Examples of Independence for Imprecise Probabilities

Inés Couso
Dpto. Estadística e I.O. y D.M.
Universidad de Oviedo
33001 - Oviedo - Spain
couso@pinon.ccu.uniovi.es

Serafín Moral
Dpto. Ciencias de la Computación
Universidad de Granada
18071 - Granada - Spain
smc@decsai.ugr.es

Peter Walley
36 Bloomfield Terrace
Lower Hutt
New Zealand

Abstract

In this paper we try to clarify the notion of independence for imprecise probabilities. Our main point is that there are several possible definitions of independence which are applicable in different types of situation. With this aim, simple examples are given in order to clarify the meaning of the different concepts of independence and the relationships between them.

Keywords. Imprecise probabilities, independence, conditioning, convex sets of probabilities.

1 Introduction

One of the key concepts in probability theory is the notion of independence. Using independence, we can decompose a complex problem into simpler components and build a global model from smaller submodels [1, 8].

We use the term *stochastic independence* to refer to the standard concept of independence in probability theory, which is usually defined as factorization of the joint probability distribution as a product of the marginal distributions.

The concept of independence is essential for imprecise probabilities too, but there is disagreement about how to define it. Comparisons of different definitions have been given by Campos and Moral [3] and Walley [12]. In this paper we aim to show that several different definitions of independence are needed in different kinds of problems. We will try to demonstrate that through simple examples which involve only two binary variables, where each variable represents the colour of a ball to be drawn from an urn. Each of the examples gives rise to a different mathematical definition of independence. We concentrate on the intuitive meaning of the definitions, making clear the assumptions under which each definition is appropriate. When possible, we give a behavioural interpretation of the definition.

Conditional independence is another fundamental concept for modeling uncertainty, but the possible definitions are even more numerous than for unconditional independence and they will not be considered here.

2 Fundamental Ideas of Imprecise Probability

In this section we give a brief introduction to imprecise probabilities, following Walley [12]. Imprecise probabilities are models for behaviour under uncertainty that do not assume a unique underlying probability distribution but correspond, in general, to a set of probability distributions. A decision maker is not required to choose between every pair of alternatives and has the option of suspending judgement.

Let Ω be a finite set of possibilities, exactly one of which must be true. A *gamble*, X, on Ω is a function from Ω to $\mathbb{R}$ (the set of real numbers). If you were to accept gamble X and ω turned out to be true then you would gain $X(\omega)$ utiles (so you would lose if $X(\omega) < 0$). A subject's beliefs are elicited by asking her to specify a *set of acceptable (*or *desirable) gambles*, i.e., gambles she is willing to accept.

The set of all gambles on Ω is denoted by $\mathcal{L}(\Omega)$. Addition and subtraction of gambles are defined pointwise, so that for gambles X and Y, for each $\omega \in \Omega$, $(X+Y)(\omega) = X(\omega) + Y(\omega)$.

There are three rules for obtaining new acceptable gambles from previous judgements of acceptability [12, 14, 7]:

R1. If min $X \geq 0$, then X is acceptable.

R2. If X is acceptable and $\lambda \geq 0$, then λX is acceptable.

R3. If X_1 and X_2 are acceptable, then $X_1 + X_2$ is acceptable.

Given a set of acceptable gambles Γ, the *closure* of Γ, denoted by $\overline{\Gamma}$, is the set of all gambles that can be obtained from gambles in Γ by applying the rules R1-R3. Closed sets of acceptable gambles correspond to closed convex sets of probability distributions. If $\mathcal{P}(\Omega)$ is the set of all probability distributions on Ω, then the closed convex set of probability distributions associated with Γ is given by $\mathcal{K}_\Gamma = \{P \in \mathcal{P}(\Omega) : E_P(X) \geq 0, \forall X \in \Gamma\}$, where E_P denotes expectation with respect to P. Conversely, from a non-empty set of probability distributions, $\mathcal{K}$, we can define a closed set of acceptable gambles by $\Gamma_\mathcal{K} = \{X \in \mathcal{L}(\Omega) : E_P(X) \geq 0, \forall P \in \mathcal{K}\}$. Sets of acceptable gambles and sets of probability distributions are dual ways of representing the same behaviour.

We shall consider only closed sets of acceptable gambles that can be generated as the closure of a *finite* set of gambles. Such sets of gambles are equivalent to closed and convex sets of probability distributions with a finite number of extreme points.

In this approach, *convexity* of a set of probability distributions has no behavioural significance. From a behavioural point of view, any non-empty set of probability distributions, $\mathcal{K}$, is indistinguishable from its convex hull, CH($\mathcal{K}$), because both $\mathcal{K}$ and CH($\mathcal{K}$) generate the same set of acceptable gambles $\Gamma_\mathcal{K}$. (But see Example 7 for one exception.) In this paper, to obtain a unique set of probability distributions which represents Γ, we will generally use the *largest* such set, which is the convex set $\mathcal{K}_\Gamma$ defined above. However, we could equally well have used a non-convex set such as the set of extreme points of CH($\mathcal{K}_\Gamma$), which is behaviourally equivalent to CH($\mathcal{K}_\Gamma$).

Another equivalent model involves *upper and lower previsions*. Given a closed set of gambles Γ, we can associate with each gamble X its lower prevision, $\underline{P}(X) = \max\{\mu \in I\!\!R : X - \mu \in \Gamma\}$. The upper prevision can be defined analogously, or through the conjugacy relationship $\overline{P}(X) = -\underline{P}(-X)$.

Upper and lower previsions are generalizations of upper and lower probabilities. When the gamble X is the indicator function of an event A, where $A \subseteq \Omega$, the upper and lower previsions $\overline{P}(X)$ and $\underline{P}(X)$ can be regarded as the upper and lower probabilities of A. By using the same notation A to refer to both the event and its indicator function, we can write these upper and lower probabilities as $\overline{P}(A)$ and $\underline{P}(A)$. Upper and lower probabilities, or more generally upper and lower previsions, can also be written as upper and lower envelopes of the closed convex set of probability distributions $\mathcal{K}$, by $\overline{P}(A) = \max\{P(A) : P \in \mathcal{K}\}$ and $\underline{P}(A) = \min\{P(A) : P \in \mathcal{K}\}$.

