1. INTRODUCTION TO PROBABILISTIC MODELS

RENATO LUI GEH
NUSP: 8536030

ABSTRACT. In this document we present the fundamentals of probabilistic models. We show the origins of reasoning through generalized propositional logic and the reasons as to why probabilities have shown themselves so sucessful in representing uncertainty. We then introduce probability calculus and how to perform simple queries by enumeration of values. Important concepts of probability, such as belief updating, independence, random variables and conditional probability are formally presented.

1. REASONING AND UNCERTAINTY IN ARTIFICIAL INTELLIGENCE

Before the advent of probability theory on uncertainty, it was believed that it was only possible to represent knowledge through the use of logic and other neo-calculist techniques. Although logic does have its praises, such as its computational ease and natural semantics that resemble human reasoning, it is problematic when one tackles the issue of uncertainty. When it comes to representing the world, propositional and first order logic must assume a binary reading of data: true or false. Indeed the world is not so simple, and this vision is, at best, simplistic, since we find that degrees of beliefs are present everywhere. For example, one can not say that all birds fly, for it may have its wing broken. Of course we could enumerate all instances where the given bird cannot fly (e.g. broken wing, bird species, etc), but this approach is exhausting and unrealistic. This issue can be ameliorated (if not solved) by the introduction of probabilities as a means to measure degrees of belief. We could then say that the bird has a 90% chance of flying, discarding the remaining 10% as a sum of all probabilities of an infinite set of improbable events that could prevent the bird from flying without having to enumerate each one of them. However the introduction of probability as a direct method of computing uncertainty was viewed with skepticism among AI researchers. Although not so welcomed at first, probability theory gained incredible support in the last decades, with influential researchers such as Judea Pearl actively supporting its use [Pea88]. In this report we first take a look at the formalisms of propositional logic when representing knowledge, analysing the semantics, properties and relationships. Once we are familiarized with the difficulties logic presents to uncertainty we introduce the basics of probability theory and how to compute simple queries.

2. PROPOSITIONAL LOGIC

Be it through probabilities to represent degrees of belief or through Boolean values to represent truths of events, one needs to define a language to describe

knowledge formally. When it comes to representing truths, an appropriate language is propositional logic. The atomic element in propositional logic is the variable. A variable X takes values in a domain Ω. Such value assignment is represented as the sentence $X = x$, for $x \in \Omega$. A sentence is defined as follows.

- A variable X is a sentence.
- If α and β are sentences, then $\alpha \wedge \beta$, $\alpha \vee \beta$ and $\neg \alpha$ are sentences.

The sentence $\alpha \wedge \beta$ represents disjointion, $\alpha \vee \beta$ is called conjunction and $\neg \alpha$ negation. The symbols $\wedge$, $\vee$ and $\neg$ are called logical connectives and are the main constructs of propositional logic. Valuations to a variable X to a value $x \in \Omega$ are described via a function ν. We say that $\nu \models \phi$ if ν satisfies the sentence ϕ. That is, the valuation of ϕ belongs to the domain Ω that ν maps. If ν does not satisfy ϕ, we say that $\nu \not\models \phi$. This means ϕ is not in the domain and therefore $\nu \models \neg \phi$. For conjunction, we say that $\nu \models \phi \vee \psi \iff \nu \models \phi$ or $\nu \models \psi$. Similarly, $\nu \models \phi \wedge \psi \iff \nu \models \phi$ and $\nu \models \psi$. By guaranteeing that negation, conjunction and disjointion obey these rules, one defines a field of sets. A field of sets is a pair $(\Omega, \mathcal{F})$ where Ω is the set of valuations and $\mathcal{F}$ is the set of possible outcomes closed under negation, union (i.e. conjunction) and intersection (i.e. disjointion).

Proposition 1. *(Exercise 7)*
Let Ω be the set of all valuations and $\mathcal{F}$ be the set of events.

$$\alpha := \{\nu \in \Omega : \nu \models \phi\}, \text{ when } \phi \text{ is a sentence.}$$

Then $(\Omega, \mathcal{F})$ is a field of sets.

