Chapter 11

First-order modal logic

Propositional logic is useful for modeling rather forms simple of reasoning, but there are many forms of reasoning that it is simply not expressive enough to capture. For example, propositional logic cannot talk about individuals, the properties they have, and relations between them. For example, suppose that Alice is American, but Bob is not. In a propositional logic, there could certainly be a primitive proposition p that is intended to express the fact "Alice is American", and another primitive proposition q to express that "Bob is American". The statement that Alice is American but Bob is not would then be expressed as $p \wedge \neg q$. But this way of expressing the statement somehow misses out on the fact that there is one property—being American—and two individuals, Alice and Bob, each of which may or may not possess the property.

First-order logic allows explicit reasoning about individuals and the properties they have. For example, in first-order logic, the fact that Alice is American and Bob could be expressed using a formula such as *American(Alice)* $\wedge$ $\neg$*American(Bob)*. This formla brings out the relationship between Alice and Bob more clearly. First-order logic can also express relations and functional connections between individuals. For example, the fact that Alice is taller than Bob can be expressed using a formula such as *Taller(Alice,Bob)*; the fact that Joe is the father of Sara can be expressed by a formula such as *Joe* = *Father(Sara)*. Finally, first-order logic can express the fact that *all* individuals have a certain property or that there is *some* individual that has a certain property by using a universal quantifier $\forall$, read "for all", and an existential quantifier $\exists$, read "there exists". For example, the formula $\exists x \forall y\, \mathit{Taller}(x, y)$ says that there is someone who is taller than everyone else; the formula $\forall x \forall y \forall z((\mathit{Taller}(x, y) \wedge \mathit{Taller}(y, z)) \Rightarrow \mathit{Taller}(x, z))$ says that the taller-than relation is transitive: if x is taller than

y and y is taller than z, then x is taller than z.

First-order modal logic combines first-order logic with modal operators. As with everything else we have looked at so far, new subtleties arise in the combination of first-order logic and modal logic that do not appear in propositional modal logic or first-order logic alone. In this section, I first review first-order logic and then consider a number of first-order modal logics.

11.1 First-Order Logic

The formal syntax of first-order logic is somewhat more complicated than that of propositional logic. The analogue in first-order logic of the set of primitive propositions is the *(first-order) vocabulary* $\mathcal{T}$, which consists of *relation symbols*, *function symbols*, and *constant symbols*. Each relation symbol and function symbol in $\mathcal{T}$ has some *arity*, that intuitively corresponds to the number of arguments it takes. If the arity is k, then the symbol is k-ary. In the earlier examples, *Alice* and *Bob* are constant symbols, *American* is a relation symbol of arity 1, *Taller* is a relation symbol of arity 2, and *Father* is a function symbol of arity 1. Because *American* is a relation symbol of arity 1, it does not make sense to write *American(Alice, Bob)*: *American* takes only one argument. Similarly, it does not make sense to write *Taller(Alice)*: *Taller* has arity 2 and takes two arguments. Intuitively, a relation symbol of arity 1 describes a property of an individual (is she an American or not?), a 2-ary relation symbol describes a relation between a pair of individuals, and so on. An example of a 3-ary relation symbol might be *Parents(a,b,c)*: a and b are the parents of c. (1-ary, 2-ary, and 3-ary relations are usually called *unary*, *binary*, and *ternary* relations, respectively, and similarly for functions.)

Besides the symbols in the vocabulary, there is an infinite supply of *variables*, which are usually denoted x and y, possibly with subscripts. Constant symbols and variables are both used to denote individuals. More complicated terms denoting individuals can be formed by using function symbols. Formally, the set of *terms* is formed by starting with variables and constant symbols, and closing off under function application, so that if f is a k-ary function symbol and t_1, ..., t_k are terms, then $f(t_1, \ldots, t_k)$ is a term. Terms are used in formulas. An *atomic formula* is either of the form $P(t_1, \ldots, t_k)$, where P is a k-ary relation symbol and $t_1, \ldots, t_k$ are terms, or of the form $t_1 = t_2$, where t_1 and t_2 are terms. Just as in propositional logic, more complicated formulas can be formed by closing off under negation and conjunction, so that if φ and ψ are formulas, then so are $\neg\varphi$ and $\varphi \wedge \psi$. But first-order logic is closed under one more feature:

quantification. If φ is a formula and x is a variable, then $\exists x\varphi$ is also a formula; $\forall x\varphi$ is an abbreviation for $\neg\exists x\neg\varphi$. Call the resulting language $\mathcal{L}^{fo}(\mathcal{T})$, or just $\mathcal{L}^{fo}$; just as in the propositional case, I often suppress the $\mathcal{T}$ if it does not play a significant role.

First-order logic can be used to reasoning about properties of addition and multiplication. The vocabulary of *number theory* consists of the binary functions symbols $+$ and $\times$, and the constant symbols 0 and 1. Examples of terms in this vocabulary are $1+(1+1)$ and $(1+1)\times(1+1)$. (Although I use infix notation, writing, for example, $1+1$ rather than $+(1,1)$, it should be clear that $+$ and $\times$ are binary function symbols.) The term denoting the sum of k 1's by k is abbreviated as k. Thus, typical formulas of number theory include $2+3=5$, $2+3=6$, $2+x=6$, and $\forall x, y(x+y=y+x)$. Clearly the first formula should be true, given the standard interpretation of the symbols, and the second to be false. It is not clear whether the third formula should be true or not, since the value of x is unknown. Finally, the fourth formula represents the fact that addition is commutative, so it should be true under the standard interpretation of these symbols. The following semantics captures these intuitions.

Semantics is given to first-order formulas using *relational structures*. Roughly speaking, a relational structure consists of a set of individuals, called the *domain* of the structure, and a way of associating with each of the elements of the vocabulary corresponding entities over the domain. Thus, a constant symbol is associated with an element of the domain, a function symbol is associated with a function on the domain, and so on. More precisely, fix a vocabulary $\mathcal{T}$. A *relational $\mathcal{T}$-structure* (sometimes simply called a relational structure or just a structure) $\mathcal{A}$ consists of a nonempty domain, denoted $\mathrm{dom}(\mathcal{A})$, an assignment of a k-ary relation $P^{\mathcal{A}} \subseteq \mathrm{dom}(\mathcal{A})^k$ to each k-ary relation symbol P of $\mathcal{T}$, an assignment of a k-ary function $f^{\mathcal{A}} : \mathrm{dom}(\mathcal{A})^k \rightarrow \mathrm{dom}(\mathcal{A})$ to each k-ary function symbol f of $\mathcal{T}$, and an assignment of a member $c^{\mathcal{A}}$ of the domain to each constant symbol c. $P^{\mathcal{A}}$, $f^{\mathcal{A}}$, and $c^{\mathcal{A}}$ are called the *denotations* of P, f, and c, respectively, in $\mathcal{A}$.

As an example, suppose that $\mathcal{T}$ consists of one binary relation symbol E. In that case, a $\mathcal{T}$-structure is simply a graph. (Recall that a graph consists of a set of nodes, some of which are connected by edges.) The domain is the set of nodes of the graph, and the interpretation of E is the edge relation of the graph, so that there is an edge from d_1 to d_2 exactly if $(d_1, d_2) \in E^{\mathcal{A}}$. For another example, consider the vocabulary of number theory discussed above. One relational structure for this vocabulary is the natural numbers, where 0, 1, $+$, and $\times$ get their standard interpretation. Another is the real numbers, where, again, all the symbols get their standard interpretation. Of course, there are many other relational structures over which these symbols can be interpreted.

A relational structure does not provide an interpretation of the variables. Technically, it turns out to be convenient to have a separate function that does this. A *valuation* V on a structure $\mathcal{A}$ is a function from variables to elements of $\mathrm{dom}(\mathcal{A})$. Recall that terms are intended to represent elements in the domain. Given a structure $\mathcal{A}$, a valuation V on $\mathcal{A}$ can be extended in a straightforward way to a function $V^{\mathcal{A}}$ (I typically omit the superscript $\mathcal{A}$ when it is clear from context) that maps terms to elements of $\mathrm{dom}(\mathcal{A})$, simply by defining $V^{\mathcal{A}}(c) = c^{\mathcal{A}}$ for each constant symbol c and then extending the definition by induction on structure to arbitrary terms, by taking $V^{\mathcal{A}}(f(t_1,\ldots,t_k)) = f^{\mathcal{A}}(V^{\mathcal{A}}(t_1),\ldots,V^{\mathcal{A}}(t_k))$.

I next want to define what it means for a formula to be true in a relational structure. Before I give the formal definition, consider a few examples. Suppose, as above, that *American* is a unary relation symbol, *Taller* is a binary relation symbol, and *Alice* and *Bob* are constant symbols. What does it mean for *American(Alice)* to be true in the structure $\mathcal{A}$? If the domain of $\mathcal{A}$ consists of people, then the interpretation $\mathit{American}^{\mathcal{A}}$ of the relation symbol *American* can be thought of as the set of all American people in $\mathrm{dom}(\mathcal{A})$. Thus *American(Alice)* should be true in $\mathcal{A}$ precisely if $\mathit{Alice}^{\mathcal{A}} \in \mathit{American}^{\mathcal{A}}$. Similarly, *Taller(Alice,Bob)* should be true if Alice is taller than Bob under the interpretation of *Taller* in $\mathcal{A}$; i.e., if $(\mathit{Alice}^{\mathcal{A}}, \mathit{Bob}^{\mathcal{A}}) \in \mathit{Taller}^{\mathcal{A}}$.

What about quantification? The English reading suggests that a formula such as $\forall x \mathit{American}(x)$ should be true in the structure $\mathcal{A}$ if every individual in $\mathit{dom}(\mathcal{A})$ is American, and $\exists x \mathit{American}(x)$ to be true if some individual in $\mathrm{dom}(\mathcal{A})$ is an American. The truth conditions will enforce this.

Recall that a structure does not give an interpretation to the variables. Thus, a structure $\mathcal{A}$ does not give us enough information to decide if a formula such as *Taller(Alice,x)* is true. That depends on the interpretation of x, which is given by a valuation. Thus, truth is defined relative to a pair $(\mathcal{A}, V)$ consisting of an interpretation and a valuation: *Taller(Alice,x)* is true in structure $\mathcal{A}$ under valuation V if $(V(\mathit{Alice}), V(x)) = (\mathit{Alice}^{\mathcal{A}}, V(x)) \in \mathit{Taller}^{\mathcal{A}}$.

As usual, the formal definition of truth in a structure $\mathcal{A}$ under valuation V proceeds by induction on the structure of formulas. If V is a valuation, x is a variable, and $d \in \mathrm{dom}(\mathcal{A})$, let $V[x/d]$ be the valuation V' such that $V'(y) = V(y)$ for every variable y except x, and $V'(x) = d$. Thus, $V[x/d]$ agrees with V except possibly on x and it assigns the value d to x.