From a finite set of lower previsions $\underline{P}(X_i) = \mu_i$ ($i = 1, \ldots, n$), we can define lower previsions for every gamble $X \in \mathcal{L}(\Omega)$ by using the technique of *natural extension* [12]. This is equivalent to considering the set of gambles $\Gamma_1 = \{X_i - \mu_i : i = 1, \ldots, n\}$, forming its closure $\overline{\Gamma}_1$, and then calculating lower previsions from $\overline{\Gamma}_1$.

3 Definitions of Independence

Consider two variable or uncertain values which may be regarded as the outcomes of two experiments. Suppose that the two outcomes are known to belong to the possibility spaces Ω_1 and Ω_2, which are assumed to be finite. The most basic condition of independence of the two experiments is that the set of possible joint outcomes is the Cartesian product $\Omega = \Omega_1 \times \Omega_2$, which is called *logical independence*. We assume throughout that this holds.

Let $\mathcal{K}_1$ and $\mathcal{K}_2$ denote the marginal convex sets of probability distributions on Ω_1 and Ω_2 respectively, which model our uncertainty about the two experiments separately. Let $\mathcal{K}$ denote the convex set of joint probability distributions on $\Omega = \Omega_1 \times \Omega_2$, which models our uncertainty about the joint experiment.

We say that a joint probability distribution P on Ω satisfies *stochastic independence* when $P(\{(\omega_1, \omega_2)\}) = P_1(\{\omega_1\})P_2(\{\omega_2\})$ for all $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$, where P_1 is the marginal probability distribution of P on Ω_1 defined by $P_1(\{\omega_1\}) = P(\{\omega_1\} \times \Omega_2)$, and similarly P_2 is the marginal distribution of P on Ω_2. In this case we write $P = P_1 \times P_2$. This is the usual definition of independence in probability theory. We assume that all marginal probabilities $P_1(\{\omega_1\})$ and $P_2(\{\omega_2\})$ are non-zero when $P_1 \in \mathcal{K}_1$ and $P_2 \in \mathcal{K}_2$, to avoid the difficulties that arise from conditioning on events of probability zero. In this case, the usual definition is equivalent to a second definition which requires equality of conditional and marginal distributions. The second definition is preferable to the first one in the case not considered here, when some marginal probabilities may be zero [12].

In this section we discuss several possible definitions of independence that can be used when the marginal probabilities are imprecise, i.e., when the marginal sets $\mathcal{K}_1$ and $\mathcal{K}_2$ are non-degenerate. Most of the definitions of independence are associated with a particular basic property of the joint set $\mathcal{K}$. In Subsection 3.3, for example, the basic property is that our uncertainty about Ω_2 does not change when we learn the value of ω_1. Given a joint set $\mathcal{K}$, we can verify whether $\mathcal{K}$ satisfies the particular basic property of independence.

However, it is not usual in practice that we are given a joint set $\mathcal{K}$. Instead, we usually need to construct the joint set from given marginal sets $\mathcal{K}_1$ and $\mathcal{K}_2$ and a judgement of independence. In general, there will be several joint sets $\mathcal{K}$ which have the same marginals $\mathcal{K}_1$ and $\mathcal{K}_2$ and satisfy a particular basic property of independence. In that case we define a unique model $\mathcal{K}$ to be the *largest* (i.e., the least informative) set of joint probability distributions which has the given marginals $\mathcal{K}_1$ and $\mathcal{K}_2$ and also satisfies the basic property. In this way, each of the basic properties can be used constructively, to determine a unique model for the joint experiment from given marginals.

Each of the following subsections 3.1–3.6 defines a different method for constructing the joint set $\mathcal{K}$ from given marginals $\mathcal{K}_1$ and $\mathcal{K}_2$. When the method is based on a particular basic property of independence, the basic property is defined in the same subsection. The definitions of independence are presented in order of increasing precision; that is, for fixed marginal sets $\mathcal{K}_1$ and $\mathcal{K}_2$, the earlier definitions produce a larger set of joint distributions, $\mathcal{K}$, than the later definitions.

3.1 Independence of the Marginal Sets and Unknown Interaction

We say that there is *independence of the marginal sets* if, for any two marginal distributions $P_1 \in \mathcal{K}_1$ and $P_2 \in \mathcal{K}_2$, there is a joint distribution P in $\mathcal{K}$ which has P_1 and P_2 as its marginals. In other words, learning the marginal probability distribution on Ω_1 would not change our uncertainty about the marginal distribution on Ω_2.

Let $\mathcal{K}_i^*$ denote the set of all joint probability distributions on Ω whose marginal distribution on Ω_i belongs to $\mathcal{K}_i$. Then the largest set of distributions that satisfies independence of the marginal sets is $\mathcal{K} = \mathcal{K}_1^* \cap \mathcal{K}_2^*$. In other words, the joint set $\mathcal{K}$ consists of all joint probability distributions whose two marginal distributions belong to $\mathcal{K}_1$ and $\mathcal{K}_2$ respectively. In this case we say that there is *unknown interaction* between the marginal experiments.

The definition of unknown interaction can also be expressed in terms of the closed sets of gambles Γ_1, Γ_2 and Γ that are determined by $\mathcal{K}_1, \mathcal{K}_2$ and $\mathcal{K}$. Given a closed set of gambles Γ on Ω, there is unknown interaction between the marginal experiments if and only if there are two sets of gambles, $\Gamma_1 \subseteq \mathcal{L}(\Omega_1)$ and $\Gamma_2 \subseteq \mathcal{L}(\Omega_2)$, such that $\Gamma = \overline{(\Gamma_1 \cup \Gamma_2)}$. Here we identify a gamble X on Ω_1 with the gamble on Ω that assigns the value $X(\omega_1)$ to each pair (ω_1, ω_2), and similarly for gambles on Ω_2.

The behavioural meaning of this definition is that the only gambles that are directly judged to be acceptable are gambles which depend on just one of the marginal outcomes; no judgments are made about any gamble that depends on both ω_1 and ω_2. This model is appropriate when our knowledge about (ω_1, ω_2) consists entirely of our knowledge about each of the components separately, and we do not know anything about how the two components are related.

