Chapter 6

Logics for Reasoning About Uncertainty

We can design a logic for reasoning about each of the methods of uncertainty discussed in Chapter 3. The general approach is much like that for the logics of knowledge and relative likelihood discussed in Chapter 2. A simple structure for each of these notions in the case of one agent has the form (W, w, X, π), where, as before, W is a set of worlds, w is the actual world, and π is an interpretation. X depends on the method of representation that we use. It can be either a probability on W if we want to reason about probability, a Dempster-Shafer belief function if we want to reason about them, and so on. We can then generalize to allow the representation of uncertainty being considered—probability, possibility, etc.—to depend on the world, and to be different for each agent. Thus, in general, a structure has the form $(W, X_1, \ldots, X_n, \pi)$. We have already seen two instances of this: if $X_i = \mathcal{K}_i$, an accessibility relation, then we get an epistemic structure; if $X_i = \mathcal{O}_i$, then we get a preferential structure.

All the other types of structures considered in this book follow the same general pattern. At the risk of boring the reader, I briefly describe them all here, just to establish the notation.

- In the case of probabilities, $X_i = \mathcal{PR}_i$, where $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$, $\mathcal{F}_{w,i}$ is an algebra of subsets of $W_{w,i} \subseteq W$, and $\mu_{w,i}$ is a probability measure on $\mathcal{F}_{w,i}$. This gives us a *probability structure*. An important special case is where $\mathcal{F}_{w,i} = 2^{W_{w,i}}$, so that all sets are measurable. A probability structure where all sets are measurable is called a *measurable probability structure*. The function $\mathcal{PR}_i$ in a probability structure is called a *probability assignment* (for agent i).

- In the case of lower probabilities, $X_i = \mathcal{LP}_i$, where $\mathcal{LP}_i(w) = (W_{w,i}, \mathcal{P}_{w,i})$ and $\mathcal{P}_{w,i}$ is a set of probability measures on $W_{w,i} \subseteq W$. This gives us a *lower probability structure*.

- In the case of belief functions, $X_i = \mathcal{BEL}_i$, where $\mathcal{BEL}_i(w) = (W_{w,i}, \mathrm{Bel}_{w,i})$ and $\mathrm{Bel}_{w,i}$ is a belief function on $W_{w,i} \subseteq W$. This gives us a *belief structure*.

- For possibility measures, $X_i = \mathcal{POSS}_i$, where $\mathcal{POSS}_i(w) = (W_{w,i}, \mathrm{Poss}_{w,i})$ and $\mathrm{Poss}_{w,i}$ is a possibility measure on $W_{w,i} \subseteq W$. This gives us a *possibility structure*.

- For ranking structures, $X_i = \mathcal{RANK}_i$, where $\mathcal{RANK}_i(w) = (W_{w,i}, \kappa_{w,i})$ and $\kappa_{w,i}$ is a ranking function on $W_{w,i} \subseteq W$. This gives us a *ranking structure*.

- Finally, for plausibility measures, $X_i = \mathcal{PL}_i$, where $\mathcal{PL}_i(w) = (W_{w,i}, \mathrm{Pl}_{w,i})$ and $\mathrm{Pl}_{w,i}$ is a plausibility measure on $W_{w,i} \subseteq W$. This gives us a *plausibility structure*.

Let $\mathcal{M}_n^{prob}$, $\mathcal{M}_n^{meas}$, $\mathcal{M}_n^{lp}$, $\mathcal{M}_n^{bel}$, $\mathcal{M}_n^{poss}$, $\mathcal{M}_n^{rank}$, and $\mathcal{M}_n^{plaus}$ be the class of all probability structures, measurable probability structures, lower probability structures, belief structures, possibility structures, ranking structures, and plausibility structures, respectively, for n agents.

In the "simple" version of $\mathcal{M}_n^{prob}$ there is just a probability measure, rather than a function from worlds to probability measures. There are similar "simple" versions for all the other notions of uncertainty. For example, a simple possibility structure has the form (W, Poss, π). I drop the subscript n to denote the simple version of these structures, writing, for example, $\mathcal{M}^{prob}$ to denote the class of simple probability structures. I leave the formal details to the reader.

What about the language? What kinds of things do we want to say? To some extent, that depends on what we *can* say, which in turn depends on the representation method being used. I actually consider two languages here. For both, the syntax includes *likelihood terms* of the form $\ell(\varphi)$. This can be interpreted as "the probability of φ", "the possibility of φ", "the plausibility of φ", and so on, depending on the underlying representation. (The ℓ stands for *likelihood*.) The first language allows arithmetic operations for forming more complicated likelihood terms, such as $\ell(\varphi) + \ell(\psi)$. This makes sense for probability and possibility, where $\ell(\varphi)$ and $\ell(\psi)$ are real numbers that can be added, but not for plausibility measures. The second language allows only comparison of likelihoods, and makes sense for all the representation methods.

6.1 Reasoning about Probability: The Measurable Case

I start by considering the case of measurable probability structures and the richer language, that allows the addition of probability terms. This allow us to say, for example, that the probability of the union of two disjoint sets is the sum of their individual probabilities. In fact, I allow more than just addition; I allow an arbitrary linear combination of likelihood terms.

The syntax is quite simple. Starting with a set Φ of primitive propositions. close off under conjunction, negation, and *(linear) likelihood formulas*, where a likelihood formula has the form $a_1\ell_{i_1}(\varphi_1) + \cdots + a_k\ell_{i_k}(\varphi_k) > b$, where $a_1, \ldots, a_k, b$ are integers, $i_1, \ldots, i_k$ are (not necessarily distinct) agents, and $\varphi_1, \ldots, \varphi_k$ are formulas. Thus, a linear likelihood formula talks about a linear combination of likelihood terms of the form $\ell_i(\varphi)$. For example, $2\ell_1(p_1 \wedge p_2) + 7\ell_1(p_1 \vee \neg p_3) \geq 3$ is a likelihood formula. Since nesting is allowed, so is $\ell_1(\ell_2(p) = 1/2) = 1/2$. $\mathcal{L}_n^{QU}(\Phi)$ (the QU stands for *quantitative uncertainty*) is the language that results from starting with Φ and closing off under conjunction, negation, and the formation of likelihood formulas.

