Probability in Epistemic Knowledge

A summary of Fagin and Halpern

Ronan Perry

1 Introduction

Logics for epistemic knowledge have broad applications and are useful in modeling the knowledge states of individuals, such as in the classic muddy children problem. Variants such dynamic epistemic logic extend the classic logic framework to better suit specific applications and situations. In their paper "Reasoning about Knowledge and Probability," Ronald Fagin and Joseph Halpern introduce probability into the language of epistemic logic, allowing agents to not only speak of worlds as possibly existing but as possibly existing with a certain probability. As the authors point out, the ability to reason about knowledge in a probabilistic manner is critical in topics such as distributed computing and fields such as economics and artificial intelligence, which often already do but which don't provide an explicit axiomatic system for such. In this paper, I will summary the key portions of Fagin and Halpern's paper, notably their added semantics and axioms which form the basis of what they call the "logic of knowledge and probability" [1]. Then, I will show how their new logic allows for discussion of the Monty Hall problem, a well known problem in logic and probability.

2 A Standard Kripke Model

I will first review the language of an epistemic Kripke model, $\mathcal{L}_e$, which consists of the following syntax:
$\alpha \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\varphi \vee \psi) \mid K_i\varphi$
Where $K_i\varphi$ denotes that agent i knows that φ and the rest are directly from sentential logic. An epistemic Kripke model for $\mathcal{L}_e$ is a model of knowledge across various agents and possible worlds. Specifically, it is the set $M = (S, \pi, R_1, ..., R_n)$ where S is the set of possible worlds in the model, π is a truth assignment evaluating a sentence at a world $s \in S$ to either true or false, and R_i is a binary relation for each agent i, $i \in [0, n]$. wR_iv if world v is a possibility from world w given what agent i knows. In their paper, Fagin and Halpern assume that R_i is an equivalence relation as they are primarily concerned with distributed systems where such an assumption makes sense. They do note that their results can be easily modified to fit a generic binary relation [1]. For convenience, let $R_i(w)$ denote the set of worlds accessible from w by agent i.

With respect to the logic of epistemic Kripke models, the following set of axioms and inferences, denoted $S5$, is sound and complete [1]:

K1. All instances of propositional tautologies

K2. $(K_i\varphi \wedge K_i(\varphi \Rightarrow \psi)) \Rightarrow K_i\psi$

K3. $K_i\varphi \Rightarrow \varphi$

K4. $K_i\varphi \Rightarrow K_iK_i\varphi$

K5. $\neg K_i\varphi \Rightarrow K_i\neg K_i\varphi$

R1. From φ and $\varphi \Rightarrow \psi$ infer ψ (modus ponens)

R2. From φ infer $K_i\varphi$

3 Adding Probability

Before we can incorporate probability into the epistemic Kripke model, we must first introduce probability theory. A probability space consists of (Ω, X, μ) where Ω is the sample space, X is a set of subsets of Ω, and μ is a probability valuation function which assigns a probability measure to each element of X. So, for this definition to be well behaved, all elements of X must be measurable. However, it may be that we wish to consider subsets of Ω which contain unmeasurable elements. In order to consider all possible subsets of Ω, Fagin and Halpern define the inner measure μ_* such that

$\mu_*(\text{A}) = \sup\{\mu(\text{B}) \mid \text{B} \subseteq \text{A and B} \in X\}$

The inner measure essentially assigns to the subset A of Ω the measure of the largest measurable subset of A, thus allowing us to consider the probabilities of all subsets of Ω [1].

To add the notion of probability to the knowledge states in our model $M = (S, \pi, R_1, ..., R_n)$ with n agents, Fagin and Halpern associate to each world $s \in S$ its own probability space. Given an agent i at world s, we wish to place a probability on the worlds in S from the perspective of that agent. The resulting model they term a "Kripke structure for knowledge and probability" and is the tuple $(S, \pi, R_1, ..., R_n, P)$. The addition P is the probability assignment mapping an agent and world to a probability space such that $P(i, s) = (S_{i,s}, X_{i,s}, \mu_{i,s})$, denoted $P_{i,s}$, where $S_{i,s} \subseteq S$ is a subset of the worlds in our model. Fagin is quick to note that $S_{i,s}$ need not necessarily be equivalent to $R_i(s)$, the set of worlds epistemically open to agent i. This allows one to consider only subsets of the possible worlds. It does need to be a subset of $R_{i,s}$ however, as otherwise it would be possible for an agent to assign a positive probability to something they know to be false. [1].