Example 1 *Suppose that we have two urns. Each of the urns has* 10 *balls which are coloured either red or white. We know that the first urn has* 5 *red,* 2 *white, and* 3 *unknown colours, and the second urn has* 3 *red,* 3 *white, and* 4 *unknown colours. One ball is chosen at random from each of the urns, but we do not assume stochastic independence and it is possible that a correlated joint procedure is used to select the two balls. For example, it could be the case that in each urn the balls are numbered from* 1 *to* 10*, and that a random number,* i*, between* 1 *and* 10 *is selected and ball number* i *is chosen from each urn. The independence here is in our complete lack of information about the interaction between the two drawings. We have information only about the two marginal distributions.*

In this example we are interested in the colours of the two balls that are drawn from the urns. The outcome ω_1 *denotes the colour of the ball drawn from the first urn, and* ω_2 *denotes the colour of the ball drawn from the second urn. The only possible colours are red and white, so that* $\Omega_1 = \Omega_2 = \{red,\ white\}$.

For $i = 1, 2$*, let* $\mathcal{K}_i$ *denote the convex set of probability distributions that models our uncertainty about the colour of the ball drawn from urn* i*. Then* $\mathcal{K}_1$ *contains all probability distributions* P_1 *on* Ω_1 *such that* $0.5 \leq P_1(\{red\}) \leq 0.8$ *and* $\mathcal{K}_2$ *contains all probability distributions* P_2 *on* Ω_2 *for which* $0.3 \leq P_2(\{red\}) \leq 0.7$*. Under the unknown interaction model, the joint set* $\mathcal{K}$ *contains all joint probability distributions on* Ω *whose marginal distributions* P_1 *and* P_2 *satisfy both constraints.*

Here $\mathcal{K}$ *is a very large set. For example,* $\mathcal{K}$ *contains the joint distribution with* $P(\{(red, red)\}) = P(\{(white, white)\}) = 0.5$*, under which the colour of the first ball completely determines the colour of the second ball. Similarly* $\mathcal{K}$ *contains the joint distribution with* $P(\{(red, white)\}) = P(\{(white, red)\}) = 0.5$*. By considering these two distributions, we see that the upper and lower probabilities that the second ball will be red, conditional on the colour of the first ball, are* 1 *and* 0*. Under this model, the effect of learning the colour of one ball is to make our beliefs about the colour of the other ball less precise than they were initially. Observing one outcome changes our uncertainty about the other outcome, contrary to the intuitive notion of independence. This is an extreme example of a phenomenon called* dilation [9].

Let S *denote the event that both balls drawn have the same colour. By considering the same two distributions, we see that the upper and lower probabilities of* S *under this model are* $\overline{P}(S) = \max\{P(S) : P \in \mathcal{K}\} = 1$ *and* $\underline{P}(S) = \min\{P(S) : P \in \mathcal{K}\} = 0$*, which indicates that the joint probabilities under this model are highly imprecise. The later concepts of independence produce increasingly narrower intervals* $[\underline{P}(S), \overline{P}(S)]$.

The lower probabilities under this model satisfy $\underline{P}(\{(red, red)\}) = 0 < 0.15 = \underline{P}_1(\{red\})\underline{P}_2(\{red\})$*, so that the model violates the factorization condition* $\underline{P}(A_1 \times A_2) = \underline{P}_1(A_1)\underline{P}(A_2)$ *(for* $A_1 \subseteq \Omega_1$*,* $A_2 \subseteq \Omega_2$*) which is satisfied by most of the later concepts of independence.*

Unknown interaction produces a very imprecise model, but it can be exploited in local computation problems [5, 6, 11]. Clearly it is not a generalization of stochastic independence. However, it does capture a basic notion of independence, in the sense that we have information about the two components separately but we do not know anything about how they are related. All the following definitions of independence are generalizations of stochastic independence.

3.2 Random Set Independence

This type of independence applies only to belief functions [4, 10], which are a special type of lower probability function. A *belief function* is a lower probability function that can be written in the form

$$\underline{P}(A) = \sum_{B \subseteq A} m(B)$$

where m is a mapping from 2^{Ω} to $[0,1]$ such that $m(\emptyset) = 0$ and $\sum_{B \subseteq \Omega} m(B) = 1$. A mapping m satisfying these conditions is called a *mass function*.

Suppose that the two marginal experiments are described by belief functions $\underline{P}_1$ and $\underline{P}_2$ which have corresponding mass functions m_1 and m_2. We say that there is *random set independence* [4, 13] if and only if uncertainty about the joint experiment is described by the belief function $\underline{P}$ on $\Omega = \Omega_1 \times \Omega_2$ whose mass function is defined by

$$m(A_1 \times A_2) = m_1(A_1)m_2(A_2)$$

when $A_1 \subseteq \Omega_1$ and $A_2 \subseteq \Omega_2$, with $m(A) = 0$ for all subsets of Ω which are not of the form $A = A_1 \times A_2$.

There is a simple way of obtaining the global set $\mathcal{K}$ from the mass function m: it is the set of probability distributions obtained by assigning each probability mass $m(A_1 \times A_2)$ arbitrarily to elements of $A_1 \times A_2$ [4]. However, the expression for $\mathcal{K}$ in terms of the marginal sets $\mathcal{K}_1$ and $\mathcal{K}_2$ is complicated, because it is not easy to express the mass function m_i as a function of $\mathcal{K}_i$.

This definition of independence can be justified under the following assumptions: (a) there are two random experiments with possibility spaces Λ_1 and Λ_2, each of which is modeled by a known probability distribution; (b) each space Λ_i is related to Ω_i through a *multivalued mapping* μ_i [4], meaning that if λ_i is the outcome of random experiment i then we learn only that the true state of Ω_i belongs to the subset $\mu_i(\lambda_i)$; (c) the probability distribution on Λ_i induces the mass function m_i on Ω_i through the multivalued mapping μ_i, for $i = 1, 2$; (d) the probability distributions on Λ_1 and Λ_2 are stochastically independent; and (e) we know nothing about the interaction between the two mechanisms for selecting the outcomes ω_1 and ω_2 from the sets $\mu_1(\lambda_1)$ and $\mu_2(\lambda_2)$. Under these assumptions, the joint mass function m, defined above, is the appropriate model for uncertainty about the joint outcome (ω_1, ω_2). These assumptions are illustrated by the next example.