$\mathcal{L}_n^{QU}$ (as usual, I suppress Φ) is rich enough to express many notions of interest. For example, we can use obvious abbreviations such as

- $\ell_i(\varphi) - \ell_i(\psi) > b$ for $\ell_i(\varphi) + (-1)\ell_i(\psi) > b$,
- $\ell_i(\varphi) > \ell_i(\psi)$ for $\ell_i(\varphi) - \ell_i(\psi) > 0$,
- $\ell_i(\varphi) < \ell_i(\psi)$ for $\ell_i(\psi) - \ell_i(\varphi) > 0$,
- $\ell_i(\varphi) \leq b$ for $\neg(\ell_i(\varphi) > b)$,
- $\ell_i(\varphi) \geq b$ for $-\ell_i(\varphi) \leq -b$,
- $\ell_i(\varphi) = b$ for $(\ell_i(\varphi) \geq b) \wedge (\ell_i(\varphi) \leq b)$.

In addition, a formula such as $\ell_i(\varphi) \geq 1/3$ can be viewed as an abbreviation for $3\ell_i(\varphi) \geq 1$. That means we can use rational numbers in formulas, viewing the resulting formula as an abbreviation for the formula that would be obtained by clearing the denominators. We can also express simple conditional probabilities such as $\ell_i(\varphi|\psi) \geq 2/3$. Since $\ell_i(\varphi|\psi) = \ell_i(\varphi \wedge \psi)/\ell_i(\psi)$, we can clear the denominator again to get $3\ell_i(\varphi \wedge \psi) \geq 2\ell_i(\psi)$.

One other important notion we can express in the this language is that of the expected value of a random variable, provided that the worlds in which the random variable takes on a particular value can be characterized by formulas. For example, suppose that you win \$2 if a coin lands heads and lose \$3 if it lands tails. Then your expected winnings are $2\ell(\mathit{heads}) - 3\ell(\mathit{tails})$. The formula $2\ell(\mathit{heads}) - 3\ell(\mathit{tails}) \geq 1$ says that your expected winnings are at least \$1.

Although $\mathcal{L}_n^{QU}$ is a very expressive language, there are important notions that cannot be expressed. One example is independence. Suppose that we want to express the fact that (according to agent i) φ is independent of ψ. Informally, (after expanding and clearing the denominator) this corresponds to the formula $\ell_i(\varphi \wedge \psi) = \ell_i(\varphi) \times \ell_i(\psi)$. There is no difficulty giving semantics to such formulas in the semantic framework I am about to describe. However, this formula is not in the language, since I have not allowed multiplication of likelihood terms in linear likelihood formulas. I return to this issue in Section 6.2.

So why not include such formulas in the language? There is a tradeoff here: Enriching the language allows us to say more, but the added expressive power comes at a price. Richer languages are typically harder to axiomatize and it is typically harder to determine the validity of formulas in a richer language. (See the notes at the end of this section for further discussion of this issue and references.) Thus, I stick to the simpler language in this book, for purposes of illustration.

Formulas in $\mathcal{L}_n^{QU}$ are either true or false; they do not get "probabilistic" truth values. A logic for reasoning about probability can still be two-valued! To give semantics to formulas in $\mathcal{L}_n^{QU}$, I use probability structures. As usual, I start with the case where there is only one agent; thus, I omit the subscripts on ℓ. I also assume for now that all sets are measurable; that is, I restrict attention to the case of measurable probability structures. This makes life simpler. In a probability structure M, the term $\ell(\varphi)$ is supposed to be interpreted as the probability of the set $[\![\varphi]\!]_M$. But if this set is not measurable, we cannot talk about its probability. As long as all sets are measurable, this problem does not arise.

Defining the truth of formulas in a simple measurable probability structure $M = (W, w, \mu, \pi)$ is straightforward. The definition in the case of primitive propositions, conjunctions, and negations, is identical to that in Chapter 2. In the case of a likelihood formula,

$$(M, w) \models a_1\ell(\varphi_1) + \cdots + a_k\ell(\varphi_k) \geq b \text{ iff } a_1\mu([\![\varphi_1]\!]_M \cap W) + \cdots a_k\mu([\![\varphi_k]\!]_M \cap W) \geq b.$$

(Recall from Chapter 2 that $[\![\varphi]\!]_M = \{w : (M, w) \models \varphi\}$.) It is easy to check that, because the probability measure is independent of the world, the truth of a likelihood formula is also independent of the actual world; $(M, w) \models a_1\ell(\varphi_1) + \cdots + a_k\ell(\varphi_k) > b$ for some $w \in W$ if and only if $(M, w') \models a_1\ell(\varphi_1) + \cdots + a_k\ell(\varphi_k) > b$ for all $w' \in W$, so that we can write $M \models a_1\ell(\varphi_1) + \cdots + a_k\ell(\varphi_k) > b$.

Example 6.1.1 Suppose that $M_1 = (W, \mu, \pi)$, where $W = \{w_1, w_2, w_3, w_4\}$, $\mu(w_1) = \mu(w_2) = .25$, $\mu(w_3) = .3$, $\mu(w_4) = .2$, and π is such that $(M, w_1) \models p \land q$, $(M, w_2) \models p \land \neg q$, $(M, w_3) \models \neg p \land q$, and $(M, w_4) \models \neg p \land \neg q$. Thus, the worlds in W correspond to the four possible truth assignments to p and q. The structure M_1 is described in Figure ??.