$(\mathcal{A}, V) \models P(t_1,\ldots,t_k)$, where P is a k-ary relation symbol and $t_1,\ldots,t_k$ are terms, iff $(V(t_1),\ldots,V(t_k)) \in P^{\mathcal{A}}$,

$(\mathcal{A}, V) \models (t_1 = t_2)$, where t_1 and t_2 are terms, iff $V(t_1) = V(t_2)$,

$(\mathcal{A}, V) \models \neg\varphi$ iff $(\mathcal{A}, V) \not\models \varphi$,

$(\mathcal{A}, V) \models \varphi_1 \wedge \varphi_2$ iff $(\mathcal{A}, V) \models \varphi_1$ and $(\mathcal{A}, V) \models \varphi_2$,

$(\mathcal{A}, V) \models \exists x\varphi$ iff $(\mathcal{A}, V[x/d]) \models \varphi$ for some $d \in \mathrm{dom}(\mathcal{A})$.

Recall that $\forall x\varphi$ is an abbreviation for $\neg\exists x\neg\varphi$. It is easy to see that $(\mathcal{A}, V) \models \forall x\varphi$ iff $(\mathcal{A}, V[x/d]) \models \varphi$ for every $d \in \mathrm{dom}(\mathcal{A})$ (Exercise 11.1). $\forall$ essentially acts like an infinite conjunction. For suppose that $\psi(x)$ is a formula with whose only free variable is x; let $\psi(c)$ is the result of substituting c for x in ψ; i.e., $\psi(x)$ is $\psi[x/c]$. I sometimes abuse notation and write $(\mathcal{A}, V) \models \varphi(d)$ for $d \in \mathrm{dom}(\mathcal{A})$ rather than $(\mathcal{A}, V[x/d]) \models \varphi$. Abusing notation still further, note that $(\mathcal{A}, V) \models \forall x\varphi(x)$ iff $(\mathcal{A}, V) \models \wedge_{d\in D}\varphi(d)$. Thus, $\forall$ acts like an infinite conjunction. Similarly, $(\mathcal{A}, V) \models \exists x\varphi(x)$ iff $(\mathcal{A}, V) \models \vee_{d\in D}\varphi(d)$, so $\exists x$ acts like an infinite disjunction.

Returning to the examples in the language of number theory, let $I\!\!N$ be the natural numbers, with the standard interpretation of the symbols 0, 1, $+$, and $\times$. Then $(I\!\!N, V) \models 2 + 3 = 5$, $(I\!\!N, V) \not\models 2 + 3 = 6$, and $(I\!\!N, V) \models \forall x\forall y(x + y = y + x)$ for every valuation V, as expected. On the other hand, $(I\!\!N, V) \models 2 + x = 6$ iff $V(x) = 4$; here the truth of the formula depends on the valuation. Identical results hold if $I\!\!N$ is replaced by $I\!\!R$, the real numbers. On the other hand, let the formula φ be $\exists x(x \times x = 2)$, which says that 2 has a square root. Then $(I\!\!R, V) \models \varphi$ and $(I\!\!N, V) \not\models \varphi$ for all valuations V.

Notice that while the truth of the formula $2 + x = 6$ depends on the valuation, this is not the case for the other formulas. Variables were originally introduced as a crutch, as "placeholders" to describe what was being quantified. It would be useful to understand when they really are acting as placeholders. Essentially, this is the case when all the variables are "bound" by quantifiers. Thus, although the valuation is necessary in determining the truth of $2 + x = 6$, it is not necessary in determining the truth of $\exists x(2 + x = 6)$, because the x in $2 + x = 6$ is bound by the quantifier $\exists x$.

Roughly speaking, an occurrence of a variable x in φ is *bound* by the quantifier $\forall x$ in a formula such as $\forall x\varphi$ or by $\exists x$ in $\exists x\varphi$; an occurrence of a variable in a formula is *free* if it is not bound. (A formal definition of what it means for an occurrence of a variable to be free is given in Exercise 11.2.) A formula in which no occurrences of variables are free is called a *sentence*. Observe that x is free in the formula $\mathit{Taller}(c, x)$, but no variables are free in the the formulas *American(Alice)* and $\exists x\mathit{American}(x)$, so the latter two formulas are sentences. It is not hard to show that the valuation does not affect the truth of a sentence. That is, if φ is a sentence, and V and V' are valuations on the structure $\mathcal{A}$, then $(\mathcal{A}, V) \models \varphi$ iff $(\mathcal{A}, V') \models \varphi$

(Exercise 11.2). In other words, a sentence is true or false in a structure, independent of any valuation.

Satisfiability and validity for first-order logic can be defined in a manner analogous to propositional logic: a first-order formula φ is *valid* if $(\mathcal{A}, V) \models \varphi$ for all structures $\mathcal{A}$ and all valuations V and it is *satisfiable* if $(\mathcal{A}, V) \models \varphi$ for some structure $\mathcal{A}$ and some valuation V.

Just as in the propositional case, φ is valid if and only if $\neg\varphi$ is not satisfiable. There are well-known sound and complete axiomatizations of first-order logic as well. Describing the axioms requires a little notation. Suppose that φ is a first-order formula in which some occurrences of x are free. Say that a term t is *substitutable in* φ if there is no subformula of φ of the form $\exists y\psi$ such that the variable y occurs in t. If t is substitutable in φ, let $\varphi[x/t]$ be the result of substituting t for all free occurrences of x. (Note that if t were not substitutable in φ, then the y in t may be inadvertently bound in t by $\exists y$.) Let AX^{fo} consist of Prop and MP (for propositional reasoning) together with the following axioms and rules of inference.

F1. $\forall x(\varphi \Rightarrow \psi) \Rightarrow (\forall x\varphi \Rightarrow \forall x\psi)$.

F2. $\forall x\varphi \Rightarrow \varphi[x/t]$, where t is substitutable in φ.

F3. $\varphi \Rightarrow \forall x\varphi$ if x does not occur free in φ.

F4. $x = x$.

F5. $x = y \Rightarrow (\varphi \Rightarrow \varphi')$, where φ is a quantifier-free formula and φ' is obtained from φ by replacing zero or more occurrences of x in φ by y.

UGen. From φ infer $\forall x\varphi$.

F1, F2, and UGen can be viewed as analogues of K1, K2, and KGen, respectively, where $\forall x$ plays the role of K_i. This analogy can be pushed further; in particular, it follows from F3 that analogues of K4 and K5 hold for $\forall x$ (Exercise 11.4).

Theorem 11.1.1 AX^{fo} *is a sound and complete axiomatization of first-order logic with respect to relational structures.*

Proof Soundness is straightforward (Exercise 11.5); as usual, completeness is beyond the scope of this book. ∎

In the context of modal logic, I have restricted to finite sets of worlds. By and large, this was without loss of generality. In particular, as far as soundness and completeness results go (for example, Theorems 2.2.4, 6.2.1, 6.3.2, 6.4.1, and 7.4.3), there is no loss of generality in restricting to finite sets. There are *finite model theorems* that show that if a formula is satisfiable at all, then it is satisfiable in a structure with only finitely many worlds. Thus, there are no new axioms added by restricting to structures with only finitely many worlds.

In a similar spirit, we might hope that, without loss of generality (or, at least, without much loss of generality), we can restrict to finite domains in the case of first-order logic. While there is no loss of generality in restricting to *countable* domains (at least, as far as satisfiability and validity are concerned), restricting to finite domains results in new axioms, as the following example shows.

Example 11.1.2 Suppose that $\mathcal{T}$ consists of the constant symbol c and the unary function symbol f. Let φ be the following formula:

$$\forall xy(x \neq y \Rightarrow f(x) \neq f(y)) \land \forall x(f(x) \neq c).$$

The first conjunct says that f is one-to-one; the second says that c is not in the range of f. It is easy to see that φ is satisfiable in the natural numbers: take c to be 0 and f to be the successor function (so that $f(x) = x + 1$). However, f is not satisfiable in a relational structure with a finite domain. For suppose that $\mathcal{A} \models \varphi$ for some relational structure $\mathcal{A}$. (Since φ is a sentence, there is no need to mention the valuation.) An easy induction on k shows that $c^{\mathcal{A}}$, $f^{\mathcal{A}}(c^{\mathcal{A}})$, $f^{\mathcal{A}}(f^{\mathcal{A}}(c))$, $\ldots$, $(f^{\mathcal{A}})^k(c^{\mathcal{A}})$ must all be distinct (Exercise 11.6). Thus, $\mathrm{dom}(\mathcal{A})$ cannot be finite. It follows that $\neg\varphi$ is valid in relational structures with finite domains although it is not valid in all relational structures (and hence is not provable in AX^{fo}). ∎

Are there some reasonable axioms that can be added to AX^{fo} to obtain a complete axiomatization of first-order logic in finite relational structures? Somewhat surprisingly, the answer is no. The set of first-order formulas valid in finite structures is not *recursively enumerable*, that is, there is no program that will generate all and only the valid formulas. It follows that there cannot be be a finite (or even recursively enumerable) axiom system that is sound and complete for first-order logic over finite structures. Essentially this says that there is no easy way to characterize that finite domains in first-order logic. (By way of contrast, the set of formulas valid in all relational structures—finite or infinite—is recursively enumerable.)

If we further restrict to *bounded* domains (that is, relational structures whose domain has cardinality at most N, for some fixed N) then there is a complete axiomatization. The following axiom characterizes structures whose domains have cardinality at most N, in that it is true in a structure $\mathcal{A}$ iff $\mathrm{dom}(\mathcal{A})$ has cardinality at most N (Exercise 11.7).

FIN$_N$. $\exists x_1 \ldots x_N \forall y(y = x_1 \vee \ldots \vee y = x_N)$.

Let AX_N^{fo} be AX^{fo} together with FIN$_N$.

Theorem 11.1.3 AX_N^{fo} *is a sound and complete axiomatization of first-order logic with respect to relational structures whose domain has cardinality at most N.*

Proof Soundness is immediate from Exercise 11.5 and 11.7. Completeness is beyond the scope of this book (although it is in fact significantly easier to prove in the bounded case than in the unbounded case). ∎

Propositional logic can be viewed as a very limited fragment of first-order logic, one without quantification, using only unary relations, and mentioning only one constant. Consider the propositional language $\mathcal{L}^{Prop}(\Phi)$. Corresponding to Φ is the first-order vocabulary Φ^* consisting of a unary relation symbol p^* for every primitive proposition p in Φ and a constant symbol a. To every propositional formula φ in $\mathcal{L}^{Prop}(\Phi)$, there is a corresponding first-order formula φ^* over the vocabulary Φ^* that results by replacing occurrences of a primitive proposition p in φ by the formula $p^*(a)$. Thus, for example, $(p \wedge \neg q)^*$ is $p^*(a) \wedge \neg q^*(a)$. Intuitively, φ and φ^* express the same proposition. More formally, there is a mapping associating with each truth assignment v over Φ a relational structure $\mathcal{A}_v$ over Φ^*, where the domain of $\mathcal{A}_v$ consists of one element d, which is the interpretation of the constant symbol a, and

$$(p^*)^{\mathcal{A}_v} = \begin{cases} \{d\} & \text{if } v(p) = \mathbf{true}, \\ = \emptyset & \text{otherwise.} \end{cases}$$

Proposition 11.1.4 *For every propositional formula φ,*

(a) $v \models \varphi$ if and only if $\mathcal{A}_v \models \varphi^*$.

(b) φ is valid if and only if φ^* is valid.

(c) φ is satisfiable if and only if φ^* is satisfiable.