Example 2 *Suppose there are two urns as in Example 1, but now the balls of unknown colour are actually not painted until after the drawings are made. We select one ball from each urn in a stochastically independent way, and if either of the selected balls are not coloured then they are painted white or red by a completely unknown procedure. If both selected balls have no colour then there can be arbitrary correlation between the colours they are assigned.*

In this case, the marginal mass functions for the colours of the two selected balls are

- *First urn:* $m_1(\{red\}) = 0.5$, $m_1(\{white\}) = 0.2$, $m_1(\{white, red\}) = 0.3$.
- *Second urn:* $m_2(\{red\}) = 0.3$, $m_2(\{white\}) = 0.3$, $m_2(\{white, red\}) = 0.4$.

For the joint colours of the two selected balls, the mass function is given by $m(A_1 \times A_2) = m_1(A_1)m_2(A_2)$. *For example, mass* 0.15 *is assigned to obtaining a red ball from each of the urns. With mass* 0.12 *we draw two unpainted balls, and in this case (since we know nothing about the procedure for assigning colours) we have absolutely no information about the two colours.*

For the event S *that the colours of the balls drawn from the two urns are the same, we obtain the lower and upper probabilities* $\underline{P}(S) = m_1(\{red\})m_2(\{red\}) + m_1(\{white\})m_2(\{white\}) = 0.21$, *and* $\overline{P}(S) = 1 - m_1(\{red\})m_2(\{white\}) - m_1(\{white\})m_2(\{red\}) = 0.79$. *The lower and upper probabilities are achieved by the extreme points of* $\mathcal{K}$ *which respectively assign the probabilities* $(0.15, 0.56, 0.23, 0.06)$ *and* $(0.56, 0.15, 0.06, 0.23)$ *to the four possible joint outcomes (RR,RW,WR,WW). The same two extreme points achieve the lower and upper probabilities that the second ball will be red, given that the first ball is red, which are* $15/71 = 0.211$ *and* $56/71 = 0.789$. *The interval* $[15/71, 56/71]$ *is wider than* $[0.3, 0.7]$, *the interval of marginal probabilities for the second drawing. Again, learning the colour of one ball changes our uncertainty about the colour of the other ball, contrary to the intuitive notion of independence.*

This definition of independence enables us to construct a joint belief function $\underline{P}$ from the marginal belief functions $\underline{P}_1$ and $\underline{P}_2$. The joint model produced by random set independence is more precise than the unknown interaction model, in the sense that the joint set $\mathcal{K}$ contains fewer probability distributions, because it requires some independence in the selection of the values ω_1 and ω_2 (except when either marginal distribution on Λ_i is degenerate). The sets A_1 and A_2 are selected from each component in a stochastically independent way, but inside the sets $A_1 \times A_2$ we allow dependent selections.

3.3 Epistemic Irrelevance and Irrelevant Natural Extension

The intuitive meaning of 'independence' is that learning the outcome of one experiment would not change our uncertainty about the other experiment. In other words, one experiment is *irrelevant* to the other [12, 2]. In behavioural terms, this means that the set of acceptable gambles concerning the second experiment does not change when we learn the outcome of the first experiment. For imprecise probabilities, irrelevance is a directional or asymmetric relation. For precise probabilities, however, irrelevance is symmetric and it agrees with stochastic independence. In this subsection we consider the property that the first experiment is irrelevant to the second, which we call epistemic irrelevance, and in the next subsection we consider the stronger property that each experiment is irrelevant to the other, which we call epistemic independence.

Let $\mathcal{K}_1, \mathcal{K}_2$ and $\mathcal{K}$ denote the marginal and joint sets of probability distributions. For a joint distribution P in $\mathcal{K}$ and $\omega_1 \in \Omega_1$, let $P_2(\cdot|\omega_1)$ denote the conditional probability distribution on Ω_2 given ω_1. This is obtained by Bayes' rule, $P_2(\{\omega_2\}|\omega_1) = P(\Omega_1 \times \{\omega_2\}|\{\omega_1\} \times \Omega_2) = P(\{(\omega_1, \omega_2)\})/P(\{\omega_1\} \times \Omega_2)$.

We say that the first experiment is *epistemically irrelevant* to the second when $\{P_2(\cdot|\omega_1) : P \in \mathcal{K}\} = \mathcal{K}_2$ for all $\omega_1 \in \Omega_1$. That is, the set of conditional probability distributions (given the outcome of the first experiment) agrees with the marginal set of distributions $\mathcal{K}_2$. This captures the intuitive idea that learning the outcome of the first experiment would not change our uncertainty about the second experiment.

Given only the two marginal convex sets $\mathcal{K}_1$ and $\mathcal{K}_2$, we can construct the *largest* set of joint distributions, $\mathcal{K}$, under which the first experiment is epistemically irrelevant to the second: this is the set of all joint distributions P for which (a) the marginal distribution on Ω_1 is in $\mathcal{K}_1$; and (b) the conditional distributions $P(\cdot|\omega_1)$ are in $\mathcal{K}_2$, for all $\omega_1 \in \Omega_1$. This set $\mathcal{K}$ is called the *irrelevant natural extension* of the marginals $\mathcal{K}_1$ and $\mathcal{K}_2$. It is the set of all joint distributions P which have the form

$$P(\{(\omega_1, \omega_2)\}) = P_1(\{\omega_1\})P_2(\{\omega_2\}|\omega_1),$$

for some $P_1 \in \mathcal{K}_1$ and $P_2(\cdot|\omega_1) \in \mathcal{K}_2$. Here ω_1 is selected according to some marginal distribution in $\mathcal{K}_1$, and then ω_2 is selected according to a distribution from $\mathcal{K}_2$ which may depend on ω_1. Note that ω_2 may be selected by a different procedure for different values of ω_1. The irrelevant natural extension satisfies independence of the marginal sets.

This model is appropriate when we are given the marginal sets $\mathcal{K}_1$ and $\mathcal{K}_2$, and we judge only that learning ω_1 should not change our uncertainty about ω_2 (but not vice versa).