It is straightforward to check that, for example,

$$(M_1, w_1) \models p \land q \land (\ell(p \land \neg q) > \ell(p \land q)).$$

Even though $p \land q$ is true at w_1, the agent considers $p \land \neg q$ to be more probable than $p \land q$. We also have

$$M_1 \models \ell(q | \neg p) = .6;$$

equivalently, sticking closer to the syntax of $\mathcal{L}^{QU}$,

$$M_1 \models \ell(\neg p \land q) - .6\ell(\neg p) = 0. \blacksquare$$

As usual, it is straightforward to extend simple probability structures to more general ones, where the probability may depend on the world and there are n agents. The definition of truth for likelihood formulas now uses probability assignments:

$$\begin{array}{l}(M, w) \models a_1 \ell_{i_1}(\varphi_1) + \cdots + a_k \ell_{i_k}(\varphi_k) \geq b \\ \quad \text{if } a_1 \mu_{w,i_1}([\![\varphi_1]\!]_M \cap W_{w,i_1}) + \cdots a_k \mu_{w,i_k}([\![\varphi_k]\!]_M \cap W_{w,i_k}) \geq b,\end{array}$$

where $\mathcal{PR}_{i_j}(w) = (W_{w,i_j}, \mu_{w,i_j})$.

Example 6.1.2 Let $M_2 = (W, \mathcal{PR}_1, \mathcal{PR}_2, \pi)$, where W and π are as in the structure M_1 in Example 6.1.1, $\mathcal{PR}_1(w_j) = (\{w_j\}, \mu_j)$, where μ_j is the unique probability measure on the singleton $\{w_j\}$ for $j = 1, \ldots, 4$, $\mathcal{PR}_2(w_1) = \mathcal{PR}_2(w_2) = (\{w_1, w_2\}, \mu_1')$, where $\mu_1'(w_1) = \mu_1'(w_2) = 1/2$, and $\mathcal{PR}_2(w_3) = \mathcal{PR}_2(w_4) = (\{w_3, w_4\}, \mu_2')$, where $\mu_2'(w_3) = \mu_2'(w_4) = 1/2$. Thus, for example,

$$\begin{array}{rl}(M, w_1) \models & p \land q \land (\ell_1(q) = 1) \land (\ell_2(q) = 1/2) \land (\ell_1(\ell_2(q) = 1/2) = 1) \land \\ & (\ell_2(\ell_1(q) = 1) = 1/2) \land (\ell_2(\ell_1(q) = 1 \lor \ell_1(\neg q) = 1) = 1).\end{array}$$

At world w_1, agent 1 is certain that q is true, while agent 2 thinks both q and $\neg q$ are equally likely. Moreover, agent 1 is certain that agent 2 thinks q has probability $1/2$, while agent 2 is certain that agent 1 assigns probability 1 to one of q or $\neg q$. $\blacksquare$

6.1.1 Axiomatizing Probabilistic Reasoning

I now present a complete axiomatization for reasoning about probability. The system, called AX_n^{prob}, divides nicely into three parts, which deal respectively with propositional reasoning, reasoning about linear inequalities, and reasoning about probability. It consists of the following axioms and rules of inference, which hold for $i = 1, \ldots, n$.

Propositional reasoning:

Prop. All substitution instances of tautologies of propositional calculus.

MP. From φ and $\varphi \Rightarrow \psi$ infer ψ (Modus ponens).

Reasoning about probability:

QU1. $\ell_i(\varphi) \geq 0$.

QU2. $\ell_i(true) = 1$.

QU3. $\ell_i(\varphi \wedge \psi) + \ell_i(\varphi \wedge \neg\psi) = \ell_i(\varphi)$.

QUGen. From $\varphi \Leftrightarrow \psi$ infer $\ell_i(\varphi) = \ell_i(\psi)$.

Reasoning about linear inequalities:

Ineq. All instances of valid formulas about linear inequalities (see below for details).

Prop and MP should be familiar from the systems K_n and $AX_{\gg}$ that we saw in Chapter 2 for reasoning about knowledge and relative likelihood. (However, note that Prop represents a different collection of axioms in each system, since the underlying language is different in each case. For example, $\neg(\ell_1(p) > 0 \wedge \neg(\ell_1(p) > 0)$ is an instance of Prop in AX_n^{prob}, which is based on the language $\mathcal{L}_n^{QU}$. It is not of Prop in K_n, since this formula is not even in the language $\mathcal{L}_n^K$.) Axioms QU1–QU3 correspond to the properties of probability: Every set gets nonnegative probability (QU1), the probability of the whole space is 1 (QU2), and finite additivity (QU3). The rule of inference QUGen is an analogue to the generalization rule Gen used in Chapter 2. The most obvious analogue—from φ infer $\ell_i(\varphi) = 1$—follows from QUGen and QU2, but is actually weaker than QUGen (and is not strong enough to give completeness).

The axiom Ineq consists of "all valid formulas about linear inequalities". To make this precise, let $\mathcal{X}$ be a a fixed infinite set of *variables*. An *inequality term* (over $\mathcal{X}$) is an expression of the form $a_1x_1 + \cdots + a_kx_k$, where $a_1, \ldots, a_k$ are integers, $x_1, \ldots, x_k$ are variables in $\mathcal{X}$, and $k \geq 1$. A *basic inequality* *formula* is a statement of the form $t \geq b$, where t is an inequality term and b is an integer. For example, $2x_3 + 7x_2 \geq 3$ is a basic inequality formula. An *inequality formula* is a Boolean combination of basic inequality formulas. I use f and g to refer to inequality formulas. An *assignment to variables* is a function A that assigns a real number to every variable. It is straightforward to define the truth of inequality formulas with respect to an assignment A to variables. For a basic inequality formula,

$$A \models a_1x_1 + \cdots + a_kx_k \geq b \text{ iff } a_1A(x_1) + \cdots + a_kA(x_k) \geq b.$$

The extension to arbitrary inequality formulas, which are just Boolean combinations of basic inequality formulas, is immediate:

$$\begin{array}{rll} A \models \neg f & \text{iff} & A \not\models f \\ A \models f \wedge g & \text{iff} & A \models f \text{ and } A \models g. \end{array}$$

As usual, an inequality formula f is *valid* if $A \models f$ for all A that are assignments to variables, and is *satisfiable* if $A \models f$ for some such A.