Fagin and Halpern then go on to express more specifically how a probability is related to an arbitrary sentence φ. It is assumed that we have defined $(M, s) \models \varphi$ for $s \in S$. Let $S_{i,s}(\varphi) = \{s' \in S_{is,} \mid (M, s') \models \varphi\}$, the set of all worlds in $S_{i,s}$ where φ is entailed. Then we can define $w_i(\varphi) \geq b$ as being true according to agent i at a world s if the measure of $S_{i,s}$ is at least b. Fagin's and Halpern's formal definition is:

$$(M, s) \models w_i(\varphi) \geq b \text{ iff } \mu_{i,s}(S_{is,}(\varphi)) \geq b$$

However, it is possible for some elements of $S_{i,s}$ to be unmeasurable and so it is noted that one should use $(\mu_{i,s})_*$ instead of $\mu_{i,s}$ to ensure that the definition is well behaved in such cases. For instance, Fagin and Halpern work through an example where an action occurs or doesn't occur depending on the result of a coin flip and the result of an input bit being either 1 or 0. This example has four possible outcomes. Although the coin has a known probability distribution, both sides being equally likely, the input bit does not have such a known probability distribution. So, although we want to consider these states and their probabilities, we cannot initially assign a probability to the set $\{\{1, heads\}, \{1, tails\}\}$ as that would give us the probability of the input bit being 1 which is unknown. We can however assign a probability of $\frac{1}{2}$ to the set $\{\{1, heads\}, \{0, heads\}\}$, the states where the coin flip results in *heads* [1].

4 Axioms of the Logic

Fagin and Halpern then go on to outline the axioms of their logic of probability and knowledge. In the introduction of Kripke Models, the axioms and inferences K1 through K2, R1, and R2 were introduced. Axiom K1 and rule R1 cover propositional reasoning while K2-K5 and R2 cover reasoning about knowledge. From probability theory, the following set of axioms are added to the logic [1]:

W1. $w_i(\varphi) \geq 0$ (nonnegativity)

W2. $w_i(\mathit{true}) = 1$ (the probability of the event *true* is 1)

W3. $w_i(\varphi \wedge \psi) + w_i(\varphi \wedge \neg\psi) = w_i(\varphi)$ (additivity)

W4. $w_i(\varphi) = w_i(\psi)$ if $\varphi \Leftrightarrow \psi$ is a propositional tautology (distributivity)

W5. $w_i(\mathit{false}) = 0$ (the probability of the event *false* is 0)

Additionally, axioms are added regarding linear inequalities so that the language of this logic can discuss the combination of sentences regarding events and probabilities [1].

I1. $(a_1 w_i(\varphi_1) + ... + a_k w_i(\varphi_k) \geq b) \Leftrightarrow (a_1 w_i(\varphi_1) + ... + a_k w_i(\varphi_k) + 0 w_i(\varphi_{k+1} \geq b)$ (algebra of zero terms)

I2. $(a_1 w_i(\varphi_1) + ... + a_k w_i(\varphi_k) \geq b) \Rightarrow (a_{j_1} w_i(\varphi_{j_1}) + ... + a_{j_k} w_i(\varphi_{j_k}) \geq b)$ if $j_1, ..., j_k$ is a permutation of $1, ..., k$ (permutation of terms)

I3. $(a_1 w_i(\varphi_1) + ... + a_k w_i(\varphi_k) \geq b) \wedge (a'_1 w_i(\varphi_1) + ... + a'_k w_i(\varphi_k) \geq b) \Rightarrow (a_1 + a'_1) w_i(\varphi_1) + ... + (a_k + a'_k) w_i(\varphi_k) \geq b$ (factorization of coefficients)