Proof See Exercise 11.8. ∎

Given that propositional logic is essentially a fragment of first-order logic, why is propositional logic of interest? Certainly, as a pedagogical matter, it is sometimes useful to focus on purely propositional formulas, without the overhead of functions, relations, and quantification. But there is a more significant reason. As I have said before (in Chapters 1 and 6, to be precise), increased expressive power comes at a price. For example, there is no algorithm for deciding whether a first-order formula is satisfiable. (Technically, this problem is undecidable.) It is easy to construct algorithms to check whether a propositional formula is satisfiable. (Technically, this problem is *NP-complete*, but that is much better than being undecidable!) If a problem can be modeled well using propositional logic, then it is worth sticking to propositional logic, rather than moving to first-order logic.

Not only can propositional logic be viewed as a fragment of first-order logic, but propositional epistemic logic can too (at least, as long as the language does not include common knowledge). Indeed, there is a translation of propositional epistemic logic that shows that, in a sense, the axioms for K_i can be viewed as consequences of the axioms for $\forall x$, although it is beyond the scope of this book to go into details.

Although first-order logic is more expressive than propositional logic, it is certainly far from the last word in expressive power. There are many ways in which it can be extended. One is to consider second- and higher-order logics. In first-order logic, there is quantification over individuals in the domain. In second-order logic, there is quantification over sets of individuals; in third-order logic, there is quantification over sets of sets of individuals, and so on. Second-order logic is very expressive. For example, the induction axiom can be expressed in second-order logic using the language of number theory. If x is a variable ranging over natural numbers (the individuals in the domain) and X is a variable ranging over sets of natural numbers, the induction axiom becomes

$$\forall X((0 \in X \wedge \forall x(x \in X \Rightarrow x+1 \in X)) \Rightarrow \forall x(x \in X)).$$

This says that any set that contains 0 and is closed under successor contains all the natural numbers.

Another way in which first-order logic can be extended is by allowing more general notions of quantification than just universal and existential quantifiers. For example, there can be a quantifier H standing for "at least half", so that a formula such as $Hx\varphi(x)$ is true (at least in a finite domain) if at least half the elements in the domain satisfy φ.

Yet a third way is to add modalities, just as in propositional logic. That is the focus of this chapter.

In this book, I do not consider second- and higher-order logics. Although they are very powerful, the increase in power is not that useful for reasoning about uncertainty. While I do not consider generalized quantifiers, the combination of modality and first-order quantification can occasionally be viewed as simulating new quantifiers. For example, in the first-order logic of probability discussed in Section 11.3, it is possible to express notions such as "at least half" (and much more general statistical notions).

11.2 First-Order Reasoning About Knowledge

The syntax for first-order epistemic logic is the obvious combination of the constructs of first-order logic—quantification, conjunction, and negation—and the modal operators $K_1, \ldots, K_n$. The semantics uses *relational Kripke structures*. In a (propositional) Kripke structure, each world is associated with a truth assignment to the primitive propositions via the interpretation π. In a relational Kripke structure, the π function associates with each world a relational structure. Formally, a relational Kripke structure for n agents over a vocabulary $\mathcal{T}$ is a tuple $(W, \mathcal{K}_1, \ldots, \mathcal{K}_n, \pi)$, where W is a set of worlds, π associates with each world in W a $\mathcal{T}$-structure (i.e., $\pi(w)$ is a $\mathcal{T}$-structure for each world $w \in W$), and $\mathcal{K}_i$ is a binary relation on W.

The semantics of first-order modal logic is, for the most part, the result of combining the semantics for first-order logic and the semantics for modal logic in a straightforward way. For example, a formula such as $K_i \mathit{American}(\mathit{President})$ is true at a world w if, in all worlds that agent i considers possible, the president is American. Note that this formula can be true even if agent i does not know who the president is. That is, there might be some world that agent i considers possible where the president is Bill, and another where the president is George. As long as the president is American in all these worlds, agent i knows that the president is American.

What about a formula such as $\exists x K_i \mathit{American}(x)$? It seems clear that this formula should be true if there is some individual in the domain at world w, say *Bill*, such that agent i knows that *Bill* is American. But now there is a problem: Although *Bill* may be a member of the domain of the relational structure $\pi(w)$, it is possible that *Bill* is not a member of the domain of $\pi(w')$ for some world w' that agent i considers possible at world w. There have been a number of solutions proposed to this problem that allow different domains at each world, but none of them are completely satisfactory (see the notes for references). For the purposes of this book, I avoid the problem by simply restricting to *common-domain Kripke structures*, that is, relational Kripke structures where the domain is the same at every world. To emphasize this point, I write the Kripke structure as $(W, D, \mathcal{K}_1, \ldots, \mathcal{K}_n, \pi)$, where D is the common domain used at each world, that is, $D = \mathrm{dom}(\pi(w))$ for all $w \in W$.

Under the restriction to common-domain structures, defining truth of formulas becomes quite straightforward. Fix a relational Kripke structure $M = (W, D, \mathcal{K}_1, \ldots, \mathcal{K}_n, \pi)$. A *valuation* V on M is a function that assigns to each variable a member of D. This means that $V(x)$ is independent of the world, although the interpretation of, say, a constant c does depend on the world. The definition of what it means for a formula φ to be true at a world w of M, given valuation V, now proceeds by the usual induction on structure. The clauses are exactly the same as those for first-order logic and propositional epistemic logic. For example,

$(M, w, V) \models P(t_1, \ldots, t_k)$, where P is a k-ary relation symbol and $t_1, \ldots, t_k$ are terms, iff $(V^{\pi(w)}(t_1), \ldots, V^{\pi(w)}(t_k)) \in P^{\pi(w)}$.

In the case of formulas $K_i\varphi$, the definition is just as in the propositional case:

$(M, w, V) \models K_i\varphi$ iff $(M, w', V) \models \varphi$ for all $w' \in \mathcal{K}_i(w)$.

First-order epistemic logic is more expressive than propositional epistemic logic. One important example of its extra expressive power is that it can distinguish between "knowing that" and "knowing who", by using the fact that variables denote the same individual in the domain at different worlds. For example, the formula $K_{Alice}\exists x(x = \mathit{President})$ says that Alice knows that someone is the president. This formula may be true in a given world where Alice does not know whether Bill or George is the president; she may consider one world possible where Bill is the president, and consider another world possible where George is the president. Therefore, although Alice knows that there is a president, she may not know exactly who the president is. The formula $\exists x K_{Alice}(x = \mathit{President})$ expresses the proposition that Alice knows who the president is. Because a valuation is independent of the world, it is easy to see that this formula says that there is one particular person who is president in every world that Alice considers possible.

What about axiomatizations? Suppose for simplicity that all the $\mathcal{K}_i$ relations are equivalence relations. In that case, the axioms K1–K5 of S5$_n$ are valid in relational Kripke structures. It might seem that by considering the first-order analogue of Prop (that is, allowing all substitution instances of first-order axioms), we should get a complete axiomatization. Unfortunately, this results in an unsound system: F2 is not sound.

Consider the following instance of F2:

$$\forall x \neg K_1(\mathit{Tall}(x)) \Rightarrow \neg K_1(\mathit{Tall}(\mathit{President})). \tag{11.1}$$

Now consider a relational Kripke structure $M = (W, D, \mathcal{K}_1, \pi)$, where

- W consists of two worlds, w_1 and w_2,
- D consists of two elements, d_1 and d_2,
- $\mathcal{K}_1(w_1) = \mathcal{K}_1(w_2) = W$,
- π is such that $\mathit{President}^{\pi(w_i)} = \{d_i\}$ and $\mathit{Tall}^{\pi(w_i)} = \{d_i\}$ for $i = 1, 2$.

Note that d_1 is not tall in w_2 and d_2 is not tall in w_1; thus, $(M, w_1) \models \forall x \neg K_1(\mathit{Tall}(x))$. On the other hand, the president is d_1 and is tall in w_1 and the president is d_2 and is tall in w_2; thus, $(M, w_1) \models K_1(\mathit{Tall}(\mathit{President}))$. It follows that (11.1) is not valid in structure M.

What is going on is that the valuation is independent of the world; hence, under a given valuation, a variable x is a *rigid designator*, that is, it denotes the same domain element in every world. On the other hand, a constant symbol such as *President* can denote different domain elements in distinct worlds. It is easy to see that F2 is valid if t is a variable. More generally, F2 is valid if t is a rigid designator (Exercise 11.9). This suggests that F2 can be salvaged by extending the definition of substitutable as follows. If φ is a first-order formula (with no occurrences of modal operators) then the definition of t being substitutable in φ is just that given above; if φ has some occurrences of modal operators, then t is substitutable in φ if t is a variable y such that there are no subformulas of the form $\exists y \psi$ in φ. With this extended definition, the hoped-for soundness and completeness result holds.

Theorem 11.2.1 $\mathrm{S5}_n$ *and* AX^{fo} *together provide a sound and complete axiomatization of first-order epistemic logic with respect to relational structures where the* $\mathcal{K}_i$ *relation is an equivalence relation.*

11.3 First-Order Reasoning About Probability

There is an obvious first-order extension of the propositional logic $\mathcal{L}^{QU}$ considered in Section 6.1. The syntax is the obvious combination of the syntax for first-order logic and that of $\mathcal{L}_n^{QU}$. I omit the formal definition. Call the resulting language $\mathcal{L}_n^{QU,fo}$. $\mathcal{L}_n^{QU,fo}$ includes formulas such as $\forall x(\ell_1(P(x)) \geq 1/2) \wedge \ell_2(\exists y Q(y)) < 1/3$—quantifiers can appear in the scope of likelihood formulas and likelihood formulas can appear in the scope of quantifiers.

Just as in Chapter 6, the likelihood (ℓ_i) can be interpreted as probability (if all sets are measurable), inner measure, lower probability, belief, or possibility, depending on the semantics. For example, in the case of probability, a structure has the form $(W, D, \mathcal{PR}_1, \ldots, \mathcal{PR}_n, \pi)$. (Note that, for the same reasons as in the case of knowledge, I am making the common-domain assumption.) Let $\mathcal{M}_n^{meas,fo}$ consist of all relational (measurable) probability structures. I leave the straightforward semantic definitions to the reader.

If this were all there was to it, this would be a very short section. However, consider the two statements "The probability that a randomly chosen bird will fly is greater than .9" and "The probability that Tweety (a particular bird) flies is greater than .9." There is no problem dealing with the second statement; it corresponds to the formula $\ell(\mathit{Flies}(\mathit{Tweety})) > .9$. (I am assuming that there is only one agent in the picture here, so I omit the subscript on ℓ.) But what about the first statement? What is the formula that should hold at a set of worlds whose probability is greater than .9?