A typical valid inequality formula is

$$\begin{array}{l} (a_1x_1 + \cdots + a_kx_k \geq b) \wedge (a_1'x_1 + \cdots + a_k'x_k \geq b') \\ \quad \Rightarrow (a_1 + a_1')x_1 + \cdots + (a_k + a_k')x_k \geq (b + b'). \end{array} \tag{6.1}$$

To get an instance of Ineq, we simply replace each variable x_i that occurs in a valid formula about linear inequalities by a likelihood term $\ell(\varphi_i)$ (of course, each occurrence of the variable x_i must be replaced by the same primitive likelihood term $\ell(\varphi_i)$). Thus, the following weight formula, which results from replacing each occurrence of x_i in (6.1) by $\ell(\varphi_i)$, is an instance of Ineq:

$$\begin{array}{l} (a_1\ell(\varphi_1) + \cdots + a_k\ell(\varphi_k) \geq b) \wedge (a_1'\ell(\varphi_1) + \cdots + a_k'\ell(\varphi_k) \geq b') \\ \quad \Rightarrow (a_1 + a_1')\ell(\varphi_1) + \cdots + (a_k + a_k')\ell(\varphi_k) \geq (b + b'). \end{array} \tag{6.2}$$

There are a number of sound and complete axiomatizations for Boolean combinations of linear inequalities. The axiom Ineq takes it for granted that we have access to all the valid formulas of this logic, just as Prop takes it for granted that we have access to all valid propositional formulas.

The following result says that AX_n^{prob} completely captures probabilistic reasoning, to the extent that we can express it in the language $\mathcal{L}_n^{QU}$.

Theorem 6.1.3 *AX_n^{prob} is a sound and complete axiomatization with respect to $\mathcal{M}_n^{meas}$ for the language $\mathcal{L}_n^{QU}$.*

Proof Soundness is straightforward (Exercise 6.1). The completeness proof is beyond the scope of this book. ∎

As in the case of likelihood, if we restrict attention to simple measurable probability structures, we get an extra property, which characterizes the fact that the probability measure is independent of the world. This property is captured by the following analogue to L4:

QU4. $\varphi \Rightarrow (\ell_i(\varphi) = 1)$ if φ is an *i-likelihood formula* of the form $a_1\ell_i(\varphi_1) + \cdots + a_k\ell_i(\varphi_k) > b$ or the negation of an i-likelihood formula.

Theorem 6.1.4 *$AX_n^{prob} \cup \{QU4\}$ is a sound and complete axiomatization with respect to simple measurable probability structures for the language $\mathcal{L}^{QU}$.*

6.2 Reasoning about Independence

As we observed earlier, the language $\mathcal{L}_n^{QU}$ does not allow us to express independence. So what can we do if we want to reason about independence? There are three approaches that we can take to allow reasoning about independence.

One approach, which I already mentioned earlier, is to extend linear likelihood formulas to polynomial likelihood formulas, which allow multiplication of terms as well as addition. Thus, a typical polynomial likelihood formula is $a_1\ell_{i_1}(\varphi_1)\ell_2(\varphi_2)^2 - a_3\ell_{i_3}(\varphi_3) > b$. Let $\mathcal{L}_n^{QU,\times}$ be the language that extends $\mathcal{L}_n^{QU}$ by using polynomial likelihood formulas rather than just linear likelihood formulas. As I observed earlier, the fact that φ and ψ are independent (according to agent i) can be expressed in $\mathcal{L}_n^{QU,\times}$ as $\ell_i(\varphi \wedge \psi) = \ell_i(\varphi) \times \ell_i(\psi)$.

An advantage of using $\mathcal{L}_n^{QU,\times}$ to express independence is that it admits an elegant complete axiomatization with respect to $\mathcal{M}_n^{meas}$. In fact, the axiomatization is just AX_n^{prob}, with one small change—we just replace Ineq by

Ineq$^+$. All instances of valid formulas about polynomial inequalities.

That is, the effect of allowing polynomial inequalities rather than just linear inequalities is that we need to reason about polynomial inequalities rather than just linear inequalities. The axioms for reasoning about probability are unaffected. Let $\mathrm{AX}_n^{prob,\times}$ be the result of replacing Ineq by Ineq$^+$ in AX_n^{prob}.

Theorem 6.2.1 *$AX_n^{prob,\times}$ is a sound and complete axiomatization with respect to $\mathcal{M}_n^{meas}$ for the language $\mathcal{L}_n^{QU,\times}$.*

There is a price to be paid for using $\mathcal{L}_n^{QU,\times}$ though, as I hinted earlier: it seems to be harder to determine if formulas in this richer language are valid. There is another problem with using $\mathcal{L}_n^{QU,\times}$ as an approach for capturing reasoning about independence. It does not extend so readily to other notions of uncertainty. As I argued in Chapter 5, it is perhaps better to think of the independence of U and V being captured by the equation $\mu(U|V) = \mu(U)$ and $\mu(V|U) = \mu(V)$, rather than by the equation $\mu(U \cap V) = \mu(U) \times \mu(V)$. It is the former definition that generalizes more directly to other approaches.

This approach can be captured directly by extending $\mathcal{L}_n^{QU}$ in a different way, by allowing conditional likelihood terms of the form $\ell_i(\varphi|\psi)$ and linear combinations of such terms. Of course, in this extended language, we can express the independence of φ and ψ (according to agent i) as $(\ell_i(\varphi|\psi) = \ell_i(\varphi)) \wedge (\ell_i(\psi|\varphi) = \ell_i(\psi))$.

