. They play an essential role in action and decision in Sequential Decision Problems under Uncertainty. Decision mechanisms need to maximize the outcome of choice by comparing the values of all available options and choosing the option which carries highest value.

**Modeling of a Complex System**

All the world systems are complex in nature. Most of the real world systems are dynamical systems containing a large number of mutually interacting and mutually exclusive entities like components, agents, processes and so on whose aggregate activity is not derivable from the summations of the activity of individual entities. At any point in time, a dynamical system is characterized by its state [10]. Changes of the state over time are described by a transition function, which determines the next state of the system as a function of its previous state. The behavior of the system is best characterized in terms of state and its evolution over time.

Any scientific method or approach of studying complex real world systems relies on modeling. A model is a representation of reality. Necessarily, it is a simplification or abstraction. Models can be developed for a variety of reasons that include understanding and learning about the behavior of the system, improving its performance and making decisions about its design or its operation. Complex systems are usually difficult to model, design, and control because the behaviors of complex systems depend on the elements of interactions.

To deal with this complexity of the real world systems, one should make several assumptions about the state space and the
transition function of dynamical systems[11]. One of these assumptions is that the system is observed and controlled at discrete steps of time and the intermediate state of the system between two steps is not relevant for determining future states. The second assumption is that the state space is discrete and finite. Finally, the most important assumption is known as Markov property: the state of the system is sufficient information for predicting its future states by using the transition function. In practice, the dynamics of a system depend on many variables but all cannot be included in a mathematical model, either for computational reasons or simply because they are unknowns. Hence, a stochastic transition function is used to overcome the effect of missing information. Presently stochastic (probabilistic) modeling plays a key role in modeling because dynamic systems become more dependent on random events. Probability theory has turned out to be a powerful means to model and analyze unreliable or unpredictable behavior exhibited by a dynamic system[12].

In general, problems having sequence of actions come under either Deterministic models or Stochastic models. Deterministic models have no components that are inherently uncertain, i.e., no parameters in the model are characterized by probability distributions, as opposed to stochastic models. For fixed starting values, a deterministic model will always produce the same result. A stochastic model will produce many different results depending on the actual values that the random variables take in each realization[9]. The mathematical model used in the study is categorized as Stochastic model. To be more specific, it is a discrete planning model.

In a discrete planning model, actions and behaviour are explained in terms of transition systems. The order in which actions can take is called ‘behaviour’. This simple idea is captured by the notion of Labelled Transition Systems (LTS). Formally, a LTS consists of a set of states, a set of labels (actions), and a transition relation T describing a change of a state. A set of states includes an initial state and also terminating or final states. The most abstract process behaviour can be described as follows: a process p performs an action ‘a’ and becomes a process q. The same can be stated as follows:

\[ p \xrightarrow{a} q \]

LTS: can be defined as follows

Definition: A LTS is a triple \((S, A, Tr)\) where

- \( S \) is a finite countable set of states
- \( A \) is a finite countable set of actions
- \( Tr \) is a transition relation such that \( Tr \subseteq S \times A \times S \)

Transition \((s, a, s')\) can be written as \( s \xrightarrow{a} s' \) which means that \( s' \) is attained from \( s \) by executing an action \( a \).

Processes are considered as agents or DMs that can execute actions in order to communicate with their environment. These actions can be observed by an eternal observer and determine the visible behavior of the process. Processes are understood as nodes of certain edge-labeled oriented graphs and a change of process states caused by performing an action is understood as moving along an edge labeled by the action name. Processes combining nondeterminism and probability can be described by means of extensions of the LTS model, in which every action-labeled transition goes from an initial state to a probability distribution over target states. This work deals with the probabilistic extension of a non-deterministic calculus. So a probability distribution on the set of the possible next moves of

the choice composition is explicitly given. They are essentially Markov decision processes [9].

**Stochastic Transition Systems (STS)**

STS describes the behavior of systems through state transition graphs in which every transition is labeled with both the action and the probability of the corresponding state change which means that each such process can be represented as a discrete-time Markov-Chain[13], whose transitions are additionally labeled with actions. STS features both nondeterministic and probabilistic behavior. It provides a concise and compositional way to describe the behavior of systems in terms of probability[6]. Nondeterministic choices can be specified in transition systems by having several transitions with different labels leaving from the same state. Most of the times probabilistic choice is considered as a refined version of nondeterministic choice. That is, a probability distribution on the set of possible next moves of the choice composition can be given explicitly. As probabilities are associated with nondeterministic choices, it becomes quite natural to assume as underlying semantic model that of Probabilistic Transition System (PTS), namely transition systems whose arcs are labeled by both an event and a probability value[14].

Definition: PTS \((S, A, Tr, p)\) is a LTS extended with a transition probability distribution

\[ p : S \times A \times S \rightarrow [0,1] \]

such that

\[ p = \begin{cases} p(s, a, s'), & \text{if } (s, a, s') \in Tr \\ 0, & \text{otherwise} \end{cases} \]

and \( \sum p(s, a, s') = 1 \), for all \( s \in S \)

PLTS constitute a framework for the description of processes with stochastic behavior. However, they become very large-both in the number of states and in the number of transitions. In order to overcome this problem, Stochastic Finite Automata-a model that allows the representation of such systems in a finite, symbolic way is introduced. The semantics of Stochastic Finite Automata can be defined in terms of PTS.