The most obvious candidate is $\ell(\forall x(\mathit{Bird}(x) \Rightarrow \mathit{Flies}(x)) > .9$. However, it might very well be the case that in each of the worlds considered possible, there is at least one bird that doesn't fly. Hence, the statement $\forall x(\mathit{Bird}(x) \Rightarrow \mathit{Flies}(x))$ holds in none of the worlds (and so has probability 0); thus, $\ell(\forall x(\mathit{Bird}(x) \Rightarrow \mathit{Flies}(x)) > .9$ does not capture the first statement. What about $\forall x \ell(\mathit{Bird}(x) \Rightarrow \mathit{Flies}(x)) > .9$ or, perhaps better $\forall x(\ell(\mathit{Flies}(x) | \mathit{Bird}(x)) > .9)$? This runs into problems if there is a constant, say Opus, that represents an individual, say a penguin, that does not fly and is a rigid designator. Then $\ell(\mathit{Flies}(\mathit{Opus}) | \mathit{Bird}(\mathit{Opus})) = 0$, contradicting both $\forall x \ell(\mathit{Flies}(x) | \mathit{Bird}(x)) > .9$ and $\forall x(\ell(\mathit{Bird}(x) \Rightarrow \mathit{Flies}(x)) > .9)$. (It is important here that *Opus* is a rigid designator. The two statements $\forall x(\ell(\mathit{Flies}(x) | \mathit{Bird}(x)) > .9)$ and $\ell(\mathit{Flies}(\mathit{Opus}) | \mathit{Bird}(\mathit{Opus})) = 0$ are consistent if *Opus* is not a rigid designator; see Exercise 11.10.)

There seems to be a fundamental difference between these two statements. The first can be viewed as a statement about what one might expect as the result of performing some experiment or trial in a given situation. It can also be viewed as capturing statistical information about the world, since given some statistical information (say, that 90% of the individuals in a population have property P), then a randomly chosen individual should have probability .9 of having property P. By way of contrast, the second statement captures a *degree of belief*. The first statement seems to assume only one possible world (the "real" world), and in this world, some probability measure over the set of birds. It is saying that, with probability greater than .9, a bird chosen at random (according to this measure) will fly. The second statement implicitly assumes the existence of a number of possible worlds (in some of which Tweety flies, while in others Tweety doesn't), with some probability over these possibilities. Not surprisingly, the possible-worlds approach is well-suited to handling the second statement, but not the first.

It is not hard to design a language appropriate for statistical reasoning suitable for dealing with the first statement. The language includes terms of the form $||\varphi||_x$, which can be interpreted as "the probability that a randomly chosen x in the domain satisfies φ". This is analogous to terms such as $\ell(\varphi)$ in $\mathcal{L}^{QU}$. More generally, there can be an arbitrary set of variables in the subscript. To understand the need for this, suppose that the formula $Son(x, y)$ says that x is the son of y. Now consider the three terms $||Son(x, y)||_x$, $||Son(x, y)||_y$, and $||Son(x, y)||_{\{x,y\}}$. The first describes the probability that a randomly chosen x is the son of y; the second describes the probability that x is the son of a randomly chosen y; the third describes the probability that a randomly chosen pair (x, y) will have the property that x is the son of y. These three statements are all quite different. By allowing different sets of random variables in the subscript, they can all be expressed in the logic.

More formally, define a *statistical likelihood term* to have the form $||\varphi||_X$, where φ is a formula and X is a set of variables. A *(linear) statistical likelihood formula* is one of the form $a_1||\varphi_1||_{X_1} + \cdots + a_k||\varphi_k||_{X_k} > b$. Formulas are now formed just as in first-order logic, except that linear statistical likelihood formulas are allowed. In this language, the statement "The probability that a randomly chosen bird will fly is greater than .9" can easily be expressed. With some abuse of notation, it is just $||Flies(x)|Bird(x)||_x > .9$. (Without the abuse, it would be $||Flies(x)||_x > .9||Flies(x) \wedge Bird(x)||_x$ or $||Flies(x)||_x - .9||Flies(x) \wedge Bird(x)||_x > 0$.)

Quantifiers can be combined with statistical likelihood formulas. For example, $\forall x(||Son(x, y)||_y > .9)$ says that for all people, the probability that a randomly chosen person is their son is greater than .9; $\forall y(||Son(x, y)||_x > .9)$ says that for all people x, the probability that x is the son of a randomly chosen person is greater than .9. Let $\mathcal{L}^{QU,stat}$ be the language that results from combining the syntax of first-order logic with statistical likelihood formulas.

As with ℓ, statistical likelihood terms can be evaluated with respect to any quantitative representation of uncertainty. For definiteness, I use probability here.

A *statistical $\mathcal{T}$-structure* is a tuple $(\mathcal{A}, \mu)$, where $\mathcal{A}$ is a relational structure and μ is a probability measure on $\mathrm{dom}(\mathcal{A})$. To simplify matters, I assume that all subsets of $\mathrm{dom}(\mathcal{A})$ are measurable, that $\mathrm{dom}(\mathcal{A})$ is finite or countable, and that μ is countably additive. That means that μ is characterized by the probability it assigns to the elements of $\mathrm{dom}(\mathcal{A})$. Let $\mathcal{M}^{meas,stat}$ consist of all statistical $\mathcal{T}$-structures of this form, where all subsets of $\mathrm{dom}(\mathcal{A})$ are measurable.

Statistical structures should be contrasted with probability structures. In a statistical structure, there are no possible worlds and thus no probability on worlds. There is essentially only one world and the probability is on the domain. There is only one probability measure, not a different one for each agent. (It would be easy to allow a different probability measure for each agent, but the implicit assumption is that the probability in a statistical structure is objective, and does not represent the agent's degree of belief.) A very important special subclass of statistical structures (which is the focus of Chapter 12) are structures where the domain is finite and the probability measure is uniform (which makes all domain elements equally likely). This interpretation is particularly important for statistical reasoning. In that case, a formula such as $||Flies(x||_x|Bird(x)) > .9$ could be interpreted as "more than 90% of birds fly".

There are a number of reasons for not insisting that μ be uniform in general. For one thing, there are no uniform probability measures in countably infinite domains where all sets are measurable. (A uniform probability probability measure in a countably infinite domain would have to assign probability 0 to each individual element in the domain, which means by countable additivity it would have to assign probability 0 to the whole domain.) For another, once we consider other forms of uncertainty besides probability, there is not always an obvious analogue of uniform probability measures. (Consider plausibility measures, for example. What would uniformity mean there?) Finally, there are times when a perfectly reasonable way of making choices might not result in all domain elements being equally likely. For example, suppose that there are seven balls, four in one urn and three in another. If an urn is chosen at random and then a ball in the urn is chosen at random, not all the balls are equally likely. The balls in the urn with four balls have probability 1/8 of being chosen; the balls in the urn with three balls have probability 1/6 of being chosen. In any case, there is no extra difficulty in giving semantics in the case that μ is an arbitrary probability measure, so that is what I will do. On the other hand, it is probably best to think in terms of uniform measures in terms of understanding the intuitions.

One more construction is needed before giving the semantics. Given a probability function μ on D, there is a standard construction for defining the *product measure* on μ^n on the product domain D^n consisting of all n-tuples of elements of D: define $\mu^n(d_1, \ldots, d_n) = \mu(d_1) \times \ldots \times \mu(d_n)$. Note that if μ assigns equal probability to every element of D, then μ^n assigns equal probability to every element of D^n.

The semantic definitions are identical to those for first-order logic; the only new clause is that for statistical likelihood formulas. Given a statistical structure $M = (\mathcal{A}, \mu)$, a valuation V, a statistical likelihood term $||\varphi||_{\{x_1,\ldots,x_n\}}$, define

$$[||\varphi||_{\{x_1,\ldots,x_n\}}]_{M,V} = \mu^n((d_1, \ldots, d_n) : (M, V[x_1/d_1, \ldots, x_n/d_n]) \models \varphi\}).$$

That is, $[||\varphi||_{\{x_1,\ldots,x_n\}}]_{M,V}$ is the probability that a randomly chosen tuple

$(d_1, \ldots, d_n)$ (chosen according to μ^n) satisfies φ. Then define

$$(M, V) \models a_1||\varphi_1||_{X_1} + \cdots + a_k||\varphi_k||_{X_k} > b \text{ iff}$$
$$a_1[||\varphi_1||_{X_1}]_{M,V} + \cdots + a_k[||\varphi_k||_{X_k}]_{M,V} > b.$$

Note that the x in $||\varphi||_x$ acts in many ways just like the x in $\forall x$; for example, both bind free occurrences of x in φ and, in both cases, the x is a dummy variable. That is, $\forall x\varphi$ is equivalent to $\forall y\varphi[x/y]$ and $||\varphi||_x > b$ is equivalent to $||\varphi[x/y]||_y > b$ if y does not appear in φ (see Exercise 11.11). Indeed, $||\cdot||_x$ can express some of the general notions of quantification referred to in Section 11.1. For example, with a uniform probability measure and a finite domain, $||\varphi||_x > 1/2$ expresses the fact that at least half the elements in the domain satisfy φ, and thus is equivalent to the formula $Hx\varphi(x)$ from Section 11.1.

Of course, statistical reasoning and reasoning about degrees of belief can be combined, by having a structure with both a probability on the domain and a probability on possible worlds. The details are straightforward, so I omit them here.

What about axioms? First consider reasoning about degrees of belief. It is easy to see that F1–F5 are sound, as are QU1–QU3, QUGen, and Ineq from Section 6.1. They are, however, not complete. In fact, there is no complete axiomatization for the language $\mathcal{L}_n^{QU,fo}$ with respect to $\mathcal{M}_n^{meas,fo}$ (even if $n = 1$); the set of formulas in $\mathcal{L}_n^{QU,fo}$ valid with respect to $\mathcal{M}_n^{meas,fo}$ is not recursively enumerable. Restricting to finite domains does not help (since first-order logic restricted to finite domains is by itself not axiomatizable), nor does restricting to finite sets of worlds. But, as in the case of first-order logic, restricting to bounded domains does help.

Let $\mathrm{AX}_{n,N}^{prob,fo}$ consist of the axioms and inference rule of AX_N^{fo} together with those of AX_n^{prob} and one other axiom:

IV. $x \neq y \Rightarrow \ell_i(x \neq y) = 1.$

IV stands for *Inequality of Variables*. It is easy to see that IV is sound, as is the analogous property for equality

EV. $x = y \Rightarrow \ell_i(x = y);$

this just follows from the fact that variables are treated as rigid, and have the same value in all worlds. EV is provable from the other axioms, so it is not necessary to state it (Exercise 11.12). In fact, the analogues of IV and EV are *both* provable in the case of knowledge, which is why they do not appear in the axiomatization of Theorem 11.2.1 (Exercise 11.13).

Theorem 11.3.1 *$AX^{prob,fo}_{n,N}$ is a sound and complete axiomatization with respect to structures in $\mathcal{M}^{meas,fo}_n$ with a domain of cardinality at most N for the language $\mathcal{L}^{QU,fo}_n$.*

Proof Soundness is immediate from the soundness of AX^{fo}_N in relational structures of size at most N, the soundness of AX^{prob}_n in the propositional case, and the validity of EV, proved in Exercise 11.12. Completeness is beyond the scope of the book. ∎

Thus, there is a sense in which the axioms of first-order logic together with those for propositional reasoning about probability capture the essence of first-order reasoning about probability.

Much the same results hold for statistical reasoning. Consider the following axioms and rule of inference, where X ranges over finite sets of variables.

PD1. $||\varphi||_X \geq 0$, where X is a set of variables.

PD2. $\forall x_1 \ldots \forall x_n \varphi \Rightarrow ||\varphi||_{\{x_1,\ldots,x_n\}} = 1$.