Example 3 *Suppose now that we have three urns. The first urn has the same contents as the first urn in Example 1. Our knowledge about the other two urns is the same as for urn 2 in Example 1, but it is not necessary that urns 2 and 3 have exactly the same composition: the unknown balls can have different colours in each urn. The procedure to select the two balls is as follows. A ball is randomly selected from the first urn. If the first ball is red then the second ball is selected randomly from the second urn, and if the first ball is white then the second ball is selected randomly from the third urn.*

Now the uncertainty about the two colours is modeled by the convex set of all joint probability distributions P of the form $P(\{(red, \omega_2)\}) = P_1(\{red\})P_2(\{\omega_2\}|red)$ and $P(\{(white, \omega_2)\}) = [1 - P_1(\{red\})]\ P_2(\{\omega_2\}|white)$, where $0.5 \leq P_1(\{red\}) \leq 0.8$, $0.3 \leq P_2(\{\omega_2\}|red) \leq 0.7$, and $0.3 \leq P_2(\{\omega_2\}|white) \leq 0.7$. This set has 8 extreme points which can be obtained from all possible combinations of the extreme values of $P_1(\{red\})$, $P_2(\{red\}|red)$ and $P_2(\{red\}|white)$. Four of these extreme points, the ones for which $P_2(\{red\}|red) \neq P_2(\{red\}|white)$, do not satisfy stochastic independence. For example, taking $P_1(\{red\}) = 0.8$, $P_2(\{red\}|red) = 0.7$ and $P_2(\{red\}|white) = 0.3$ gives the probability distribution $(0.56, 0.24, 0.06, 0.14)$ for the possible joint outcomes (RR, RW, WR, WW). This illustrates that, although ω_1 is epistemically irrelevant to ω_2, it may be stochastically relevant in the sense that the occurrence of ω_1 may change the physical probability distribution of ω_2: the physical probability that the second ball is red may depend on the colour of the first ball.

Taking $S = \{RR, WW\}$ to be the event that the two colours are the same, the upper and lower probabilities of S are $\overline{P}(S) = 0.7$ (achieved by the preceding distribution) and $\underline{P}(S) = 0.3$.

Under this model, learning the colour of the first ball does not change our uncertainty about the colour of the second ball. The conditional upper and lower probabilities that the second ball is red, given the colour of the first ball, are 0.7 and 0.3, the same as the marginal upper and lower probabilities. However, we find that the conditional upper and lower probabilities that the first ball is red, given the colour of the second ball, are $28/31 = 0.903$ and 0.3, which differ from, and are less precise than, the marginal upper and lower probabilities 0.8 and 0.5. This is another example of dilation. Thus the second experiment is epistemically relevant to the first. This shows that epistemic irrelevance is not a symmetric relation.

3.4 Epistemic Independence and Independent Natural Extension

The next concept of independence, which we call *epistemic independence* [12, 2], is characterized by the property that our uncertainty about either of the two outcomes does not change when we obtain some information about the other outcome. In other words, each experiment is epistemically irrelevant to the other experiment. Unlike epistemic irrelevance, epistemic independence is a symmetric relation.

The mathematical definition is as follows. As in the previous subsection, for a joint distribution P in $\mathcal{K}$ and $\omega_1 \in \Omega_1, \omega_2 \in \Omega_2$, let $P_2(\cdot|\omega_1)$ denote the conditional probability distribution on Ω_2 given ω_1, defined by Bayes' rule, and let $P_1(\cdot|\omega_2)$ denote the conditional probability distribution on Ω_1 given ω_2. We say that the two experiments are *epistemically independent* when each one is epistemically irrelevant to the other, i.e., when $\{P_2(\cdot|\omega_1) : P \in \mathcal{K}\} = \mathcal{K}_2$ for all $\omega_1 \in \Omega_1$, and $\{P_1(\cdot|\omega_2) : P \in \mathcal{K}\} = \mathcal{K}_1$ for all $\omega_2 \in \Omega_2$.

The behavioural meaning of epistemic independence is that the set of acceptable gambles concerning either experiment would not change if we learned the outcome of the other experiment. Defining $\Gamma_2^{\omega_1}$ to be the set of gambles on Ω_2 that are acceptable conditionally on observing ω_1, and defining $\Gamma_1^{\omega_2}$ similarly, there is epistemic independence if and only if $\Gamma_2^{\omega_1} = \Gamma_2$ and $\Gamma_1^{\omega_2} = \Gamma_1$ for all $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$. Equivalently, for all real-valued functions $X(\omega_2)$, the lower previsions that are generated by $\mathcal{K}$ or Γ satisfy $\underline{P}_2(X|\omega_1) = \underline{P}_2(X)$, and similarly for functions $Y(\omega_1)$.

Given only the two marginal convex sets $\mathcal{K}_1$ and $\mathcal{K}_2$, we can construct the *largest* set of joint distributions, $\mathcal{K}$, under which the two experiments are epistemically independent: this is the set of all joint distributions P for which (a) the conditional distributions $P_2(\cdot|\omega_1)$ are in $\mathcal{K}_2$, for all $\omega_1 \in \Omega_1$; and (b) the conditional distributions $P_1(\cdot|\omega_2)$ are in $\mathcal{K}_1$, for all $\omega_2 \in \Omega_2$. This set $\mathcal{K}$ is called the *independent natural extension* of the marginals $\mathcal{K}_1$ and $\mathcal{K}_2$, since it is simply the natural extension of the sets of conditional probability distributions that are determined by epistemic independence ([12], Section 9.3). The independent natural extension $\mathcal{K}$ can be constructed as the intersection of two irrelevant natural extensions (defined in the previous subsection), or by using linear programming methods to find the extreme probability distributions that satisfy the linear constraints (a) and (b). As in the case of epistemic irrelevance, a joint distribution P in $\mathcal{K}$ can have different conditional distributions $P_2(\cdot|\omega_1)$ for different values of ω_1, and similarly for the distributions conditional on ω_2. The independent natural extension satisfies independence of the marginal sets.

The independent natural extension is the appropriate model when we are given the two marginal sets $\mathcal{K}_1$ and $\mathcal{K}_2$ (or the corresponding sets of acceptable gambles or upper and lower previsions), together with a judgement that the experiments are epistemically independent, but we are not willing to make stronger assumptions, e.g., that there are underlying stochastic mechanisms which are stochastically independent, which would justify the concept of strong independence defined in the next subsection.