Proof. To prove $(\Omega, \mathcal{F})$ is a field of sets we must show that $\mathcal{F}$ is a non-empty set closed under union, intersection and complement. Let $\phi, \psi \in \Omega$, $\nu \models (\phi \vee \psi) \iff \nu \models \phi$ or $\nu \models \psi$. But from our earlier assumption, $\phi, \psi \in \Omega$; therefore $\mathcal{F}$ is closed under union. Similarly, $\nu \models (\phi \wedge \psi) \iff \nu \models \phi$ and $\nu \models \psi$. Since both sentences $\phi, \psi \in \Omega$ then $\mathcal{F}$ is closed under intersection. Finally, for an arbitrary ϕ, $\nu \models \neg \phi \iff \nu \not\models \phi$. Assume $\phi \in \Omega$. Then ϕ satisfies in Ω and therefore $\neg \phi$ does not satisfy lest we create a contradiction. Thus, by the definition of satisfiabilitiy $\nu \models \neg \phi$. Now consider $\nu \not\models \neg \phi$, then $\nu \models \neg(\neg \phi)$ and so $\neg \phi$ is in Ω. Therefore, $\mathcal{F}$ is closed under complements. $\square$

3. Probability Calculus

Whilst in propositional logic we had a binary set of possible values to variables, in probability theory we have a contiguous interval of possible valuations. These possible values represent a certain degree of belief that an agent may consider for an event. That is, for an event α to occur in the set of possible outcomes ω, we may define the degree of belief, or probability, as a function $\mathbb{P}$ such that:

$$\mathbb{P}(\alpha) \stackrel{\text{def}}{=} \sum_{\omega \models \alpha} \mathbb{P}(\omega)$$

In other words, the probability of event α to occur is the summation of every worlds w_i in which α is satisfiable. Of course one must guarantee that $\mathbb{P}$ follows the axioms of probability.

Definition. *A function $\mathbb{P}$ is a probability function if, for every $\alpha \in \Omega$ where Ω is the set of valuations in the field of sets $(\Omega, \mathcal{F})$, $\mathbb{P}$ obeys the following probability axioms:*

(1) $\mathbb{P}(\alpha) \in \mathbb{R}, \mathbb{P}(\alpha) \leq 0$
(2) $\mathbb{P}(\Omega) = 1$
(3) $\mathbb{P}(\bigcup_i^n \alpha_i) = \sum_i^n \mathbb{P}(\alpha)$ if $\bigcap_i^n \alpha_i = \emptyset$, that is, α_i are disjoint sets.

From these axioms we derive the following properties.

Proposition 2. *(Exercise 1.1)*
For any event α, it follows that $\mathbb{P}(\alpha) = 1 - \mathbb{P}(\alpha^c)$, where α^c is the complement of α.

Proof. Let Ω be the space of outcomes in which α is contained in. Since α and α^c are disjoint and exhaustive then from the axiom $\mathbb{P}(\Omega) = 1$, we have that $\mathbb{P}(\alpha) + \mathbb{P}(\alpha^c) = \mathbb{P}(\Omega) = 1$. From there we have $\mathbb{P}(\alpha) = 1 - \mathbb{P}(\alpha^c)$. $\square$

Proposition 3. *(Exercise 1.2)*
$\mathbb{P}(\emptyset) = 0$

Proof. Assume that $\mathbb{P}(\emptyset) > 0$. From the probability axioms we have that $\mathbb{P}(\alpha \cup \beta) = \mathbb{P}(\alpha) + \mathbb{P}(\beta)$ if α and β are disjoint. Then $\mathbb{P}(\emptyset \cup \emptyset) = \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset)$, since an empty set is disjoint with itself. But $\emptyset \cup \emptyset = \emptyset$, therefore $\mathbb{P}(\emptyset) = 2\mathbb{P}(\emptyset)$. Which is a contradiction if $\mathbb{P}(\emptyset) > 0$. Thus $\mathbb{P}(\emptyset) = 0$. $\square$

Proposition 4. *(Exercise 1.3)*
If $\alpha \subseteq \beta$ are events then $\mathbb{P}(\alpha) \leq \mathbb{P}(\beta)$. This is called monotonicity.