There is, however, a slight technical difficulty with this approach. Consider for simplicity a simple structure M. What is the truth value of a formula such as $\ell(\varphi|\psi) > b$ in M if $\mu([\![\psi]\!]_M = 0$? To some extent we can deal with this probem by considering *conditional probability structures*, where μ is a conditional probability measure, as defined in Section 5.1.1. But even if μ is a conditional probability measure, we need to deal with the case that $\mu([\![\varphi]\!]_M) = \emptyset$. Besides this technical problem, it is not clear how to axiomatize this extension of $\mathcal{L}_n^{QU}$ without allowing polynomial terms. In particular, it is not clear how to capture the fact that $\ell(\varphi|\psi) \times \ell(\psi) = \ell(\varphi \wedge \psi)$ if expressions of the form $\ell(\varphi|\psi) \times \ell(\psi)$ are allowed in the language. On the other hand, if multiplicative terms are allowed, then using the language $\mathcal{L}_n^{QU,\times}$ allows us to express independence without having to deal with the technical problem of giving semantics to formulas with terms of the form $\ell(\varphi|\psi)$ if $\mu([\![\psi]\!]_M) = 0$.

A third approach to reasoning about independence is just to add formulas directly to the language that talk about independence. Using the notation of Chapter 5, we can add formulas of the form $I(\psi_1, \psi_2|\varphi)$ or $I^{rv}(\psi_1, \psi_2|\varphi)$, with the obvious interpretation. Note that when viewed as random variables, a formula has only two possible values—true or false—so $I^{rv}(\psi_1, \psi_2|\varphi)$ is equivalent to $I(\psi_1, \psi_2|\varphi) \wedge I(\psi_1, \psi_2|\neg\varphi)$. Of course, the notation can extended as in Chapter 5 to allow sets of formulas to be arguments of I^{rv}.

I and I^{rv} inherit all the properties of the corresponding operators on events and random variables, respectively, considered in Chapter 5. In addition, if the language contains both facilities for talking about independence (via I or I^{rv}) and for talking about probability in terms of ℓ, there will in general be some interaction between the two. For example, $(\ell(p) = 1/2) \wedge (\ell(q) = 1/2) \wedge I(p, q|true) \Rightarrow \ell(p \wedge q)$ is certainly valid. No work has been done to date on getting axioms for such a combined language.

6.3 Reasoning about Lower Probability, Inner Measure, Belief, and Possibility

The language $\mathcal{L}_n^{QU}$ is appropriate not just for reasoning about probability, but also for reasoning about lower probability, inner measure, belief, and possibility. That is, we can interpret formulas in this language perfectly well in a number of different types of structures. All that changes is the class of structures considered and the interpretation of ℓ. Again, at the risk of boring the reader, I summarize the details here.

- In a lower probability structure $M = (W, \mathcal{LP}_1, \ldots, \mathcal{LP}_n, \pi)$,

  $$(M, w) \models a_1\ell_{i_1}(\varphi_1) + \cdots + a_k\ell_{i_k}(\varphi_k) \geq b$$
  $$\text{if } a_1(\mathcal{P}_{w,i_1})_*([\![\varphi_1]\!]_M \cap W_{w,i_1}) + \cdots + a_k(\mathcal{P}_{w,i_k})_*([\![\varphi_k]\!]_M \cap W_{w,i_k}) \geq b,$$

  where $\mathcal{LP}_i(w) = (W_{w,i}, \mathcal{P}_{w,i})$ for $i = 1, \ldots, n$. Thus, ℓ is now being interpreted as a lower probability. We could equally well interpret ℓ as an upper probability; either choice would work perfectly well.

- In a belief structure $M = (W, \mathcal{BEL}_1, \ldots, \mathcal{BEL}_n, \pi)$,

  $$(M, w) \models a_1\ell_{i_1}(\varphi_1) + \cdots + a_k\ell_{i_k}(\varphi_k) \geq b$$
  $$\text{if } a_1\mathrm{Bel}_{w,i_1}([\![\varphi_1]\!]_M \cap W_{w,i_1}) + \cdots + a_k\mathrm{Bel}_{w,i_k}([\![\varphi_k]\!]_M \cap W_{w,i_k}) \geq b,$$

  where $\mathcal{BEL}_i(w) = (W_{w,i}, \mathrm{Bel}_{w,i})$ for $i = 1, \ldots, n$.

- In the case of probability structure $M = (W, \mathcal{PR}_1, \ldots, \mathcal{PR}_n)$ where not all sets are necessary measurable, ℓ is interpreted as an inner measure, so

  $$(M, w) \models a_1\ell_{i_1}(\varphi_1) + \cdots + a_k\ell_{i_k}(\varphi_k) \geq b$$
  $$\text{if } a_1(\mu_{w,i_1})_*([\![\varphi_1]\!]_M \cap W_{w,i_1}) + \cdots + a_k(\mu_{w,i_k})_*([\![\varphi_k]\!]_M \cap W_{w,i_k}) \geq b,$$

  where $\mathcal{PR}_i(w) = (W_{w,i}, \mathcal{F}_{w,i}, \mu_{w,i})$ for $i = 1, \ldots, n$. This is a generalization of the measurable case; if all sets are in fact measurable, then the inner measure agrees with the measure. Again, we could equally well use outer measure here instead of inner measure.

- Finally, in a possibility structure $M = (W, \mathcal{POSS}_1, \ldots, \mathcal{POSS}_n, \pi)$, ℓ is interpreted as a possibility measure, so

  $$(M, w) \models a_1\ell_{i_1}(\varphi_1) + \cdots + a_k\ell_{i_k}(\varphi_k) \geq b$$
  $$\text{if } a_1\mathrm{Poss}_{w,i_1}([\![\varphi_1]\!]_M \cap W_{w,i_1}) + \cdots + a_k\mathrm{Poss}_{w,i_k}([\![\varphi_k]\!]_M \cap W_{w,i_k}) \geq b,$$

  where $\mathcal{POSS}_i(w) = (W_{w,i}, \mathrm{Poss}_{w,i})$ for $i = 1, \ldots, n$.