**Stochastic Finite Automata**

A finite state automaton is a model of computation consisting of a set of states, an input alphabet, an initial state, a set of transition rules and a set of final states. Transitions are the rules in the following form: (current-state) and (condition) then (active-new-state)

The transitions rules may be given by a function or a relation, mapping or relating the current state and the actual input symbol to the next state. A finite state automaton can decide whether an input string is accepted or not. To this end, the finite state automaton performs a computation beginning with the initial state reading the first symbol from the input string. The computation consists of a series of transitions. In each transition, the next input symbol is read from the input string and the current state is changed according to the transition rules to establish a new state. The computation terminates when the automaton has read the last symbol from the input string. The automaton will accept the input string if it terminates in an accepting / final state.

A Finite Automaton can be defined mathematically as follows[15]

Definition: A finite state automaton is a five-tuple

\[ M = (Q, \Sigma, \delta, q_0, F) \]

Where,

\( Q \): a non empty set of finite states \( S \).

\( \Sigma \): a finite set of input symbols.

\( \delta : Q \times \Sigma \rightarrow Q \): a transition function \((\delta)\) mapping from \( Q \times \Sigma \)
into \( Q \), i.e., from a state \( q \in Q \) on taking an input symbol \( s \in \Sigma \) there exists transitions to another state / states \( q_0 \in Q : \) start state or initial state.

\( F : \) a set of final states or accepting states and \( F \subseteq Q \)

The language of the automaton is known as Regular Language which is the set of all accepted input strings over the input alphabet.

A finite state automaton is called stochastic if the transition rules are defined by transition probabilities and initial and final states are defined by probability distributions. A stochastic finite state automaton \( M^* \) is pair \( (M, p) \) can be defined as follows.

\[
M^* = (Q, A, q_0, \delta, p, F) \quad \text{Where}
\]

\( Q : \) is a finite set of states
\( A : \) is a non empty set of actions
\( q_0 \in Q : \) is the initial states
\( F \subseteq Q : \) is a set of final states
\( \delta \subseteq Q \times A \times Q : \) is a finite set of transitions between states
\( p : \) is a function \( \delta \rightarrow [0,1] \) such that for all \( q \in Q \) and for all \( a \in A \), \[ \sum_{q \in Q} p(q, a, q') = 1 \]

The function \( p \) can be generalized to \( p^*: Q \times A^* \times Q \rightarrow [0,1] \) in the following way:

\[
p^*(q, a, q') = 1 \quad \text{if} \quad q = q' \]

\[
= 0 \quad \text{if} \quad q \neq q'
\]

and \( p^*(q, xa, q^{*}) = \sum_{q \in Q} p^*(q, x, q').p(q', a, q^{*}) \)

for all \( x \in A^* \) and \( a \in A \) and \( q, q', q^* \in Q \).

A stochastic finite state automaton \( M^* \) induces a function \( \varrho: A^* \rightarrow [0,1] \) as follows:

\[
\varrho(x) = \sum_{q \in F} P^*(q_0, x, q)
\]

Therefore, \( M^* \) also induces a weighted language in \( A^* \).

It has already been established that the Stochastic Regular Language as a Mathematical Model for the Language of Sequential Actions for Decision Making under Uncertainty [16]. PTS can be extended using the outcomes \( O \) and when treated with the transition probability of an action as the transition probability of the outcomes of an action and it may be called as Modified Probabilistic Transition System (MPTS). Definition: MPTS is a five tuple \((S, A, O, Tr, p)\) in such a way that the transition probability distribution \( p \) is as \( p: S \times A \times O \times S \rightarrow [0,1] \) such that

\[
p = \{ p(s, a, o, s') | (s, a, o, s') \in Tr \}
\]

\[
= 0, \quad \text{otherwise}
\]

and \( \sum p(s, a, o, s') = 1 \), for all \( s \in S \) and \( o \in O \). The transition relation is defined as \( Tr \subseteq S \times A \times O \times S \rightarrow [0,1] \)

\( p(s, a, o, s') \) can be written as \( s \xrightarrow{a,o,p} s' \). If \( p \) is the transition probability function given above, we can define for all \( s \in S, a \in A \)

\( o \in O, S' \subseteq S \):

\[
p(s, a, o, s^*) = \sum p \quad s \xrightarrow{a,o,p} s', s' \in S'
\]

This can also be called as Modified Probabilistic Transition System (MPTS).

Using MPTS, Stochastic Finite Automata can be extended by adding two components \( O \) and \( U \) and it is a pair \((M, P)\) such that \( M = (Q, A, \delta^*, q_0, O, U, F) \) where \( \delta^* \subseteq Q \times A \times O \times U \times Q \) is a finite set of transitions between states and \( P \) is a function \( \delta^* \rightarrow [0,1] \) which means that probability can be assigned to the occurrence of the outcomes.

Conclusion

It has already been established by the researchers that for every Stochastic Regular Language there exists an equivalent Stochastic Finite Automaton[17]. Hence, it can conveniently be inferred from [16] and [17] that the extended Stochastic Finite Automata can be called as Stochastic Finite Automata for Sequential Decision Making under Uncertainty. From this, it is concluded that the Stochastic Finite Automata is a Mathematical Model for Sequential Decision Making under Uncertainty.

References