Example 4 *Suppose again that there are two urns, and our knowledge about their composition is exactly as in Example 1. One ball is drawn from each urn. All we know about the stochastic mechanism for drawing the two balls, i.e., the joint probability distribution P, is that (a) whatever the colour of the first ball, the conditional probability that the second ball is red lies between* 0.3 *and* 0.7*, and (b) whatever the colour of the second ball, the conditional probability that the first ball is red lies between* 0.5 *and* 0.8*. Since these upper and lower bounds agree with the marginal upper and lower probabilities, learning the colour of either ball does not change our uncertainty about the colour of the other ball. Here we allow the possibility that there may be some physical interaction between the two drawings, as in Example 3, but we assume that the drawings are epistemically independent.*

In this case the appropriate model is the independent natural extension of the two marginals. The upper and lower probabilities of any events in the joint space Ω, and the extreme points of the joint convex set $\mathcal{K}$, can be found using linear programming techniques. For example, for the event S that the two balls drawn have the same colour, the upper and lower probabilities under this model are found to be $\overline{P}(S) = 40/59 = 0.678$ *and* $\underline{P}(S) = 19/59 = 0.322$*. The upper probability is achieved by the joint probability distribution* $(28/59, 12/59, 7/59, 12/59)$ *for (RR, RW, WR, WW), which is an extreme point of $\mathcal{K}$ but not a product of marginal distributions. The independent natural extension $\mathcal{K}$ is strictly contained in the set $\mathcal{K}$ produced by the weaker judgement of epistemic irrelevance in Example 3, but because it has extreme points which do not factorize, it strictly contains the set $\mathcal{K}$ that is produced by the following definition of strong independence.*

3.5 Independence in the Selection and Strong Independence

In some problems we know that there are underlying stochastic mechanisms for the two experiments that are stochastically independent. We say that there is *independence in the selection* when every joint probability distribution P that is an extreme point of $\mathcal{K}$ factorizes as $P = P_1 \times P_2$, where P_i is the marginal distribution of P on Ω_i [3, 12]. (Such a model was called an *independent envelope* in [12].) That is, writing ext($\mathcal{K}$) for the set of extreme points of $\mathcal{K}$, there is independence in the selection if and only if

$$\text{ext}(\mathcal{K}) \subseteq \{P_1 \times P_2 \,:\, P_1 \in \mathcal{K}_1,\, P_2 \in \mathcal{K}_2\}.$$

This definition is intended to capture the idea of non-interaction in the random selection of ω_1 and ω_2. It does not require that there is complete independence between the procedures that produce ω_1 and ω_2, and it is possible that we know some relationship between them, but there is independence in the stochastic part of the selection. In some cases we may have information about linkage between the underlying marginal probability distributions P_1 and P_2, which rules out some of the combinations $P_1 \times P_2$ such that $P_1 \in \mathcal{K}_1$ and $P_2 \in \mathcal{K}_2$. Independence in the selection implies epistemic independence of the marginal experiments.

Independence in the selection is appropriate when the following assumptions are satisfied: (a) the two outcomes result from random experiments, each governed by a unique (but unknown) probability distribution; (b) we know that the two probability distributions belong to the sets $\mathcal{K}_1$ and $\mathcal{K}_2$ respectively; and (c) the random experiments are stochastically independent. Often (c) can be justified through knowledge about causal unrelatedness of the two experiments.

Example 5 *Assume that we have two urns with the same composition as in Example 1, and we also know that:*

- *the* 7 *balls in the two urns whose colours are unknown are all the same colour;*
- *the drawings from the two urns are stochastically independent.*

There is some interaction between the experiments, in the sense that having more red balls in the first urn implies having more in the second. There are two extreme marginal probability distributions for each urn:

- *First urn:* P_1^1 *(*8 *red and* 2 *white) and* P_1^2 *(*5 *red and* 5 *white).*
- *Second urn:* P_2^1 *(*7 *red and* 3 *white) and* P_2^2 *(*3 *red and* 7 *white).*

The only possible joint probability distributions are $P_1^1 \times P_2^1$ *and* $P_1^2 \times P_2^2$*. The first piece of additional information rules out the two joint distributions* $P_1^1 \times P_2^2$ *and* $P_1^2 \times P_2^1$*. Thus* $\mathcal{K} = \text{CH}\{P_1^1 \times P_2^1, P_1^2 \times P_2^2\}$*.*

Consider the event A *that a white ball is drawn from the first urn and a red ball is drawn from the second urn. The only possible values of the probability* $P(A)$ *are* $\underline{P}(A) = 0.14$ *and* $\overline{P}(A) = 0.15$*. Also* $\underline{P}(A) = \underline{P}(\{(white, red)\}) = 0.14 > 0.06 = \underline{P}_1(\{white\})\underline{P}_2(\{red\})$ *and* $\overline{P}(A) = \overline{P}(\{(white, red)\}) = 0.15 < 0.35 = \overline{P}_1(\{white\})\overline{P}_2(\{red\})$*, so this model violates the factorization conditions.*

For the event S *that both balls are the same colour, this model gives* $\underline{P}(S) = 0.5$ *and* $\overline{P}(S) = 0.62$*, which are much more precise than the previous values.*

Independence in the selection captures an important aspect of independence, but it cannot be regarded as a definition of a complete absence of any relationship between the two components. In an extreme case, a functional relationship between the two components satisfies the definition. Imagine, for example, the situation in which the 20 balls in the two urns are either all red or all white. In that case there is independence in the selection, but the colour of a randomly selected ball from one urn completely determines the colour of a randomly selected ball from the other urn.

The largest set of joint probability distributions that satisfies independence in the selection is

$$\mathcal{K} = \text{CH}\{P_1 \times P_2 \,:\, P_1 \in \mathcal{K}_1,\, P_2 \in \mathcal{K}_2\}.$$

When this equality is satisfied, we say that there is *strong independence*. (This model was called a *type-1 extension* in [12] and [2].)

This definition is intended to capture a complete lack of interaction between the two components. This concept of independence has been studied in [3, 12, 13], and it has been adopted by most of the authors who have modeled independence using imprecise probabilities.

Strong independence is appropriate when assumptions (a)-(c) are satisfied and also: (d) we do not know of any relationship between the two marginal probability distributions that would enable us to rule out some of the possible combinations of marginal distributions. In other words, there is independence of the marginal sets as defined in subsection 3.1. Strong independence is equivalent to independence in the selection plus independence of the marginal sets.