Can we characterize these notions of uncertainty axiomatically? Clearly all the axioms and inference rules other than QU3 (finite additivity) are still valid in $\mathcal{M}_n^{lp}$, $\mathcal{M}_n^{bel}$, and $\mathcal{M}_n^{prob}$. In the case of $\mathcal{M}_n^{bel}$, we have an obvious replacement for QU3: the analogues of B2 and B3.

QU5. $\ell_i(\mathit{false}) = 0$.

QU6. $\ell_i(\bigvee_{j=1}^n \varphi_j) \geq \sum_{j=1}^n \sum_{\{I:|I|=j\}} (-1)^{j+1} \ell_i(\wedge_{k \in I} \varphi_k)$.

It turns out that QU1, QU2, QU5, QU6, together with Prop, MP, and QUGen, gives a sound a complete axiomatization for reasoning about belief functions. This is not so surprising, since the key axioms just capture the properties of belief functions in an obvious way. What is perhaps more surprising is that these axioms also capture reasoning about inner measures. As I observed in Section 3.3, every inner measure is a belief function, but not every belief function is an inner measure. This suggests that, although inner measures satisfy the analogue of B3 (namely, Equation (3.6)), they may satisfy additional properties. In a precise sense, the following theorem shows they do not. Let AX_n^{bel} consist of QU1, QU2, QU5, QU6, QUGen, Prop, MP, and Ineq.

Theorem 6.3.1 *AX_n^{bel} is a sound and complete axiomatization with respect to both $\mathcal{M}_n^{bel}$ and $\mathcal{M}_n^{prob}$ for the language $\mathcal{L}_n^{QU}$.*

Proof Soundness is again straightforward (Exercise 6.2), and completeness is beyond the scope of this book. However, Exercises 6.3, 6.4, and 6.5 explain why the same axioms characterize belief structures and probability structures, even though not every belief function is an inner measure. ∎

In the case of simple belief structures or simple generalized probability structures, we also need QU4 to get completeness, just as with probability. I shall not belabor this point here.

For possibility structures, we need to replace QU6 by the key axiom that characterizes possibility, namely that the possibility of a union of two disjoint sets is the max of their individual possibilities. The following axiom does the job:

QU7. $(\ell_i(\varphi \wedge \psi) \geq \ell_i(\varphi \wedge \neg\psi)) \Rightarrow \ell_i(\varphi) = \ell_i(\varphi \wedge \psi)$.

Let AX_n^{poss} consist of QU1, QU2, QU5, QU7, QUGen, Prop, MP, and Ineq. As expected, we get

Theorem 6.3.2 *AX_n^{poss} is a sound and complete axiomatization with respect to $\mathcal{M}_n^{poss}$ for the language $\mathcal{L}_n^{QU}$.*

What about lower probability? As we have seen (Exercise 3.9), lower probabilities do not satisfy the analogue of B3. It follows that QU6 is not valid in $\mathcal{M}_n^{lp}$. All the other axioms in AX_n^{bel} are valid though (Exercise 6.6). Since lower probabilities are superadditive (Exercise 3.9), the following axiom is also valid in $\mathcal{M}_n^{lp}$

- $\ell_i(\varphi \wedge \psi) + \ell_i(\varphi \wedge \neg\psi) \leq \ell_i(\varphi)$.

However, superadditivity does not completely characterize lower probabilities. For example, the following property holds for lower probabilities:

$$\mathcal{P}_*(U \cup V) \geq \mathcal{P}_*(U) + \mathcal{P}_*(V) + \mathcal{P}_*(\overline{U \cap V}) - 1. \qquad (6.3)$$

To see this, note that if $\mathcal{P}$ is a set of probability measures, then for all $\mu \in \mathcal{P}$,

$$\mu(U \cup V) = \mu(U) + \mu(V) - \mu(U \cap V).$$

It immediately follows that

$$\mu(U \cup V) \geq \mathcal{P}_*(U) + \mathcal{P}_*(V) - \mathcal{P}^*(U \cap V).$$

And since $\mathcal{P}^*(U \cap V) = 1 - \mathcal{P}_*(\overline{U \cap V})$ (Exercise 3.9),

$$\mu(U \cup V) \geq \mathcal{P}_*(U) + \mathcal{P}_*(V) + \mathcal{P}_*(\overline{U \cap V}) - 1.$$

Equation (6.3) now follows immediately.

It is not hard to show that (6.3) does not follow from superadditivity (Exercise 6.7). We can use the inclusion-exclusion rule to generate other properties of lower probability as well; each of these properties corresponds to an axiom (Exercise 6.8). Do these properties (and the corresponding axioms) suffice to characterize lower probabilities? I believe that this remains an open question.

6.4 Reasoning about Relative Likelihood

The class $\mathcal{M}_n^{rank}$ of *ranking structures* can be defined just as all the other types of structures we have seen so far. The only reason that I am treating them separately here is that the language $\mathcal{L}_n^{QU}$ is simply inappropriate for reasoning about ranking. Ranks are always nonnegative integers; it is impossible to have, say, $\ell_i(\varphi) = 1/2$ if ℓ_i is to be interpreted as a ranking.

We could instead consider a variant of $\mathcal{L}_n^{QU}$ that just allows the coefficients in likelihood formulas to be in $I\!N^*$. There is no difficulty in obtaining a complete axiomatization for this language with respect to ranking structures. It is very similar to that for possibility structures, except that QU2 and QU5 needs to be changed to reflect the fact that for ranking structures, 0 and ∞ play the same role as 1 and 0 in possibility structures, and QU7 needs to be modified to use min rather than max. Rather than belaboring the details here, instead I consider what happens if we use another language altogether for reasoning about ranking—$\mathcal{L}_n^{\gg}$.