I4. $(a_1 w_i(\varphi_1) + ... + a_k w_i(\varphi_k) \geq b) \Leftrightarrow (d a_1 w_i(\varphi_1) + ... + d a_k w_i(\varphi_k) \geq db)$ for $d > 0$ (multiplication of positive coefficient)

I5. $(w_i(\varphi) \geq b) \vee (w_i(\varphi) \leq b)$ (dichotomy)

I6. $(w_i(\varphi) \geq b) \Rightarrow (w_i(\varphi) > c)$ if $b > c$ (monotonicity)

In the case that there are non-measurable elements, axiom W3 fails and is instead replaced by a new axiom W6 which is based on the *inclusion-exclusion rule* but which will be omitted as it is technically complicated and not particularly necessary. These axioms comprise a sound and complete axiomatization for the logic of knowledge and probability [1].

Fagin and Halpern go on to consider additional conditions that may be added to capture certain properties of the system and of the knowledge of agents. These conditions are in correspondence with additional axioms. One such desired condition, stated previously but not formalized, is that $S_{i,s} \subseteq R_i(s)$ which guarantees that a positive probability is not assigned to something which agent i knows to be false. This condition is denoted CONS, for *consistent* [1].

CONS. For all i and s, if $P_{i,s} = (S_{i,s}, X_{i,s}, \mu_{i,s})$, then $S_{i,s} \subseteq R_i(s)$

W7. $K_i(\varphi) \Rightarrow (w_i(\varphi) = 1)$

Another possibly desired condition is that all of the agents agree on the probability space at any given point. This is relevant in cases where probability distributions are common knowledge, such as a fair coin is flipped for all to see. This condition is denoted OBJ for *objective* [1].

OBJ. $P_{i,s} = P_{j,s}$ for all i, j and s

W8. $(a_1 w_i(\varphi_1) + ... + a_k w_i(\varphi_k) \geq b) \Rightarrow (a_1 w_j(\varphi_1) + ... + a_k w_j(\varphi_k) \geq b)$

In some cases, an assumption is made that the probability space is the same in all worlds accessible to an agent i. That is, $P_i = (S, X_i, \mu_i)$, where $S = R_i(s)$ is the set of all worlds. As Fagin and Halpern point out, this is a common assumption, especially in economics works. The probability of a specific event is the conditional probability of that event given the set of worlds accessible to an agent i at a state s. This condition is denoted SDP, for *state-determined probability* [1].

SDP. For all i, s, and t, if $t \in R_i(s)$, then $P_{i,s} = P_{i,t}$

W10. $\varphi \Rightarrow K_i(\varphi)$ if φ is an i-probability formula for the negation of an i-probability formula

A less strict version of SDP if UNIF, for *uniformity*, which only requires that $R_i(s)$ can be partitioned into subsets such that the probability space of each world in a given subset is the same [1].

UNIF. For all i, s, and t, if $P_{i,s} = (S_{i,s}, X_{i,s}, \mu_{i,s}$ and $t \in S_{i,s}$, then $P_{i,t} = P_{i,s}$

W9. $\varphi \Rightarrow (w_i(\varphi) = 1)$ if φ is an i-probability formula for the negation of an i-probability formula

The final condition described is that all formulas define measurable sets and is denoted MEAS [1].

MEAS. For all i and s and for every formula φ, the set $S_{i,s}(\varphi) \in X_{i,s}$.

The union of our required axioms with a subset of these additional axioms is sound and complete for model which satisfy the conditions corresponding to those additional axioms [1].

Lastly, Fagin and Halpern incorporate probability into the definition of common knowledge, which is relevant in cases where it is not necessarily possible for true common knowledge to exist but yet there is some threshold of certainty which is good enough, say in a scenario in the Two Generals Problem.

Let G denote the set of all agents $\{1, ..., n\}$. Then $E_G^b\varphi$ says that "everyone knows φ with at least certainty b" and $C_G^b\varphi$ says that "it is common knowledge among everyone that φ with certainty at least b". Fagin's and Halpern's definitions are as follows.

$$(M, s) \models E_G^b\varphi \text{ iff } (M, s) \models K_i(w_i(\varphi) \geq b) \text{ for all } i \in G$$

For common knowledge, a new term must be introduced. $(F_G^b)^0\varphi = \textit{true}$ and $(F_G^b)^{k+1} = E_G^b(\varphi \wedge (F_G^b)^k\varphi)$. Thus,

$$(M, s) \models C_G^b\varphi \text{ iff } (M, s) \models (F_G^b)^k\varphi \text{ for all } k \geq 1$$

With that, Fagin and Halpern conclude their formulation of the logic for probability and knowledge [1].