Proof. Assume that $\mathbb{P}(\alpha) > \mathbb{P}(\beta)$ when $\alpha \subseteq \beta$. Now consider $\alpha = \emptyset$: $\mathbb{P}(\emptyset) > \mathbb{P}(\beta)$. But $\mathbb{P}(\emptyset) = 0 \Rightarrow 0 > \mathbb{P}(\beta)$, and from the probability axiom, $\mathbb{P}(\beta) \geq 0$ which is a contradiction. Therefore $\mathbb{P}(\alpha) \geq \mathbb{P}(\beta)$. $\square$

Proposition 5. *(Exercise 1.4)*
For every event α, we have that $0 \leq \mathbb{P}(\alpha) \leq 1$.

Proof. From the first probability axiom we have that $\mathbb{P}(\alpha) \geq 0$. Consider $\mathbb{P}(\alpha) > 1$. From $\mathbb{P}(\alpha) = 1 - \mathbb{P}(\alpha^c)$ we have that $\mathbb{P}(\alpha^c) < 0$ which contradicts the first probability axiom. Thus $0 \leq \mathbb{P}(\alpha) \leq 1$. $\square$

Proposition 6. *(Exercise 1.5)*
For any events α and β (not necessarily disjoint), we have that $\mathbb{P}(\alpha \cup \beta) = \mathbb{P}(\alpha) + \mathbb{P}(\beta) + \mathbb{P}(\alpha \cap \beta)$.

Proof. From the third axiom, $\mathbb{P}(\beta) = \mathbb{P}(\alpha \cap \beta) + \mathbb{P}(\beta \setminus (\alpha \cap \beta))$, where $\alpha \cap \beta$ and $\beta \setminus (\alpha \cap \beta)$ are disjoint. The third axiom of probability, $\mathbb{P}(\bigcup_i X_i) = \sum_i \mathbb{P}(X_i)$, is for disjoint sets. Therefore: $\mathbb{P}(\alpha \cup (\beta \setminus (\alpha \cap \beta))) = \mathbb{P}(\alpha) + \mathbb{P}(\beta \setminus (\alpha \cap \beta))$. Substituting $\mathbb{P}(\beta \setminus (\alpha \cap \beta))$ with $\mathbb{P}(\beta) - \mathbb{P}(\alpha \cap \beta)$, $\mathbb{P}(\alpha \cup (\beta \setminus (\alpha \cap \beta))) = \mathbb{P}(\alpha) + \mathbb{P}(\beta) - \mathbb{P}(\alpha \cap \beta)$. But $\alpha \cup (\beta \setminus (\alpha \cap \beta)) = \alpha \cup \beta$, resulting in $\mathbb{P}(\alpha \cup \beta) = \mathbb{P}(\alpha) + \mathbb{P}(\beta) - \mathbb{P}(\alpha \cap \beta)$. □

Proposition 7. *(Exercise 1.6)*
If $\mathbb{P}(\alpha) = 1$ for some event α then $\mathbb{P}(\beta) = \mathbb{P}(\alpha \cap \beta)$ for any event β.

Proof. $\mathbb{P}(\alpha \cup \beta) = \mathbb{P}(\alpha) + \mathbb{P}(\beta) - \mathbb{P}(\alpha \cap \beta)$. Since $\mathbb{P}(\alpha) = 1$ and if α happens with certainty then $\mathbb{P}(\alpha \cup \beta) = \mathbb{P}(\Omega) = 1$, then $1 = 1 + \mathbb{P}(\beta) - \mathbb{P}(\alpha \cap \beta) \Rightarrow \mathbb{P}(\beta) = \mathbb{P}(\alpha \cap \beta)$. □

3.1. *Conditional probability*

Belief updating in probability is done through conditional probabilities. Intuitively, one could assume belief updating as a change in belief given new evidence. For example, we could assume that the probability of a train arriving on time to our station is $\mathbb{P}(OnTime = true) = 0.3$. But say we receive news from our neighboring station that the train has arrived on time to their station. We could safely assume that this fact increases our belief that the train will arrive on time, since we received news it has not yet delayed. That is, given the evidence that the train has not delayed on the i-th station, our probability of our train arriving on time is $\mathbb{P}(OnTime = true | StationDelayed_i = false) = 0.8$. A probability function of the form $\mathbb{P}(X|E)$ is called a conditional or posterior probability. The variable X is called the query variable and E the evidence.