PD3. $||\varphi \wedge \psi||_X + ||\varphi \wedge \neg\psi||_X = ||\varphi||_X$.

PD4. $||\varphi||_X = ||\varphi[x/z]||_{X[x/z]}$, where z is a variable that does not appear in X or φ.

PD5. $||\varphi \wedge \psi||_{X\cup Y} = ||\varphi||_X \times ||\psi||_Y$ if none of the free variables of φ is contained in Y, none of the free variables of ψ is contained in X, and X and Y are disjoint.

PDGen. From $\varphi \equiv \psi$ infer $||\varphi||_X = ||\psi||_X$.

PD1, PD3, and PDGen are the obvious analogues of QU1, QU3, and QUGen. respectively. PD2 is an extension of QU2. PD4 allows renaming of variables bound by "statistical" quantification. As we saw earlier, there is an analogous property for first-order logic, namely $\forall x\varphi \Rightarrow \forall y\varphi[x/y]$, which follows easily from F2 and F3 (Exercise 11.11). PD5 allows reasoning based on the independence of the random variables.

F1–F5 continue to be sound for statistical reasoning, except that the notion of substitutability in F2 must be modified to take into account that $||\cdot||_y$ acts like a quantifier, so that t not substitutable in φ if the variable y occurs in t and there is a term $||\cdot||_y$ in φ.

As in the case of degrees of belief, there is no complete axiomatization for the language $\mathcal{L}^{QU,stat}$ with respect to $\mathcal{M}^{meas,stat}$; the set of formulas in $\mathcal{L}^{QU,stat}$ valid with respect to $\mathcal{M}^{meas,stat}$ is not recursively enumerable. And again, while restricting to structures with finite domains does not help, restricting to bounded domains does. Let AX_N^{stat} consist of the axioms and inference rule of AX_N^{fo} together with PD1–PD5 and PDGen.

Theorem 11.3.2 *AX_N^{stat} is a sound and complete axiomatization with respect to structures in $\mathcal{M}^{meas,stat}$ with a domain of cardinality at most N for the language $\mathcal{L}^{QU,fo}$.*

11.4 First-Order Conditional Logic

In Section 7.4 we saw that conditional logic can be used to reason about defaults and counterfactuals. Moreover, we can give semantics to conditional logic using a number of different approaches, including possibility structures, ranking structures, PS structures (for sequences of probability structures), and preferential structures, all of which can be viewed as special cases of qualitative plausibility structures. All the different semantic approaches are characterized by the same axiom system, AX_n^{cond}, occasionally with C5 and C6 added, as appropriate.

These results suggest that all the different semantic approaches are essentially the same, at least as far as conditional logic is concerned. A more accurate statement would be that these approaches are the same as far as *propositional* conditional logic is concerned. Some significant differences start to emerge once the additional expressive power of first-order quantification is allowed.

Just as with probabilistic reasoning, for all these approaches, it is possible to consider a "degrees of belief" version, with some measure of likelihood over the possible worlds, and a "statistical" version, with some measure of likelihood on the domain. For the purposes of this section, I focus on the degrees of belief version. There are no new issues that arise for the statistical version, beyond those that already arise in the degrees of belief version.

Perhaps the most significant issue that emerges in first-order conditional logic is the importance of allowing structures with not only infinite domains but infinitely many possible worlds. Because the focus of the book thus far has been on structures with finitely many worlds, it is perhaps worth reviewing how all the different representations of likelihood that apply in default reasoning are affected by the move to infinitely many worlds.

- For probability, σ-algebras (which are closed under countable unions) and countably additive probability measures are typically considered. Countable adiitivity will not play a role in this section.

- For possibility, the possibility of an infinite set is taken to be the sup of the possibility of its elements. Thus, there may be no world $w \in W$

with $\mathrm{Poss}(w) = 1$ (as long as there worlds in W with possibility arbitrarily close to 1).

- No changes are necessary for ranking functions or plausibility measures to deal with infinitely many worlds. In particular, the notion of a qualitative plausibility measure is unchanged, and all the properties of qualitative plausibility measures discussed in Section 7.2 continue to hold even if W is infinite.

- For relative likelihood defined by a partial order $\succeq$ on worlds, there are two ways to go, as discussed in Section 2.3.1. In the first approach, a relational preferential structure has the form $(W, D, \mathcal{O}_1, \ldots, \mathcal{O}_n, \pi)$ and if $\mathcal{O}_i(w) = (W_{w,i}, \succeq_i)$, then there does not exist an infinite sequence of worlds $w_0, w_1, \ldots \in W_{w,i}$ such that $w_0 \prec_{w,i} w_1 \prec_{w,2} \prec \ldots$. In the second approach, such infinitely increasing sequences are allowed, but $\succ^s$ is defined as in Exercise 2.13. I consider both approaches in this section, although most of the technical details are left to exercises.

Let $\mathcal{M}_n^{qual,fo}$, $\mathcal{M}_n^{ps,fo}$, $\mathcal{M}_n^{poss,fo}$, $\mathcal{M}_n^{rank,fo}$, $\mathcal{M}_n^{pref,fin,fo}$, and $\mathcal{M}_n^{pref,fo}$ be the class of all relational qualitative plausibility structures, PS structures, possibility structures, ranking structures, preferential structures with no infinite increasing sequences, and preferential structures where $\succ^s$ is defined using the definition of Exercise 2.13, respectively, for n agents. Let $\mathcal{L}_n^{\rightarrow,fo}(\mathcal{T})$ be the obvious first-order analogue of the $\mathcal{L}_n^{\rightarrow}(\Phi)$.

Let's start with plausibility, where things work out quite nicely. Clearly the axioms of AX_n^{cond} and AX^{fo} are sound in $\mathcal{M}_n^{qual,fo}$. To get completeness, it is also necessary to include the analogue of IV. Let $N_i\varphi$ be an abbreviation for $\neg\varphi \rightarrow_i$ *false*. It is easy to show that if $M = (W, D, \mathcal{PL}_1, \ldots, \mathcal{PL}_n, \pi) \in \mathcal{M}_n^{qual,fo}$, then $(M, w) \models N_i\varphi$ iff $\mathrm{Pl}_{w,i}([\![\neg\varphi]\!]_M) = \perp$; i.e., $N_i\varphi$ asserts that the plausibility of $\neg\varphi$ is the same as that of the empty set, so that φ is true "almost everywhere" (Exercise 11.14). Thus, $N_i\varphi$ is the plausibilistic analogue of $\ell_i(\varphi) = 1$. Let $\mathrm{AX}^{cond,fo}$ consist of all the axioms and inference rules of AX^{cond} (for propositional reasoning about conditional logic) and AX^{fo}, together with the plausibilistic version of IV:

IVPl. $x \neq y \Rightarrow N_i(x \neq y)$.

The validity of IVPl in $\mathcal{M}^{qual,fo}$ follows from the fact that variables are rigid, just as in the case of probability (Exercise 11.15).

Theorem 11.4.1 $\mathrm{AX}_n^{cond,fo}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}_n^{qual,fo}$ *for the language* $\mathcal{L}_n^{\rightarrow,fo}$.

In the propositional case, adding C6 to AX_n^{cond} gives a sound and complete axiomatization of $\mathcal{L}_n^{\rightarrow}$ with respect to PS structures (Theorem 7.4.4). The analogous result holds in the first-order case.

Theorem 11.4.2 $\mathrm{AX}_n^{cond,fo} + \{C6\}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}_n^{ps,fo}$ *for the language* $\mathcal{L}_n^{\rightarrow,fo}$.

What about the other types of structures considered in Chapter 7? It turns out that more axioms besides C5 and C6 are required. To see why, consider the following example, known as the *lottery paradox*.

Example 11.4.3 Consider a lottery. Assume that the lottery is fair, so that any particular individual is highly unlikely to win, but someone is almost certainly guaranteed to win. Thus, the lottery has the following two properties:

$$\forall x(true \rightarrow \neg Winner(x)) \tag{11.2}$$

$$true \rightarrow \exists x\, Winner(x). \tag{11.3}$$

Let the formula *Lottery* be the conjunction of (11.2) and (11.3). (I am assuming here that there is only one agent doing the reasoning, so I drop the subscript 1 on $\rightarrow$.)

Lottery is satisfiable in $\mathcal{M}_1^{qual,fo}$. Define $M_{lot} = (W_{lot}, D_{lot}, \mathcal{PL}_{lot}, \pi_{lot})$ as follows:

- D_{lot} is a countable domain consisting of the individuals $d_1, d_2, d_3, \ldots$;
- W_{lot} consists of a countable number of worlds $w_1, w_2, w_3, \ldots$;
- $\mathcal{PL}_{lot}(w) = (W, \mathrm{Pl}_{lot})$, where Pl_{lot} gives the empty set plausibility 0, each non-empty finite set plausibility 1/2, and each infinite set plausibility 1;
- the denotation of *Winner* in world w_i according to π_{lot} (that is, $Winner^{\pi_{lot}(w_i)}$) is the singleton set $\{d_i\}$ (that is, in world w_i the lottery winner is individual d_i).

It is straightforward to check that Pl_{lot} is qualitative (Exercise 11.16). Abusing notation slightly, let $Winner(d_i)$ be the formula that is true if individual d_i wins. (Essentially, I am treating d_i as a constant in the language which denotes individual $d_i \in D_{lot}$ in all worlds.) By construction, $[\![\neg Winner(d_i)]\!]_{M_{lot}} = W - \{w_i\}$, so

$$\mathrm{Pl}_{lot}([\![\neg Winner(d_i)]\!]_{M_{lot}}) = 1 > 1/2 = \mathrm{Pl}([\![Winner(d_i)]\!]_{M_{lot}}).$$

That is, the plausibility of individual d_i losing is greater than the plausibility of individual d_i winning, for each $d_i \in D_{lot}$. Thus, M_{lot} satisfies (11.2). On the other hand, $[\![\exists x\, Winner(x)]\!]_{M_{lot}} = W$, so $\mathrm{Pl}_{lot}([\![\exists x\, Winner(x)]\!]_{M_{lot}}) > \mathrm{Pl}_{lot}([\![\neg\exists x\, Winner(x)]\!]_{M_{lot}})$; hence M_{lot} satisfies (11.3).

It is also possible to construct a relational PS structure (in fact, using the same set W_{lot} of worlds and the same interpretation π_{lot}) that satisfies *Lottery* (Exercise 11.17). On the other hand, there is there is no relational ranking structure that satisfies *Lottery*. To see this, suppose that $M = (W, D, \mathcal{RANK}, \pi)$ is a relational ranking structure such that $(M, w) \models$ *Lottery*. Suppose that $\mathcal{RANK}(w) = (W', \kappa)$. For each $d \in D$, let W_d be the subset of worlds in W' where d is the winner of the lottery; that is, $W_d = \{w \in W' : d \in Winner^{\pi(w)}\}$. It must be the case that $\kappa(W' - W_d) < \kappa(W_d)$ (i.e., $\kappa(W' \cap [\![\neg Winner(d)]\!]_M) < \kappa(W' \cap [\![Winner(d)]\!]_M)$), otherwise (11.2) would not be true at world w. Let w_0 be a world in W' such that $\kappa(w_0) = 0$. (There must be some world with this property; there may be more than one.) Clearly $w_0 \notin W_d$ for all $d \in D$, for otherwise $\kappa(W_d) = 0 \leq \kappa(W' - W_d)$. That means no individual d wins in w_0; that is, $Winner^{\pi(w_0)} = \emptyset$. Thus, $w_0 \in [\![\neg\exists x\, Winner(x)]\!]_M \cap W'$. But that means that

$$\kappa([\![\neg\exists x\, Winner(x)]\!]_M \cap W') \leq \kappa([\![\exists x\, Winner(x)]\!]_M \cap W'),$$

so $(M, w) \not\models true \rightarrow \exists x\, Winner(x)$. This contradicts the initial assumption that $(M, w) \models$ *Lottery*. Much the same argument shows that there is no nontrivial relational preferential structure in $\mathcal{M}_n^{pref,fin,fo}$ that satisfies *Lottery* (Exercise 11.18).