Example 6 *Consider the two urns of Example 1, and suppose that a ball is selected from each urn in a stochastically independent way. The possible relative frequencies of red balls in each urn are* $0.5, 0.6, 0.7$ *and* 0.8 *for the first urn, and* $0.3, 0.4, 0.5, 0.6$ *and* 0.7 *for the second urn. Because the two drawings are stochastically independent, the probability of drawing (say) two red balls is the product of the two relative frequencies, which can take any of* 16 *possible values ranging from* $0.5 \times 0.3 = 0.15$ *to* $0.8 \times 0.7 = 0.56$*. The interval* $[0.15, 0.56]$ *is the convex hull of the possible probabilities of drawing two red balls, and it represents our uncertainty about this event.*

More generally, the set of possible joint proba-

bility distributions concerning the two colours is $\{P_1 \times P_2 : P_1 \in \mathcal{K}_1, P_2 \in \mathcal{K}_2\}$*, where* $\mathcal{K}_i$ *is the set of* 4 *or* 5 *probability distributions concerning the colour drawn from urn* i*. The convex hull of this set has* 4 *extreme points which can be obtained by combining either of the two extreme points of* $\mathcal{K}_1$ *with either of the two extreme points of* $\mathcal{K}_2$*. (Clearly it would make no difference to this model, or to upper and lower probabilities and previsions, if we replaced* $\mathcal{K}_i$ *here by its convex hull.) The four extreme distributions on* $\Omega = \{RR, RW, WR, WW\}$ *are* $(0.56, 0.24, 0.14, 0.06)$*,* $(0.24, 0.56, 0.06, 0.14)$*,* $(0.35, 0.15, 0.35, 0.15)$ *and* $(0.15, 0.35, 0.15, 0.35)$*. For the event* S *that both balls have the same colour, strong independence produces the upper and lower probabilities* $\overline{P}(S) = 0.62$ *and* $\underline{P}(S) = 0.38$*, which are more precise than the probabilities produced by epistemic independence, but less precise than the probabilities in Example 5, where the upper probability was the same but the lower probability was* 0.5*.*

The model in Example 5 determines some probabilities much more precisely than strong independence. Consider the event A *that a white ball is drawn from the first urn and a red ball is drawn from the second urn. Under strong independence, the possible probabilities* $P(A)$ *range from* 0.06 *to* 0.35*. Under the model in Example 5, the interval* $[0.14, 0.15]$ *was much narrower, because the additional information produced a smaller joint set* $\mathcal{K}$*.*

Strong independence implies the earlier independence properties of independence of the marginal sets, epistemic irrelevance, epistemic independence and independence in the selection. In particular, learning the outcome of one experiment does not change our uncertainty about the other experiment, in accordance with the intuitive notion of independence.

The probabilities produced by strong independence are always at least as precise as those produced by all the definitions given in the previous subsections, i.e., the joint set $\mathcal{K}$ under strong independence is always a subset of the joint sets under the earlier models. However, there is an important case in which strong independence agrees with random set independence, the irrelevant natural extension and the independent natural extension: whenever $A_1 \subseteq \Omega_1$ and $A_2 \subseteq \Omega_2$, these four models produce the same upper and lower probabilities for the product set $A_1 \times A_2$, which are given by the factorization formulae $\overline{P}(A_1 \times A_2) = \overline{P}_1(A_1)\overline{P}_2(A_2)$ and $\underline{P}(A_1 \times A_2) = \underline{P}_1(A_1)\underline{P}_2(A_2)$. The other models do not always satisfy these factorization properties.

3.6 Repetition Independence

Now suppose that the two experiments have the same set of possible outcomes, $\Omega_1 = \Omega_2$, and that each experiment is governed by the same probability distribution, but we know only that this probability distribution belongs to the set $\mathcal{K}_1$. In this case we not only have identical information about the two experiments, but we also know that the two outcomes are identically distributed. If we also know that the two experiments are stochastically independent, then the joint probability distribution is of the form $P \times P$, for some P in $\mathcal{K}_1$. Under these assumptions the convex set of joint probability distributions on $\Omega_1 \times \Omega_2$ is

$$\mathcal{K} = \mathrm{CH}\{P \times P : P \in \mathcal{K}_1\}.$$

When this condition is satisfied, we say that there is *repetition independence* [12, 13]. The additional knowledge that the two experiments are identical, i.e., that they are governed by the same probability distribution, reduces the set $\mathcal{K}$ and produces a more precise global model than we would obtain from strong independence. Repetition independence is a special type of independence in the selection, where we know that the two marginal probability distributions are the same. Again, repetition independence implies epistemic independence.

Repetition independence is the appropriate definition of independence in statistical problems, where we have stochastically independent repetitions of a random experiment and our only knowledge about the underlying probability distribution is that it belongs to the set $\mathcal{K}_1$.

Example 7 *Suppose that we have two identical urns, and our knowledge about them is the same as our knowledge about the first urn in Example 1: we know that each urn has* 5 *red balls,* 2 *white balls, and* 3 *balls of unknown colours. Now we know also that the number of red balls is the same in each urn. Suppose that a ball is selected from each urn in a stochastically independent way. The possible relative frequencies of red balls in each urn are* $0.5, 0.6, 0.7$ *and* 0.8*, with the same relative frequency in each urn.*

Consider the event A *that a white ball is drawn from the first urn and a red ball is drawn from the second urn. Because the drawings are stochastically independent, the only possible values of* $P(A)$ *are now* $0.25, 0.24, 0.21$ *and* 0.16*. Compare with the strong independence model, under which the values of* $P(A)$ *range from* 0.1 *to* 0.4*. Since* $\underline{P}(A) = 0.16 > 0.1 = \underline{P}_1(\{white\})\underline{P}_2(\{red\})$ *and* $\overline{P}(A) = 0.25 < 0.4 = \overline{P}_1(\{white\})\overline{P}_2(\{red\})$*, the repetition independence model violates the factorization conditions.*

We would obtain a slightly different joint model here if we replaced the marginal set $\mathcal{K}_1$*, which contains just the four possible marginal distributions with* $P(\{red\}) = 0.5, 0.6, 0.7$ *or* 0.8*, by its convex hull* $\mathrm{CH}(\mathcal{K}_1)$*. For example, consider the gamble* X *which takes the value* 3 *if* RW *occurs,* 1 *if* RR*, and* 0 *otherwise. If the joint model* $\mathcal{K}$ *is constructed using repetition independence from the marginal* *set $\mathcal{K}_1$, we find that the upper prevision of X is 1.120, achieved when $P(\{red\})$ is 0.7 or 0.8. But if the joint model is constructed from the marginal set $\mathrm{CH}(\mathcal{K}_1)$, the upper prevision of X is 1.125, achieved when $P(\{red\})$ is 0.75. In this case, convexity of the marginal set of probability distributions does have some behavioural significance, because it affects the joint model that is formed using repetition independence.*

	$\underline{P}(S)$	$\overline{P}(S)$
Unknown Interaction	0.00	1.00
Random Set Independence	0.21	0.79
Irrelevant Natural Extension	0.30	0.70
Independent Natural Extension	0.32	0.68
Strong Independence	0.38	0.62

Table 1: Probability intervals for the event S under different definitions of independence.