Definition. *A conditional probability $\mathbb{P}(\alpha|\beta)$ is defined as:*

$$\mathbb{P}(\alpha|\beta) = \frac{\mathbb{P}(\alpha \cap \beta)}{\mathbb{P}(\beta)}, \mathbb{P}(\beta) > 0$$

Just as in probabilities of the form $\mathbb{P}(X)$, which are called prior probabilities, conditional probabilities obey the three probability axioms:

(1) $\mathbb{P}(\alpha|\beta) \geq 0$
(2) $\mathbb{P}(\Omega|\beta) = 1$
(3) $\mathbb{P}(\alpha \cup \beta|\gamma) = \mathbb{P}(\alpha|\gamma) + \mathbb{P}(\beta|\gamma)$ for α and β disjoint.

Proposition 8. *(Exercise 2)*
Let $\mathbb{P}$ be a probability function (or some probability space). Show that for any event β with $\mathbb{P}(\beta) > 0$ the function $\mathbb{P}'(\alpha) = \mathbb{P}(\alpha|\beta)$ is also a probability fuction.

Proof. From the definition of conditional probability:

$$\mathbb{P}(\alpha|\beta) = \frac{\mathbb{P}(\alpha \cap \beta)}{\mathbb{P}(\beta)}$$

Since $\mathbb{P}(\beta) > 0$, $\mathbb{P}(\alpha \cap \beta)/\mathbb{P}(\beta)$ is defined for all possible values and $\mathbb{P}(\alpha|\beta) \geq 0$ (first axiom), because $\mathbb{P}(\alpha \cap \beta) \geq 0$ and $\mathbb{P}(\beta) > 0$.

$$\mathbb{P}(\alpha|\beta) = \frac{\mathbb{P}(\alpha) + \mathbb{P}(\beta) - \mathbb{P}(\alpha \cup \beta)}{\mathbb{P}(\beta)} \leq 1$$

Because $\mathbb{P}(\alpha) + \mathbb{P}(\beta) - \mathbb{P}(\alpha \cup \beta) \leq \mathbb{P}(\beta) \Rightarrow \mathbb{P}(\alpha) \leq \mathbb{P}(\alpha \cup \beta)$ which is true since if $X \subseteq Y$ then $\mathbb{P}(X) \leq \mathbb{P}(Y)$. Substituting X for α and Y for $\alpha \cup \beta$ we have that $\mathbb{P}(\alpha) \leq \mathbb{P}(\alpha \cup \beta)$. Thus $\mathbb{P}(\alpha|\beta) \leq 1$ (second axiom). For the third axiom suppose α and γ are disjoint.

$$\mathbb{P}(\alpha \cup \gamma|\beta) = \frac{\mathbb{P}((\alpha \cup \gamma) \cap \beta)}{\mathbb{P}(\beta)} = \frac{\mathbb{P}((\alpha \cap \beta) \cup (\gamma \cap \beta))}{\mathbb{P}(\beta)}$$

$\alpha \cap \beta$ and $\gamma \cap \beta$ are disjoint since α and γ are disjoint. Therefore we may apply axiom three:

$$\mathbb{P}(\alpha \cap \gamma|\beta) = \frac{\mathbb{P}(\alpha \cap \beta) + \mathbb{P}(\gamma \cap \beta)}{\mathbb{P}(\beta)} = \frac{\mathbb{P}(\alpha \cap \beta)}{\mathbb{P}(\beta)} = \frac{\mathbb{P}(\gamma \cap \beta)}{\mathbb{P}(\beta)} = \mathbb{P}(\alpha|\beta) + \mathbb{P}(\gamma|\beta)$$

$\square$

3.2. *Chain Rule*

Proposition 9. *(Exercise 3)*
For any sequence of events $\alpha_1, \ldots, \alpha_n$ we have that:
$$\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_n) = \mathbb{P}(\alpha_1) \prod_{i=2}^{n} \mathbb{P}(\alpha_i|\alpha_1 \cap \ldots \cap \alpha_{i-1})$$