There is a relational possibility structure that satisfies *Lottery*. Consider the possibility structure $M_1 = (W_{lot}, D_{lot}, \mathcal{POSS}, \pi_{lot})$, where all the components besides $\mathcal{POSS}$ are just as in the plausibility structure M_{lot} and $\mathcal{POSS}(w) = (W, \mathrm{Poss})$, where $\mathrm{Poss}(w_i) = i/(i+1)$. This means that if $i > j$, then it is more possible that individual d_i wins than individual d_j. Moreover, this possibility approaches 1 as i increases. It is not hard to show that M_1 satisfies *Lottery* (Exercise 11.19). Poss corresponds to the ordering on worlds that has the infinitely increasing sequence $w_0 \prec w_1 \prec w_2 \prec \ldots$. There is also preferential structure in $\mathcal{M}_n^{pref,fo}$ that uses this ordering and satisfies *Lottery* (Exercise 11.20). ∎

Although *Lottery* is satisfiable in $\mathcal{M}_n^{poss,fo}$ and $\mathcal{M}_n^{pref,fo}$, a slight variant of it is not, as the following example shows.

Example 11.4.4 Consider a *crooked lottery*, where there is one individual who is more likely to win than the rest, but is still unlikely to win. This can be expressed using the following formula *Crooked*:

$$\begin{array}{l}\neg\exists x(\mathit{Winner}(x) \rightarrow \mathit{false})\wedge\\ \quad \exists y\forall x(x \neq y \Rightarrow ((\mathit{Winner}(x) \vee \mathit{Winner}(y)) \rightarrow \mathit{Winner}(y))).\end{array}$$

The first conjunct of *Crooked* states that each individual has some plausibility of winning; in the language of plausibility, this means that if $(M, w) \models$ *Crooked*, then $\mathrm{Pl}(W_w \cap [\![\mathit{Winner}(d)]\!]_M) > \perp$ for each domain element d. Roughly speaking, the second conjunct states that there is an individual who is at least as likely to win than the rest. More precisely, it says if $(M, w) \models$ *Crooked*, d^* is the individual guaranteed to exist by the second conjunct, and d is any other individual, then it must be the case that $\mathrm{Pl}(W_w \cap [\![\mathit{Winner}(d) \wedge \neg \mathit{Winner}(d^*)]\!]_M) < \mathrm{Pl}(W_w \cap [\![\mathit{Winner}(d^*)]\!]_M)$. This follows from the observation that if $(M, w) \models (\varphi \vee \psi) \rightarrow \psi$, then either $\mathrm{Pl}(W_w \cap [\![\varphi \vee \psi]\!]_M) = \perp$ (which cannot happen for the particular φ and ψ in the second conjunct because of the first conjunct of *Crooked*) or $\mathrm{Pl}(W_w \cap [\![\varphi \wedge \neg\psi]\!]_M) < \mathrm{Pl}(W_w \cap [\![\psi]\!]_M)$.

Take the crooked lottery to be formalized by the formula *Lottery* $\wedge$ *Crooked*. It is easy to model the crooked lottery using plausibility. Consider the structure $M'_{lot} = (W_{lot}, D_{lot}, \mathcal{PL}'_{lot}, \pi_{lot})$, which is identical to M_{lot} except that $\mathcal{PL}'_{lot}(w) = (W, \mathrm{Pl}'_{lot})$, where

- $\mathrm{Pl}'_{lot}(\emptyset) = 0$;
- if A is finite, then $\mathrm{Pl}'_{lot}(A) = 3/4$ if $w_1 \in A$ and $\mathrm{Pl}'_{lot}(A) = 1/2$ if $w_1 \notin A$;
- if A is infinite, then $\mathrm{Pl}_{lot}(A) = 1$.

It is easy to check that Pl'_{lot} is qualitative, that M'_{lot} satisfies *Crooked*, taking d_1 to be the special individual whose existence is guaranteed by the second conjunct (since $\mathrm{Pl}([\![\mathit{Winner}(d_1)]\!]_{M'_{lot}}) = 3/4 > 1/2 = \mathrm{Pl}([\![\mathit{Winner}(d_i) \cap \neg \mathit{Winner}(d_1)]\!]_{M'_{lot}})$ for $i > 1$), and that $\mathrm{Pl}'_{lot} \models$ *Lottery* (Exercise 11.21). On the other hand, *Lottery* $\wedge$ *Crooked* is not satisfiable in either $\mathcal{M}_n^{poss,fo}$ or $\mathcal{M}_n^{pref,fo}$ (Exercise 11.22). Intuitively, the problem in the case of possibility is that the possibility of d_1 winning has to be at least as great as that of d_i winning for $i \neq 1$, yet it must be less than 1. However, the possibility of *someone* winning must be one. This is impossible. A similar problem occurs in the case of preferential structures. ∎

Examples 11.4.3 and 11.4.4 show that $\mathrm{AX}_n^{cond,fo}$ (even with C5 and C6) is not a complete axiomatization for the language $\mathcal{L}_n^{\rightarrow,fo}$ with respect to any of $\mathcal{M}_n^{poss,fo}$, $\mathcal{M}_n^{rank,fo}$, $\mathcal{M}_n^{pref,fin,fo}$, or $\mathcal{M}_n^{pref,fo}$: $\neg$*Lottery* is a valid formula in $\mathcal{M}_1^{rank,fo}$ and $\mathcal{M}_1^{pref,fin,fo}$ and is not provable in $\mathrm{AX}_1^{cond,fo}$ even with C5 and C6 (if it were, it would be valid in plausibility structures that satisfy C5 and C6, which Example 11.4.3 show it is not); similarly $\neg(\mathit{Lottery} \wedge \mathit{Crooked})$ is a valid formula in $\mathcal{M}_1^{poss,fo}$ and $\mathcal{M}_1^{pref,fo}$ and is not provable in $\mathrm{AX}_1^{cond,fo}$, even with C5 and C6. These examples show that first-order conditional logic can distinguish these different representations of uncertainty although propositional conditional logic cannot.

Both the domain D_{lot} and the set W_{lot} of worlds in M_{lot} are infinite. This is not an accident. The formula *Lottery* is not satisfiable in any plausibility structure with either a finite domain or a finite set of worlds (or, more accurately, it is satisfiable in such a structure only if $\perp = \top$). This follows from the following more general result.

Proposition 11.4.5 *Suppose that $M = (W, D, \mathcal{PL}_1, \ldots, \mathcal{PL}_n, \pi)$ and either W or D is finite. If x does not appear free in ψ, then the following axiom is valid in M:*

C7. $\forall x(\psi \rightarrow_i \varphi(x)) \Rightarrow (\psi \rightarrow_i \forall x\varphi(x))$.

Proof Suppose that $(M, w) \models \forall x(\psi \rightarrow_i \varphi(x))$ and $\mathcal{PL}_i(w) = (W_{w,i}, \mathrm{Pl}_{w,i})$. Let $U_0 = W_{w,i} \cap [\![\psi]\!]_M$. If $\mathrm{Pl}_{w,i}(U_0) = \perp$, then clearly $(M, w) \models \psi \rightarrow_i \forall\varphi$, so suppose that $\mathrm{Pl}_{w,i}(U_0) > \perp$. If D is finite, suppose that $D = \{d_1, \ldots, d_n\}$. For each j, let $U_j = \{w' \in W_{w,i} : (M, w') \models \psi \wedge \varphi(d_j)\}$. Since $(M, w) \models \forall x(\psi \rightarrow_i \varphi)$, it must be the case, $\mathrm{Pl}_{w,i}(U_j) > \mathrm{Pl}_{w,i}(W_{w,i} \cap \overline{U_j})$ for all j. Using Pl4$'$ (which holds by Proposition 7.2.9, since $\mathrm{Pl}_{w,i}$ is qualitative, and so satisfies Pl4), an easy argument by induction shows that $\mathrm{Pl}(W_{w,i} \cap U_0 \cap U_1 \cap \ldots \cap U_n) > \mathrm{Pl}(W_{w,i} \cap U_0 \cap \overline{U_1 \cap \ldots \cap U_n})$ (Exercise 11.23). But $W_{w,i} \cap U_0 \cap U_1 \cap \ldots \cap U_n = [\![\psi \wedge \forall x\varphi]\!]_M \cap W_{w,i}$ and $W_{w,i} \cap U_0 \cap \overline{U_1 \cap \ldots \cap U_n} = [\![\psi \wedge \neg\forall x\varphi(x)]\!]_M \cap W_{w,i}$. Thus, $(M, w) \models \psi \Rightarrow \forall x\varphi(x)$.

The argument if W is finite is similar. For each $d \in D$, let $U_d = \{w' \in W_{w,i} : (M, w') \models \psi \wedge \varphi(d)\}$. Since W is finite, there can only be finitely many distinct sets U_d. Let $d_1, \ldots, d_n \in D$ be such that for all $d \in D$, $U_d = U_{d_j}$ for some $j \in \{1, \ldots, n\}$. Again from Pl4$'$ it follows that $\mathrm{Pl}(W_{w,i} \cap U_0 \cap U_{d_1} \cap \ldots \cap U_{d_n}) > \mathrm{Pl}(W_{w,i} \cap U_0 \cap \overline{U_{d_1} \cap \ldots \cap U_{d_n}})$. Since $[\![\varphi(d)]\!]_M = [\![\varphi(d_j)]\!]_M$ for some $j \in \{1, \ldots, n\}$, it follows that $W_{w,i} \cap U_0 \cap U_1 \cap \ldots \cap U_n = [\![\psi \wedge \forall x\varphi]\!]_M \cap W_{w,i}$ and $W_{w,i} \cap U_0 \cap \overline{U_1 \cap \ldots \cap U_n} = [\![\psi \wedge \neg\forall x\varphi(x)]\!]_M \cap W_{w,i}$. So, again, $(M, w) \models \psi \rightarrow_i \forall x\varphi(x)$. ∎

Corollary 11.4.6 *Suppose that $M = (W, D, \mathcal{PL}_1, \ldots, \mathcal{PL}_n, \pi)$ and either W or D is finite. Then $M \models \forall x(\mathit{true} \rightarrow \mathit{Winner}(x)) \Rightarrow \mathit{true} \rightarrow \forall x\neg\mathit{Winner}(x)$. Hence $M \models \mathit{Lottery} \Rightarrow (\mathit{true} \rightarrow \mathit{false})$.*

Proof It is immediate from Proposition 11.4.5 that $M \models \forall x(\mathit{true} \rightarrow \mathit{Winner}(x)) \Rightarrow \mathit{true} \rightarrow \forall x \neg \mathit{Winner}(x)$. Thus, if $(M, w) \models \mathit{Lottery}$, then $(M, w) \models \mathit{true} \rightarrow \forall x \neg \mathit{Winner}(x) \wedge \mathit{true} \rightarrow \exists x \, \mathit{Winner}(x)$. From the AND rule (C2) and right weakening (RC2), it follows that $(M, w) \models \mathit{true} \rightarrow \mathit{false}$. Thus, $M \models \mathit{Lottery} \Rightarrow (\mathit{true} \rightarrow \mathit{false})$. ∎

Corollary 11.4.6 shows that if W or D is finite, then if each person is unlikely to win the lottery, then it is unlikely than anyone will win. This is not the type of behavior we want. To avoid it (at least in the framework of plausibility measures and thus in all the other representations that can be used to model defaults, which can all be viewed as instances of qualitative plausibility measures), an infinite domain and an infinite number of possible worlds are both required. The structure M_{lot} shows that $\mathit{Lottery} \wedge \neg(\mathit{true} \rightarrow \mathit{false})$ is satisfiable in a structure with an infinite domain and an infinite set of worlds. In fact M_{lot} shows that $\forall x(\mathit{true} \rightarrow \neg \mathit{Winner}(x) \wedge \neg(\mathit{true} \rightarrow \forall x \neg \mathit{Winner}(x))$ is satisfiable.