We can view $\mathcal{L}_n^{\gg}$ as a sublanguage of $\mathcal{L}_n^{QU}$, writing $\ell_i(\varphi) > \ell_i(\psi)$ rather than $\varphi \gg_i \psi$. For consistency with the earlier part of this chapter, I do that here. The semantics is just what we would expect. If $M = (W, w, \kappa, \pi)$ is a simple ranking structure, then

$$(M, w) \models \ell(\varphi) > \ell(\psi) \text{ if } \kappa([\![\varphi]\!]_M) < \kappa([\![\psi]\!]_M).$$

Note that greater likelihood corresponds to smaller rank.

For a ranking structure $M = (W, \mathcal{RANK}_1, \ldots, \mathcal{RANK}_n, \pi)$,

$$(M, w) \models \ell_i(\varphi) > \ell_i(\psi) \text{ if } \kappa_{w,i}([\![\varphi]\!]_M \cap W_{w,i}) < \kappa_{w,i}([\![\psi]\!]_M \cap W_{w,i}),$$

where $\mathcal{RANK}_i(w) = (W_{w,i}, \kappa_{w,i})$.

Of course, structures of all the types we have considered so far can be used to give semantics to formulas in $\mathcal{L}_n^{\gg}$ in the obvious way, using the fact that $\mathcal{L}_n^{\gg}$ can in fact be viewed as a special case of $\mathcal{L}_n^{QU}$. What about axiomatizations? As it happens, we have already seen all the relevant axioms in Chapter 2. As far as reasoning about relative likelihood goes, possibility measures and ranking functions are characterized by precisely the same axioms. Moreover, these are the axioms that characterize totally ordered relative likelihood; that is, the axiom system AX^M (or, equivalently, AX_n^{tot}) from Chapter 2 gives the desired sound and complete axiomatization. The extra structure of the real numbers (in the case of possibility measures) or $\mathbb{N}^*$ (in the case of ranking functions) plays no role if we are just considering statements of relative likelihood.

Theorem 6.4.1 *AX^M is a sound and complete axiomatization of the language $\mathcal{L}_n^{\gg}$ with respect to $\mathcal{M}_n^{rank}$ and $\mathcal{M}_n^{poss}$.*

Proof Soundness is straightforward, since possibility measures and ranking functions induce total orders. Completeness can be proved along the same lines as discussed in Exercise 2.24; details are left to the reader (Exercise 6.9). ∎

What happens when we interpret $\mathcal{L}_n^{\gg}$ in other types of structures? It is easy to see that in all the classes of structures we have considered, RL1, RL3, and RL6 are sound; $>$ is still bound to be irreflexive, orderly, and transitive. In $\mathcal{M}_n^{prob}$, $\mathcal{M}_n^{bel}$, and $\mathcal{M}_n^{lp}$, the ordering is modular, so RL5 holds as well. It does not hold in $\mathcal{M}_n^{plaus}$, since the domain of plausibility values may be partially ordered. RL2 and RL7 do not hold for probability, belief, lower probability, or plausibility either. They are not in general qualitative, nor do they satisfy the union property. Let AX_n^{ord} consist of RL1, RL3, RL6, MP, Gen.

Theorem 6.4.2

(a) AX_n^{ord} is a sound and complete axiomatization with respect to $\mathcal{M}_n^{plaus}$ for the language $\mathcal{L}_n^{\gg}$.

(b) $AX_n^{ord} \cup \{RL5\}$ is sound with respect to all of $\mathcal{M}_n^{meas}$, $\mathcal{M}_n^{prob}$, $\mathcal{M}_n^{bel}$, and $\mathcal{M}_n^{lp}$ for the language $\mathcal{L}_n^{\gg}$.

Proof Again, soundness is straightforward and completeness in part (a) can be proved along the same lines as discussed in Exercise 2.24; see Exercise 6.10. ∎

$\mathrm{AX}_n^{ord} \cup \{\mathrm{RL5}\}$ is *not* a complete axiomatization with respect to $\mathcal{M}_n^{meas}$, $\mathcal{M}_n^{prob}$, $\mathcal{M}_n^{bel}$, and $\mathcal{M}_n^{lp}$ for the language $\mathcal{L}_n^{\gg}$. For example, a formula such as $\ell(p) > \ell(\neg p)$ is true in a structure $M = (W, w, \mu, \pi) \in \mathcal{M}^{meas}$ iff $\mu([\![p]\!]_M) > 1/2$. Thus, the following formula is valid in $MMeas$:

$$(\ell(p) > \ell(\neg p) \wedge \ell(q) > \ell(\neg q)) \Rightarrow \ell(p) > \ell(\neg q).$$

However, this formula is not provable in AX^M, let alone $\mathrm{AX}_n^{ord} \cup \{RL5\}$ (Exercise 6.11). This example suggests that it will be difficult to find an elegant collection of axioms that is complete for $\mathcal{M}_n^{meas}$ with respect to $\mathcal{L}_n^{\gg}$. Even though this simple language does not have facilities for numeric reasoning, it is possible to express numeric properties. Other examples can be used to show that there are formulas valid in $\mathcal{M}_n^{prob}$, $\mathcal{M}_n^{bel}$, and $\mathcal{M}_n^{lp}$ that are not provable in AX^M (Exercise 6.12).

Exercises

6.1 Show that AX_n^{prob} is sound with respect to $\mathcal{M}_n^{meas}$ for the language $\mathcal{L}_n^{QU}$.

6.2 Show that AX_n^{bel} is sound with respect to both $\mathcal{M}_n^{bel}$ and $\mathcal{M}_n^{prob}$ for the language $\mathcal{L}_n^{QU}$.

6.3 Given a simple probability structure $M = (W, w, \mathcal{F}, \mu, \pi)$, define a simple belief structure $M' = (W, w, \mathrm{Bel}, \pi)$ by taking $\mathrm{Bel}(U) = \mu_*(U)$.