Proof. We will prove by induction. For the base $n = 2$ we have, directly from the definition of conditional probability:

$$\mathbb{P}(\alpha_2|\alpha_1) = \frac{\mathbb{P}(\alpha_1 \cap \alpha_2)}{\mathbb{P}(\alpha_1)} \Rightarrow \mathbb{P}(\alpha_1 \cap \alpha_2) = \mathbb{P}(\alpha_1)\mathbb{P}(\alpha_2|\alpha_1)$$

For the k-th case:

$$\mathbb{P}(\alpha_k|\alpha_1 \cap \ldots \cap \alpha_{k-1}) = \frac{\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_k)}{\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k-1})}$$

(i) $\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_k) = \mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k-1})\mathbb{P}(\alpha_k|\alpha_1 \cap \ldots \cap \alpha_{k-1})$

But $\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k-1})$ can be transformed into

(ii) $\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k-1}) = \mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k-2})\mathbb{P}(\alpha_{k-1}|\alpha_1 \cap \ldots \cap \alpha_{k-2})$

Applying (ii) in (i) we have:

$$\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_k) = \mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k-2})\mathbb{P}(\alpha_{k-1}|\alpha_1 \cap \ldots \cap \alpha_{k-2})\mathbb{P}(\alpha_k|\alpha_1 \cap \ldots \cap \alpha_{k-1})$$

For the $(k+1)$-th case:

$$\mathbb{P}(\alpha_{k+1}|\alpha_1 \cap \ldots \cap \alpha_k) = \frac{\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k+1})}{\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_k)}$$

$$\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_{k+1}) = \mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_k)\mathbb{P}(\alpha_{k+1}|\alpha_1 \cap \ldots \cap \alpha_k)$$

But by induction hypothesis $\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_k)$ is true. Therefore for $n \geq 2$ we have:

$$\mathbb{P}(\alpha_1 \cap \ldots \cap \alpha_n) = \mathbb{P}(\alpha_1)\mathbb{P}(\alpha_2|\alpha_1) \ldots \mathbb{P}(\alpha_n|\alpha_1 \cap \ldots \cap \alpha_{n-1})$$

$$= \mathbb{P}(\alpha_1) \prod_{i=2}^{n} \mathbb{P}(\alpha_i|\alpha_1 \cap \ldots \cap \alpha_{i-1})$$

$\square$

3.3. *Total Probability Rule*

Proposition 10. *(Exercise 4)*
Let $\alpha_1, \ldots, \alpha_n$ be a partition of Ω (i.e. $\bigcup_i \alpha_i = \Omega$ and $\alpha_i \cap \alpha_j = \emptyset$ for $i \neq j$, $\alpha_i \neq \emptyset$). Then for any event β it follows that:

$$\mathbb{P}(\beta) = \sum_{i=1}^{n} \mathbb{P}(\beta|\alpha_i)\mathbb{P}(\alpha_i)$$

Proof. Since $\alpha_1, \ldots, \alpha_n$ are disjoint, then:

(i) $\mathbb{P}(\beta \cap (\bigcup_i \alpha_i)) = \mathbb{P}((\beta \cap \alpha_1) \cup \ldots \cup (\beta \cap \alpha_n))$

From the definition of conditional probability, for an arbitrary i we have:

(ii) $\mathbb{P}(\beta|\alpha_i) = \dfrac{\mathbb{P}(\beta \cap \alpha_i)}{\mathbb{P}(\alpha_i)} \Rightarrow \mathbb{P}(\beta \cap \alpha_i) = \mathbb{P}(\beta|\alpha_i)\mathbb{P}(\alpha_i)$

Applying (ii) in (i) we have:

$$\mathbb{P}(\beta \cap (\bigcup_i \alpha_i)) = \mathbb{P}(\beta|\alpha_1)\mathbb{P}(\alpha_1) + \cdots + \mathbb{P}(\beta|\alpha_n)\mathbb{P}(\alpha_n)$$