Recall that in Section 8.3, it was shown that the definition of $B_i\varphi$ in terms of plausibility, as $\mathit{true} \rightarrow_i \varphi$ is equivalent to that in terms of a binary relation $\mathcal{B}_i$ (as in Section 2.2.1) *provided that the set of possible worlds is finite* (cf. Exercise 8.14). The lottery paradox shows that they are not equivalent with infinitely many worlds. It is not hard to show that B_i defined in terms of a $\mathcal{B}_i$ relation satisfies the property $\forall x B_i \varphi \Rightarrow B_i \forall x \varphi$ (Exercise 11.24). But under the identification of $B_i\varphi$ with $\mathit{true} \rightarrow_i \varphi$ this is precisely C7 which, as we have seen, does not hold in general.

C7 can be viewed as an instance of an infinitary AND rule since, roughly speaking, it says that if $\psi \rightarrow \varphi(d)$ holds for all $d \in D$, then $\psi \rightarrow \wedge_{d \in D} \varphi(d)$ holds. It was shown in Chapter 7 that Pl4 sufficed to give the (finitary) AND rule. There is a natural generalization of Pl4 that suffices for the infinitary version:

Pl4*. If $\{A_i : i \in I\}$ are pairwise disjoint sets, $A = \cup_{i \in I} A_i$, $0 \in I$, and for all $i \in I - \{0\}$, $\mathrm{Pl}(A - A_i) > \mathrm{Pl}(A_i)$, then $\mathrm{Pl}(A_0) > \mathrm{Pl}(A - A_0)$.

Recall that Pl4 states that if A_0, A_1, and A_2 are disjoint, $\mathrm{Pl}(A_0 \cup A_1) > \mathrm{Pl}(A_2)$, and $\mathrm{Pl}(A_0 \cup A_2) > \mathrm{Pl}(A_1)$, then $\mathrm{Pl}(A_0) > \mathrm{Pl}(A_1 \cup A_2)$. It is easy to check by induction that an analogous statement holds for any finite number of sets (Exercise 11.25). Pl4* asserts that a condition of this type holds even for an infinite collection of sets. This is not implied by Pl1–5. Consider the structure M_{lot} from Example 11.4.3. Take A_0 to be empty and take A_i, $i > 1$, to be the singleton consisting of the world w_i. Then $\mathrm{Pl}_{lot}(A - A_i) = 1 > 1/2 = \mathrm{Pl}_{lot}(A_i)$, but $\mathrm{Pl}_{lot}(A_0) = 0 < 1 = \mathrm{Pl}(\cup_{i>0} A_i)$. Hence, Pl4* does not hold for plausibility structures in general. It does, however, hold for certain subclasses:

Proposition 11.4.7 *Pl4* holds in every structure in* $\mathcal{M}_n^{rank,fo}$ *and* $\mathcal{M}_n^{pref,fin,fo}$.

Proof See Exercise 11.26. ∎

The following proposition tells us that C7 does indeed follow from Pl4*.

Proposition 11.4.8 *C7 is valid in all plausibility structures satisfying Pl4*.*

Proof See Exercise 11.27. ∎

Propositions 11.4.7 and 11.4.8 explain why the lottery paradox cannot be captured in $\mathcal{M}_n^{rank,fo}$ or $\mathcal{M}_n^{pref,fin,fo}$. Neither Pl4* nor C7 hold in general in $\mathcal{M}_n^{poss,fo}$ or $\mathcal{M}_n^{pref,fo}$. Indeed, the structure M_1 described in Example 11.4.3 and its analogue in $\mathcal{M}_1^{pref,fo}$ provide counterexamples (Exercise 11.28), which is why *Lottery* holds in these structures. So why is $\neg$(*Lottery* $\wedge$ *Crooked*) valid in $\mathcal{M}_1^{poss,fo}$ and $\mathcal{M}_1^{pref,fo}$? The following two properties of plausibility help to explain why. The first is a slightly weaker infinitary version of Pl4 than Pl4*; the second is an infinitary version of Pl5.

Pl4†. If $\{A_i : i \in I\}$ are pairwise disjoint sets, $A = \cup_{i\in I} A_i$, $0 \in I$, and for all $i \in I - \{0\}$, $\mathrm{Pl}(A_0) > \mathrm{Pl}(A_i)$, then $\mathrm{Pl}(A_0) \not< \mathrm{Pl}(A - A_0)$.

Pl5*. If $\{A_i : i \in I\}$ are sets such that $\mathrm{Pl}(A_i) = \perp$, then $\mathrm{Pl}(\cup_i A_i) = \perp$.

It is easy to see that Pl4† is implied by Pl4*. For suppose that Pl satisfies Pl4* and the preconditions of Pl4†. By Pl3, $\mathrm{Pl}(A_0) > \mathrm{Pl}(A_i)$ implies that $\mathrm{Pl}(A - A_i) > \mathrm{Pl}(A_i)$. By Pl4*, $\mathrm{Pl}(A_0) > \mathrm{Pl}(A - A_0)$ and therefore $\mathrm{Pl}(A_0) \not< \mathrm{Pl}(A - A_0)$, so Pl satisfies Pl4†. However, Pl4† can hold in structures that do not satisfy Pl4*. In fact, the following proposition shows that Pl4† holds in every structure in $\mathcal{M}_n^{poss,fo}$ and $\mathcal{M}_n^{pref,fo}$ (including the ones that satisfy *Lottery*, and hence do not satisfy Pl4*).

Proposition 11.4.9 *Pl4† holds in every relational structure structure in* $\mathcal{M}_n^{pref,fin,fo}$ *and* $\mathcal{M}_n^{poss,fo}$.

Proof See Exercise 11.29. ∎

Pl5* is an infinitary version of Pl5. It is easy to verify that it holds in all the approaches we consider, except plausibility measures and ϵ-semantics.

Proposition 11.4.10 *Pl5* holds in every plausibility structure in* $\mathcal{M}_n^{rank,fo}$, $\mathcal{M}_n^{poss,fo}$, $\mathcal{M}_n^{pref,fin,fo}$, *and* $\mathcal{M}_n^{pref,fo}$.

Proof Exercise 11.30. ∎.

Pl5* has elegant axiomatic consequences.

Proposition 11.4.11 *The axiom*

C8. $\forall x N_i\varphi \Rightarrow N_i(\forall x\varphi)$

is sound in qualitative plausibility structures satisfying Pl5; the axiom*

C9. $\forall x(\varphi(x) \rightarrow_i \psi) \Rightarrow ((\exists x\varphi(x)) \rightarrow_i \psi)$*, if* x *does not appear free in* ψ

is sound in structures satisfying Pl4 and Pl5*.*

Proof See Exercise 11.31. ∎

Axiom C9 can be viewed as an infinitary version of the OR Rule (C3), just as C7 can be viewed as an infinitary version of the AND Rule (C2). Abusing notation yet again, the antecedent says that $\wedge_{d\in D}(\varphi(d) \rightarrow_i \psi)$, while the conclusion says that $(\vee_{d\in D}\varphi(d)) \rightarrow_i \psi$.

When Pl4† and Pl5* hold, the crooked lottery is (almost) inconsistent.

Proposition 11.4.12 *The formula* $\mathit{Lottery} \wedge \mathit{Crooked} \Rightarrow (\mathit{true} \rightarrow \mathit{false})$ *is valid in structures satisfying Pl4*† *and Pl5*.*

Proof See Exercise 11.32. ∎

Since Pl4† and Pl5* are valid in $\mathcal{M}_n^{poss,fo}$, it immediately follows that $\mathit{Lottery} \wedge \mathit{Crooked}$ is unsatisfiable in $\mathcal{M}_n^{poss,fo}$.

To summarize, this discussion shows that once we move from propositional logic to first-order conditional logic, significant differences arise between approaches that were shown to be equivalent in the propositional case. This vindicates the intuition that there are significant differences between these approaches; the propositional language is simply too weak to capture them. Using plausibility makes it possible to delineate the key properties that distinguish the various approaches, properties such as Pl*, Pl4†, and Pl5*, which manifest themselves in axioms such as C7, C8, and C9.

Conditional logic was introduced in Section 7.4 as a tool for reasoning about defaults. Does the preceding analysis have anything to say about default reasoning? For that matter, how should defaults even be captured in first-order conditional logic? Statements like "birds typically fly" are similar in spirit to statements like "90% of birds fly". Using $\forall x(\mathit{Bird}(x) \rightarrow \mathit{Flies}(x))$ to represent this formula is just as inappropriate as using $\forall x(\ell(\mathit{Flies}(x)|\mathit{Bird}(x)) > .9)$ to represent "90% of birds fly". We saw earlier that the latter statement is perhaps best represented statistically, using a probability on the domain, not a probability on possible worlds. Similarly, it seems that "birds typically fly" should be represented using statistical plausibility. On the other hand, conclusions about individual birds (such as "Tweety is a bird, so Tweety (by default) flies") are similar in spirit to statements like "The probability that Tweety (a particular bird) flies is greater than .9"; these are best represented using plausibility on possible worlds.

Drawing the conclusion "Tweety flies" from "birds typically fly" would then require some way of connecting statistical plausibility with plausibility on possible worlds. There are no techniques given in this chapter for doing that; that is the subject of Chapter 12.