5 The Monty Hall Problem

The addition of probability theory to epistemic models allows one to discuss problems such as the Monty Hall problem which involve probabilistic components. The Monty Hall problem was a logic and probability puzzle that famously appeared in the "Ask Marilyn" column of PARADE magazine, run by Marilyn vos Savant who briefly appeared in the Guinness Book of World Records for her record high IQ [3]. The question posed was

> "Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the other, goats. You pick door, say #1, and the host, who knows what's behind the doors, opens another door, say #3, which has a goat. He says to you, 'Do you want to pick door #2?' Is it to your advantage to switch your choice of doors?"
>
> Craig F. Whitaker
> Columbia, MD

The name of the problem comes from the name of the host of the TV show "Let's Make a Deal", Monty Hall [2]. Despite the common intuition that there is no advantage to switching, vos Savant answered that it is always advantageous to switch, much to the uproar of many individuals who wrote in to criticize her. However, she was indeed correct [3]. Because one is initially more likely to select a door with a goat behind it, given some assumptions, there is a $\frac{2}{3}$ chance that one will win the car after choosing to switch doors. We shall see this more rigorously below after applying the logic of knowledge and probability.

In this problem, Monty Hall acts not as an agent but rather is simply part of the game, what Fagin and Halpern like to term "nature" [1]. Thus, there is only one agent in consideration and so for convenience I will drop the agent indicator subscript. There are three possible states for the doors to be in. Let C_i denote that the car is behind the *ith* door. Thus, our states are $\{C_1, C_2, C_3\}$. We need to consider the progression of the game through three states and their models: M_0, before the agent has selected a door, M_1, after the agent has selected a door but before Monty Hall opens one, and M_2, where Monty Hall has opened a door.

Consider the first model M_0. Based on the assumptions underlying the problem, the states are equally likely and so $w(C_1) = w(C_2) = w(C_3) = \frac{1}{3}$. After the agent picks a door but before Monty Hall opens a door, our probability model does not change and so in M_1 it is still the case that $w(C_1) = w(C_2) = w(C_3) = \frac{1}{3}$. Note that both M_0 and M_1 both satisfy the SDP condition as for both, the probability space is the same regardless of which world the agent is in. Additionally, these models also naturally satisfy CONS as we don't wish to assign probabilities to false worlds.

Thus, without loss of generality due to uniformity regarding the agent's pick of door, we need only consider one of the three choices. Assume that door 1 was chosen. In the subsequent model M_2, Monty Hall a opens door, where O_i denotes the opening of the *ith* door. Thus, the possible worlds from the perspective of the agent in M_1 consist of combinations of cases of C_i and O_i. For the case C_2, we have that $w(C_2) = \frac{1}{3}$ and because door 1 was chosen, $K(C_2 \Rightarrow O_3)$. So, $w(C_2 \wedge O_3) = \frac{1}{3}$. By similar reasoning, $w(C_3 \wedge O_2) = \frac{1}{3}$. For the case C_1, things are slightly different. $K(C_1 \Rightarrow (O_2 \vee O_3))$ and so $w(C_1 \wedge (O_2 \vee O_3)) = \frac{1}{3}$. Note that, because we don't know the the probability distribution behind the opening of the doors, $C_1 \cap O_2$ is not a measurable set. However, we can say that $w(C_1 \wedge O_2) + w(C_1 \wedge O_3) = \frac{1}{3}$. Thus, because of axiom W1, $0 \leq w(C_1 \wedge O_2) \leq \frac{1}{3}$ and $0 \leq w(C_1 \wedge O_3) \leq \frac{1}{3}$. If a reasonable assumption is made that both cases are equally likely, we have that $w(C_1 \wedge O_2) = w(C_1 \wedge O_3) = \frac{1}{6}$.