4 Conclusions

In probability theory, there is essentially only one concept of independence: the different definitions are essentially equivalent, except in their treatment of events that have probability zero. In this paper we have shown that independence is more complex when imprecise probabilities are involved. There are several definitions of independence which are not equivalent.

We have presented six definitions of independence which produce a joint model from given marginals, in order of increasing precision: the later definitions produce smaller joint convex sets $\mathcal{K}$ and more precise inferences. Table 1 summarizes the (lower, upper) probability intervals for the event S, that the two balls drawn from the urns of Example 1 have the same colour, under these different definitions, excluding repetition independence. (Repetition independence is inapplicable in this example because the two marginal sets are different.)

All of the definitions of independence given in Section 3 seem to be useful in particular kinds of application. Clearly, the choice of an appropriate definition must depend on the type of application. We hope that our discussion of the assumptions on which each concept is based, and our examples, will help in selecting the most appropriate concept.

Which independence concepts are likely to be the most useful and most frequently applicable in applications? We think that, when ω_1 and ω_2 are the outcomes of two random experiments that are stochastically independent, strong independence will usually be the most appropriate concept. If there is additional information about the linkage between the stochastic mechanisms then a more precise model, involving independence in the selection or repetition independence, may be appropriate. All three of these models are based on assumptions of stochastic independence. In examples of drawing balls from urns, it is natural to assume both that there are underlying marginal probability distributions (determined by the composition of each urn) and that the drawings are physically unrelated and therefore stochastically independent. Under these assumptions, strong independence, or one of its modifications, is appropriate.

However, in many practical applications there is no underlying stochastic mechanism which would justify strong independence, and then the choice of an appropriate definition of independence is less clear. The authors of this paper disagree slightly about which independence concepts are most useful in these cases. One of us (PW) thinks that, because of their simple behavioural meaning, epistemic independence (or irrelevance) and independent (or irrelevant) natural extension are likely to be the most frequently applicable concepts; they are purely epistemic concepts which require no assumptions about underlying stochastic mechanisms. The other two authors (IC and SM) think that strong independence is applicable even in cases in which no underlying stochastic mechanism is assumed: it is enough that the values ω_1 and ω_2 are produced by physical procedures that are causally unrelated. However, without an assumption of underlying stochastic independence, no behavioural justification for strong independence is available at present.

We believe that the most important role for a concept of independence is in constructing a joint or global model from simpler components. In most applications of the concept, we do not extract independence from a global model, but rather we perceive independence as a primitive concept and then use it to construct the global model. In the formulation adopted in this paper, we would construct the joint convex set $\mathcal{K}$ from the marginal convex sets $\mathcal{K}_1$ and $\mathcal{K}_2$, using only the judgement that the marginal experiments are independent. Most of the concepts defined in this paper can be used in this way, but some are easier to use than others. For example, to construct the independent natural extension it is necessary to use linear programming methods, whereas construction of the joint model using random sets independence, irrelevant natural extension or strong independence involves simpler calculations.

We are presently studying the computational aspects of epistemic independence and the other definitions. In future work, we also plan to extend the independence concepts to conditional independence, to try to find other characterizations of the concepts which will help to clarify their meaning, and to study more closely the relationships between the concepts.

Acknowledgements

Peter Walley's research on this project was supported by FAPESP and the Escola Politécnica, Universidade de São Paulo. The authors thank Fabio Cozman and two referees for helpful comments.

References

[1] E. Castillo, J.M. Gutiérrez and A.S. Hadi. *Expert Systems and Probabilistic Network Models*. Springer Verlag, New York, 1997.

[2] F. Cozman. Irrelevance and independence relations in quasi-Bayesian networks. In G.F. Cooper and S. Moral, eds, *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 89–96. Morgan & Kaufmann, San Mateo, 1998.

[3] L.M. de Campos and S. Moral. Independence concepts for convex sets of probabilities. In Ph. Besnard and S. Hanks, eds, *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, 108–115. Morgan & Kaufmann, San Mateo, 1995.

[4] A.P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[5] P. Hansen and B. Jaumard. Probabilistic satisfiability. To appear in: J. Kohlas and S. Moral, eds, *Algorithms for Uncertain and Defeasible Reasoning*, Kluwer, Dordrecht, 1999.

[6] S. Moral. Algorithms for imprecise probabilities. To appear in: J. Kohlas and S. Moral, eds, *Algorithms for Uncertain and Defeasible Reasoning*, Kluwer, Dordrecht, 1999.

[7] S. Moral and N. Wilson. Revision rules for convex sets of probabilities. In G. Coletti, D. Dubois and R. Scozzafava, eds, *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence*, 113–128. Plenum Press, New York, 1995.

[8] J. Pearl. *Probabilistic Reasoning with Intelligent Systems*. Morgan & Kaufman, San Mateo, 1988.

[9] T. Seidenfeld and L. Wasserman. Dilation for sets of probabilities. *Annals of Statistics*, 21:1139–1154, 1993.

[10] Ph. Smets. Belief functions. In Ph. Smets, A. Mamdani, D. Dubois and H. Prade, eds, *Non-Standard Logics for Automated Reasoning.* Academic Press, London, 1988.

[11] J.F. Verdegay-López. *Representación y Combinación de la Información con Incertidumbre Mediante Convexos de Probabilidades*. PhD thesis, Universidad de Granada, 1997.

[12] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[13] P. Walley and T.L. Fine. Towards a frequentist theory of upper and lower probability. *Annals of Statistics*, 10:741–761, 1982.

[14] N. Wilson and S. Moral. A logical view of probability. In A. Cohn, ed., *Proceedings of the 11th European Conference on Artificial Intelligence (ECAI'94)*, 386–390. Wiley, London, 1994.