Exercises

11.1 Show that $(\mathcal{A}, V) \models \forall x\varphi$ iff $(\mathcal{A}, V[x/d]) \models \varphi$ for every $d \in \mathrm{dom}(\mathcal{A})$

11.2 Inductively define what it means for an occurrence of a variable x to be free in a first-order formula as follows:

- if φ is an atomic formula ($P(t_1, \ldots, t_k)$ or $(t_1 = t_2)$ then every occurrence of x in φ is free,
- an occurrence of x is free in $\neg\varphi$ iff the corresponding occurrence of x is free in φ,
- an occurrence of x is free in $\varphi_1 \wedge \varphi_2$ iff the corresponding occurrence of x in φ_1 or φ_2 is free,
- an occurrence of x is free in $\exists y\varphi$ iff the corresponding occurrence of x is free in φ and x is different from y.

A *sentence* is a formula in which no occurrences of variables are free.

(a) Show that if φ is a formula and V and V' are valuations that agree on all of the variables that are free in φ, then $(\mathcal{A}, V) \models \varphi$ iff $(\mathcal{A}, V') \models \varphi$.

(b) Show that if φ is a sentence and V and V' are valuations on the structure $\mathcal{A}$, then $(\mathcal{A}, V) \models \varphi$ iff $(\mathcal{A}, V') \models \varphi$.

11.3 Show that if all the symbols in the formula φ are contained in $\mathcal{T}' \subseteq \mathcal{T}$ and $\mathcal{A}$ and $\mathcal{A}'$ are two relational $\mathcal{T}$-structures such that $\mathrm{dom}(\mathcal{A}) = \mathrm{dom}(\mathcal{A}')$ and $\mathcal{A}$ and $\mathcal{A}'$ agree on the denotations of all the symbols in $\mathcal{T}'$, then $(\mathcal{A}, V) \models \varphi$ iff $(\mathcal{A}', V) \models \varphi$.

11.4 Show that the following two formulas, which are the analogues of K4 and K5 for $\forall x$, are valid in relational structures:

$$\forall x\varphi \Rightarrow \forall x\forall x\varphi$$
$$\exists x\varphi \Rightarrow \forall x\exists x\varphi.$$

11.5 Show that all the axioms of AX^{fo} are valid in relational structures and that UGen preserves validity.

11.6 Show that the domain elements $c^{\mathcal{A}}$, $f^{\mathcal{A}}(c^{\mathcal{A}})$, $f^{\mathcal{A}}(f^{\mathcal{A}}(c))$, ..., $(f^{\mathcal{A}})^k(c^{\mathcal{A}})$ defined in Example 11.1.2 must all be distinct.

11.7 Show that $\mathcal{A} \models \mathrm{FIN}_N$ iff $|\mathrm{dom}(\mathcal{A})| \le N$.

* **11.8** Prove Proposition 11.1.4.

11.9 Show that F2 is valid if t is a rigid designator.

11.10 Show that

$$\forall x(\ell(\mathit{Flies}(x)|\mathit{Bird}(x)) > .9) \land \ell(\mathit{Flies}(\mathit{Opus})|\mathit{Bird}(\mathit{Opus})) = 0$$

is satisfiable if *Opus* is not a rigid designator.

11.11 Show that

$$\forall x\varphi \Rightarrow \forall y\varphi[x/y], \text{ if } y \text{ does not appear in } \varphi$$

is provable in AX^{fo}.

11.12 (a) Show that IV and EV are valid in $\mathcal{M}_n^{meas,fo}$.

(b) Show that EV is provable in $\mathrm{AX}_{n,N}^{prob,fo}$. (Hint: use QU2, F4, QUGen, and F2.)

* **11.13** State analogues of IV and EV for knowledge and show that they are both provable using the axioms of $\mathrm{S5}_n$. (Hint: the argument for EV is similar in spirit to that for probability given in Exercise 11.12(b). For IV, use EV and K5, and show that $\neg K\neg K\varphi \Leftrightarrow K\varphi$ is provable in $\mathrm{S5}_n$.)

11.14 Show that if $M \in \mathcal{M}^{qual,fo}$, then $(M,w) \models N_i\varphi$ iff $\mathrm{Pl}_{w,i}([\![\neg\varphi]\!]_M) = \bot$.

11.15 Show that every instance of IVPl is valid in $\mathcal{M}^{qual,fo}$.

11.16 Show that the plausibility measure Pl_{lot} constructed in Example 11.4.3 is qualitative.

11.17 Construct a relational PS structure that satisfies *Lottery*.

11.18 Show that if $M \in \mathcal{M}_n^{pref,fin,fo}$ then $M \models \mathit{Lottery} \Rightarrow (\mathit{true} \rightarrow \mathit{false})$. That is, the only way *Lottery* can be true in a world $w \in W$ is if $W_w = \emptyset$.

11.19 Show that the relational possibility structure M_1 constructed in Example 11.4.3 satisfies *Lottery*.

11.20 Show that there is a relational preferential structure $M = (W_{lot}, D_{lot}, \mathcal{O}_1, \pi) \in \mathcal{M}_n^{pref,fo}$ such that $M \models \mathit{Lottery}$ where $\mathcal{O}_1(w) = (W, \prec)$ and $w_0 \prec w_1 \prec w_2 \prec \ldots$.

11.21 Show that the plausibility measure Pl'_{lot} constructed in Example 11.4.12 is qualitative and that $M'_{lot} \models \mathit{Lottery} \wedge \mathit{Crooked}$.

11.22 Show that $\mathit{Crooked} \wedge \mathit{Lottery}$ is not satisfiable in either $\mathcal{M}_1^{poss,fo}$ or $\mathcal{M}_1^{pref,fo}$.

11.23 Show by induction on n that Pl4 implies the generalization of Pl4′ to n sets. More precisely, show that if Pl satisfies Pl4 and $\mathrm{Pl}(U_0 \cap U_i) > \mathrm{Pl}(U_0 \cap \overline{U_i})$ for $i = 1, \ldots, n$, then

$$\mathrm{Pl}(U_0 \cap U_1 \cap \ldots \cap U_n) > \mathrm{Pl}(U_0 \cap \overline{U_1 \cap \ldots \cap U_n}).$$

(You may use Proposition 7.2.9.)

11.24 Show that $\forall x K_i \varphi \Rightarrow K_i \forall x \varphi$ is valid in relational epistemic structures.

11.25 Show by induction on n that if the plausibility measure Pl satisfies Pl4, $A_0, \ldots, A_n$ are pairwise disjoint, $A = \cup_{i=1}^n A_i$, and $\mathrm{Pl}(A - A_i) > \mathrm{Pl}(A_i)$ for $i = 1, \ldots, n$, then $\mathrm{Pl}(A_0) > \mathrm{Pl}(\cup_{i=1}^n A_i)$.

11.26 Prove Proposition 11.4.7.

11.27 Prove Proposition 11.4.8.

11.28 Show that the structure M_1 described in Example 11.4.3 and its analogue in $\mathcal{M}_1^{pref,fo}$ satisfy neither Pl^* nor C7.

11.29 Prove Proposition 11.4.9.

11.30 Prove Proposition 11.4.10. Also show that Pl5* does not necessarily hold in structures in $\mathcal{M}^{qual,fo}$ and $\mathcal{M}_n^{ps,fo}$.

11.31 Prove Proposition 11.4.11.

11.32 Prove Proposition 11.4.12.

Notes

The discussion of first-order logic here is largely taken from [Fagin, Halpern, Moses, and Vardi 1995], which in turn is based on that of Enderton [1972]. The axiomatization of first-order logic given here is essentially that given by Enderton, who also proves completeness. A discussion of generalized quantifiers can be found in [Ebbinghaus 1985]. Trakhtenbrot [1950] proved that the set of first-order formulas valid in finite relational structures is not recursively enumerable (from which it follows that there is no complete axiomatization for first-order logic over finite structures). The fact that there is a translation from propositional epistemic logic to first-order logic, as mentioned in Section 11.1, seems to have been observed independently by a number of people. The first treatment of these ideas in print seems to be due to van Benthem [1974]; details and further discussion can be found in his book [1985]. Finite model theorems are standard in the propositional modal logic literature; they are proved for epistemic logic in [Halpern and Moses 1992], for the logic of probability in [Fagin, Halpern, and Megiddo 1990], and for conditional logic in [Friedman and Halpern 1994b].

The discussion in Section 11.2 on first-order reasoning about knowledge is also largely taken from [Fagin, Halpern, Moses, and Vardi 1995]. Garson [Garson 1984] discusses in detail a number of ways of dealing with what is called the problem of "quantifying-in": how to give semantics to a formula such as $\exists x K_i(P(x))$ without the common domain assumption. The distinction between "knowing that" and "knowing who" is related to an old and somewhat murky distinction between knowledge *de dicto* (literally, "knowledge of words") and knowledge *de re* (literally, "knowledge of things"). See Plantinga [Plantinga 1974] for a discussion.

Section 11.3 on first-order reasoning about probability is largely taken from [Halpern 1990a], including the discussion of the distinction between the two interpretations of probability (the statistical interpretation and the degree of belief interpretation), the axiom systems $\mathrm{AX}_{n,N}^{prob,fo}$ and AX_N^{stat}, and Theorems 11.3.1 and 11.3.2. The idea of there being two types of probability is actually an old one. For example, Carnap [1950] talks about probability$_1$ and probability$_2$. Probability$_2$ corresponds to relative frequency or statistical information; probability$_1$ corresponds to what Carnap calls *degree of confirmation*. This is not quite the same as degree of belief; the degree of confirmation considers to what extent a body of evidence supports or confirms a belief. However, there is some commonality in spirit. Skyrms [Skyrms 1980] talks about first- and second-order probabilities, where first-order probabilities represent propensities or frequency—essentially statistical information—while second-order probabilities represent degrees of belief. These are called first- and second-order probabilities since typically one has a degree of belief about statistical information. Bacchus [1988] was the first to observe the difficulty in expressing statistical information using a possible-worlds model; he suggested using the language $\mathcal{L}^{QU,stat}$. He also provided an axiomatization in the spirit of AX_N^{stat} which was complete with respect to structures where probabilities could be non-standard reals; see [**?**] for details. Abadi and I [1994] proved that there could not be a complete axiomatization for either $\mathcal{L}_n^{QU,fo}$ or $\mathcal{L}_n^{QU,stat}$ with respect to $\mathcal{M}_n^{meas,fo}$ or $\mathcal{M}^{meas,stat}$, respectively.

The material in Section 11.4 on first-order conditional logic is largely taken from [Friedman, Halpern, and Koller 1996], including the analysis of the lottery paradox, the definitions of Pl4*, Pl4†, Pl5*, and all the technical results. Other papers that consider first-order conditional logic include [Delgrande 1987; Brafman 1997; Lehmann and Magidor 1990; Schlechta 1995; Schlechta 1996]. Brafman [1997] considered a preference order on the domain, which can be viewed as an instance of statistical plausibility. He assumed that there were no infinitely increasing sequences, and showed that, under this assumption, the analogue of C7, together with F1–F5, UGen, and analogues of C1–C4 in the spirit of PD1–PD4 provide a complete axiomatization. This suggests that adding C7 to the axioms will provide a complete axiomatization of $\mathcal{L}_n^{\rightarrow,fo}$ for $\mathcal{M}_n^{pref,fin,fo}$, and that adding C5–C7 will provide a complete axiomatization for $\mathcal{M}_n^{rank,fo}$. This has not yet been proved though. Lehmann and Magidor [1990] and Delgrande [1988] consider ways of using conditional logic for default reasoning.