Because door 1 was chosen, M_2 can be based on either O_2 or O_3. The agent's view on C_1 being *true* is the conditional probability of C_1 given the set of possible worlds in M_1, as specified by SDP. That is, given an M_2 where O_2 is true, our probability measure says $\mu_{C_1}(C_1) = \mu(C_1 \cap O_2)/\mu(O_2)$. We know that $\mu(O_2) = \mu((O_2 \cap C_1) \cup (O_2 \cap C_3))$ since C_1 and C_3 span the O_2 sample space. Unfortunately, as mentioned previously, $C_1 \cap O_2$ is not a measurable set. However, for our purposes it suffices to know that $\mu(C_1 \cap O_2) \in [0, \frac{1}{3}]$. Thus, $\mu_{C!}(C_1)$ is constrained to the interval $[0, \frac{1}{2}]$ and so in model M_2, after we have picked door 1 and Monty Hall has opened door 2, $w(C_1) \in [0, \frac{1}{2}]$.

In M_2, if O_2 is *true* then C_2 is necessarily *false* and so we are only left to consider C_3. Similar to the above case, the probability measure here says $\mu_{C_3}(C_3) = \mu(C_3 \cap O_2)/\mu(O_2)$. Although we know that in the numerator, $\mu(C_3 \cap O_2) = \frac{1}{3}$, we still face the issue where the denominator is non-measurable. Considering the range of possible values like before, we conclude that $w(C_1) + w(C_3) = 1$ and $w(C_3) \in [\frac{1}{2}, 1]$.

By the same logic, it is easy to show that if O_3 occurs in M_2, then $w(C_1) \in [0, \frac{1}{2}]$ and $w(C_1) + w(C_2) = 1$. Thus, $w(C_1) \leq w(C_2)$ and $w(C_1) \leq w(C_3)$ in the O_3 and O_2 cases respectively. Equality in one forces inequality in the other and can only hold for one of infinitely many possible probability spaces for how Monty Hall chooses the door to open. If we take the common assumption that Monty Hall has no preference when it comes to choosing a door to open, then $w(C_1) = \frac{1}{3} < w(C_2) = \frac{2}{3}$ and $w(C_1) = \frac{1}{3} < w(C_3) = \frac{2}{3}$ in the O_3 and O_2 cases respectively. Since the agent initially chose door 1, it is always to their advantage to switch. Since the logic is symmetric for any initially selected door, it is always advantageous for the agent to switch their choice of door. On a final side note, at M_2, the CONS condition is still satisfied but SDP is no longer as different starting worlds will lead to different probability spaces. As an easy example, if we again select door 1, $C_2 \Rightarrow O_3 \Rightarrow \mu_{C_2}(C_3) = 0$ whereas $C_3 \Rightarrow O_2 \Rightarrow \mu_{C_3}(C_2) = 0$ in M_2.

6 Conclusion

In discussing problems such as the Monty Hall problem, the logic of knowledge and probability developed by Fagin and Halpern allows us to make probabilistic arguments and decisions which wouldn't have been possible in the classic epistemic logic language. Thus, it is a key stepping stone in being able to reason about new problems and develop additional logics. In the analysis of the Monty Hall problem, various models were considered as the game progressed in dynamic sense but the dynamics were never formalized. Fagin's and Halpern's work allows for future research in developing formal dynamic and probabilistic epistemic logics. Such a topic would be strongly connected to ideas of Bayesian inference. In Bayesian inference, we have some unknown truth but also some ideas of what said truth could be, ie. possible worlds. By observing events or data, we are able to formulate a prior distribution, the probability of a possible world given the observations. This in turn gives us an idea as to what the most likely worlds are. The axiomatization of this logic for probability and knowledge seems well suited to future research in areas such as statistics, decision theory, and game theory.

References

[1] R. Fagin and J. Y. Halpern, Reasoning About Knowledge and Probability, *Journal of the ACM*, 41(2):340-367, 1994.

[2] Selvin, Steve. A problem in probability (letter to the editor), *American Statistician*, 29(1):67, 1975.

[3] vos Savant, Marilyn. "Ask Marilyn", Parade Magazine: 16, 1990