(Since every inner measure is a belief function, Bel is indeed a belief function.) Show that $(M, w') \models \varphi$ iff $(M', w') \models \varphi$, for every formula $\varphi \in \mathcal{L}^{QU}$ and $w' \in W \cup \{w\}$.

This shows that every formula satisfiable in a simple probability structure is satisfiable in a belief structure. The next exercise shows that the converse also holds. This explains why simple belief structures and simple probability structures are characterized by exactly the same axiom system. It is not hard to show that this result also holds for general (not just simple) belief structures and probability structures; this problem is left to Exercise 6.5.

6.4 Given a belief structure $M = (W, w_0, \mathrm{Bel}, \pi)$, define a probability structure $M' = (W', w_0, \mathcal{F}, \mu, \pi')$ as follows: Let $W' = \{(U, u) : U \subseteq W, u \in U\}$. For $U \subseteq W$, define $U^* = \{(U, u) : u \in U\}$. Note that $U^* \subseteq W'$. Moreover, if $U \neq V$, then U^* and V^* are disjoint; also, $W' = \cup_{U \subseteq W} U^*$. Take the sets $U^* for U \subseteq W$ to be a basis for $\mathcal{F}$. (That is, the sets in $\mathcal{F}$ consist of all possible unions of sets of the form U^*.) Let m be the mass function corresponding to Bel. Defin $\mu(U^*) = m(U)$, and extend μ to all the sets in $\mathcal{F}$ by finite additivity. Finally, define $\pi'(U, u) = \pi(u)$ for all $(U, u) \in W'$ and $\pi'(w_0) = \pi(w_0)$. Show that $(M, u) \models \varphi$ iff $(M', (U, u)) \models \varphi$ for all sets U such that $u \in U$ and that $(M, w_0) \models \varphi$ iff $(M', w_0) \models \varphi$.

Note that this result essentially says that a belief function on W can be viewed as the inner measure corresponding to a probability measure defined on W', at least as far as sets definable by formulas are concerned. That is, $\mathrm{Bel}([\![\varphi]\!]_M) = \mu_*([\![\varphi]\!]_{M'})$ for all formulas φ.

6.5 Extend the results of Exercises 6.4 and 6.5 to general (not just simple) belief structures and probability structures.

6.6 Show that all the axioms in AX_n^{bel} other than QU6 are valid in $\mathcal{M}_n^{lp}$.

6.7 Show that (6.3) does not follow from superadditivity by defining a set W and a function f associating with each subset of W a real number in $[0, 1]$ such that (a) $f(\emptyset) = 0$, (b) $f(W) = 1$, (c) f satisfies superadditivity (i.e., $f(U \cup V) \geq f(U) + f(V)$), but (d) f does not satisfy (6.3) (that is, there exist sets U and V such that $f(U \cup V) \geq f(U) + f(V) + f(U \cap V) - 1$).

6.8 Using the inclusion-exclusion rule, construct a sequence of axioms $IE_1, IE_2, IE_3, \ldots$ that are valid in $\mathcal{M}_n^{lp}$ such that IE_n is an inequality that holds for $\mathcal{P}_*(U_1 \cup \ldots \cup U_n)$ and IE_2 is (6.3).

* **6.9** Prove Theorem 6.4.1. (Hint: soundness is straightforward; for completeness, use the ideas of Exercise 2.24.)

* **6.10** Prove Theorem 6.4.2.

6.11 Show that $(\ell(p) > \ell(\neg p) \wedge \ell(q) > \ell(\neg q)) \Rightarrow \ell(p) > \ell(\neg q)$ is not provable in AX^M. (Hint: if it were provable in AX^M, it would be valid in $\mathcal{M}_n^{poss}$ and $\mathcal{M}_n^{rank}$.)

* **6.12** Show that there are formulas valid in $\mathcal{M}_n^{prob}$, $\mathcal{M}_n^{bel}$, and $\mathcal{M}_n^{lp}$ that are not provable in AX^M.

Notes

The logics $\mathcal{L}^{QU}$ and $\mathcal{L}^{QU,\times}$ were introduced in [Fagin, Halpern, and Megiddo 1990], and used for reasoning both about probability and belief functions. $\mathcal{L}^{QU}$ can be viewed as a formalization of Nilsson's [1986] *probabilistic logic*; it is a fragment of a propositional probabilistic *dynamic logic* introduced by Feldman [1984]. (Dynamic logic is a logic for reasoning about actions.) Theorems 6.1.3, 6.2.1, and 6.1.4 are proved in [Fagin, Halpern, and Megiddo 1990]. (Actually, slightly simplified versions of these theorems are proved, where only simple probability structures are considered and the syntax is restricted so that for a likelihood term of the form $\ell(\varphi)$, the formula φ is required to be a propositional formula. Nevertheless, the same basic proof techniques work for the slightly more general result stated here.) In addition, a finite collection of axioms is given that characterizes Ineq. Finally, the complexity of these logics is also considered, and it is shown that while the satisfiability problem for $\mathcal{L}^{QU}$ in simple probability structures (whether measurable or not) is *NP*-complete, no worse than that of propositional logic, but for $\mathcal{L}^{QU,\times}$, the complexity seems to go up to polynomial space (no although lower bound is known). On the other hand, results of [Fagin and Halpern 1994] show that once we consider general probability structures, the complexity of the satisfiability problem for both languages is *PSPACE*-complete.

See the notes in Chapter 5 for references regarding I and I^{rv}. Relatively little work has been done on logical characterizations of independence (in particular, the operator I).

Theorem 6.3.1 and Exercises 6.3, 6.4, and 6.5 were first proved by Fagin and me [1991b], using the results of [Fagin, Halpern, and Megiddo 1990]. Fariñas del Cerro and Herzig [1991] provide a complete axiomatization for $\mathcal{M}_n^{poss}$ similar in spirit to AX_n^{poss} (although their axiomatization is not quite complete as stated; see [Halpern 1997a]).