Since $\alpha_1, \ldots, \alpha_n$ are disjoint and exhaustive in Ω, then $\beta \cap (\alpha_1 \cup \ldots \cup \alpha_n) = \beta \cap \Omega = \beta$. Therefore:

$$\mathbb{P}(\beta) = \sum_i \mathbb{P}(\beta|\alpha_i)\mathbb{P}(\alpha_i)$$

$\square$

3.4. *Baye's Rule*

Sometimes computing the posterior probability through the joint distribution is expensive if not impossible. A solution to this problem is Baye's Rule.

$$\mathbb{P}(\alpha|\beta) = \frac{\mathbb{P}(\beta|\alpha)\mathbb{P}(\alpha)}{\mathbb{P}(\beta)}$$

This expression facilitates finding $\mathbb{P}(\alpha|\beta)$ when one already has $\mathbb{P}(\beta|\alpha)$ computed. A typical example to this is on medical applications. Consider a disease α and symptoms β. The diagnosis of a disease given the symptoms ($\mathbb{P}(\alpha|\beta)$) the patient may be experiencing can be easily computed when one knows the symptoms the disease causes $\mathbb{P}(\beta|\alpha)$. Baye's rule can be seen as the update of a probability after one computes the "future probability" one would get.

Proposition 11. *(Exercise 5)*
Extended Baye's Rule. For any events α, β and γ with positive probability, it follows that:

$$\mathbb{P}(\beta|\alpha \cap \gamma) = \frac{\mathbb{P}(\alpha|\beta \cap \gamma)\mathbb{P}(\beta|\gamma)}{\mathbb{P}(\alpha|\gamma)}$$

Proof. Applying the definition of conditional probability on $\mathbb{P}(\alpha \cap \beta \cap \gamma)$ we have:

$$\mathbb{P}(\beta|\alpha \cap \gamma) = \frac{\mathbb{P}(\alpha \cap \beta \cap \gamma)}{\mathbb{P}(\alpha \cap \gamma)} = \frac{\mathbb{P}(\alpha \cap \beta \cap \gamma)}{\mathbb{P}(\alpha|\gamma)\mathbb{P}(\gamma)}$$

After we decompose $\mathbb{P}(\alpha \cap \gamma)$ we now once again apply conditional probability:

$$\mathbb{P}(\beta|\alpha \cap \gamma) = \frac{\mathbb{P}(\alpha \cap \beta \cap \gamma)}{\mathbb{P}(\alpha|\gamma)\mathbb{P}(\gamma)} = \frac{\mathbb{P}(\alpha|\beta \cap \gamma)\mathbb{P}(\beta \cap \gamma)}{\mathbb{P}(\alpha|\gamma)\mathbb{P}(\gamma)}$$

Once again we decompose $\mathbb{P}(\beta \cap \gamma)$:

$$\mathbb{P}(\beta|\alpha \cap \gamma) = \frac{\mathbb{P}(\alpha|\beta \cap \gamma)\mathbb{P}(\beta \cap \gamma)}{\mathbb{P}(\alpha|\gamma)\mathbb{P}(\gamma)} = \frac{\mathbb{P}(\alpha|\beta \cap \gamma)\mathbb{P}(\beta|\gamma)\mathbb{P}(\gamma)}{\mathbb{P}(\alpha|\gamma)\mathbb{P}(\gamma)}$$

Since we assume $\mathbb{P}(\alpha), \mathbb{P}(\beta), \mathbb{P}(\gamma) > 0$, then:

$$\mathbb{P}(\beta|\alpha \cap \gamma) = \frac{\mathbb{P}(\alpha|\beta \cap \gamma)\mathbb{P}(\beta|\gamma)}{\mathbb{P}(\alpha|\gamma)}$$

$\square$

3.5. *Independence*

Two events are considered independent of each other when the outcome of the first does not impact the outcome of the second. An example to this is the toss of two coins. Ideally, the first toss should not change the probability of the second. However, when it comes to modelling the real world one finds that true independency between events is unrealistic. Be that as it may, it is still possible for an event to gain independency in relation to a second event when faced with a third. For instance, consider

References

[Pea88]   Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, 1988.