1

A Brief History of Quantification

We start approaching the concept of quantification through some glimpses of its historical development, right from the Aristotelian beginnings. This is not just a soft way to introduce a technical concept. To understand how people have tried to think about quantification earlier, both their insights and their difficulties, is useful for us who try to think about it today. In fact, the evolution of notions of quantification is quite interesting, both from a historical and a systematic perspective. We hope this will be evident in what follows. But our own perspective is mainly systematic. We shall use these glimpses from the history of ideas as occasions to introduce a number of semantic and methodological issues that will be recurring themes of this book.

Section 1.1 looks at the early history of quantification, beginning with Aristotle, who initiated the logical study of the four quantifiers `all`, `some`, `not`, and `not all`, that medieval philosophers arranged in the 'square of opposition'. A closer look at that square reveals interesting facts about how negation combines with quantification, and also brings up the still debated issue of whether the truth of `all As are B` requires that there are some `As`. In the rich medieval discussion of quantification, we focus only on the distinction between categorematic and syncategorematic terms, a distinction closely tied to the issue of the meaning or significance of the quantifier expressions themselves (not just the sentences containing them), and interestingly related to several modern themes. From the Middle Ages we jump to the emergence of modern logic at the end of the nineteenth century (section 1.2). The tentative accounts of quantification in Peirce, Peano, and Russell, and the clear and explicit account in Frege, gradually turn, via Tarski's formal treatment of truth, into the modern model-theoretic perspective, where the truth conditions of quantified sentences in formal languages are relative to interpretations or *models* (section 1.3). In particular, we discuss the role of the universe of quantification, and the subtle mechanisms by which natural languages are able to restrict that universe.

Much of the discussion, and the confusion, around quantification since medieval times concerns what quantifier expressions signify or denote, or make larger expressions (like `all men`, `some sailors`, `no cats`) signify or denote. The most obvious candidates are individuals or sets of individuals, but all attempts along these lines seem to meet with insuperable difficulties. In a postscript to the chapter (section 1.4), we give a formal proof that, indeed, no systematic account along these lines *can* succeed. One needs to go one level up in abstraction (to sets of sets of individuals), and that is where the modern theory of quantification begins.

1.1 EARLY HISTORY OF QUANTIFIERS

1.1.1 Aristotelian beginnings

When Aristotle invented the very idea of logic more than 2,300 years ago, he focused precisely on the analysis of quantification. Operators like *and* and *or* were added later (by Stoic philosophers). Aristotle's syllogisms can be seen as a formal rendering of certain inferential properties, hence of aspects of the meaning, of the expressions `all`, `some`, `no`, `not all`. These provide four prime examples of the kind of quantifiers that this book is about.

A *syllogism* has the form:[1]

$$\frac{\begin{array}{l} Q_1 A\ B \\ Q_2 B\ C \end{array}}{Q_3 A\ C}$$

where each of Q_1, Q_2, Q_3 is one of the four expressions above. Later on, these expressions were often presented diagrammatically in the *square of opposition*; see Fig. 1.1.[2] Positions in the square indicate certain logical relations between the quantifiers involved. Thus, there are really two kinds of logical 'laws' at work here: the (valid) syllogistic inference schemes with their characteristic form, and the 'oppositions' depicted in the square, which do not have that form when written as inference schemes, but are nevertheless considered to be valid. As to the former, here are two typical examples of syllogistic schemes:

(1.1)
$$\frac{\begin{array}{l} \texttt{all}\ A\ B \\ \texttt{no}\ B\ C \end{array}}{\texttt{no}\ A\ C}$$

This scheme is clearly *valid*: no matter what the properties *A*, *B*, *C* are, it always holds that *if* the two premises are true, then so is the conclusion.

(1.2)
$$\frac{\begin{array}{l} \texttt{all}\ A\ B \\ \texttt{some}\ B\ C \end{array}}{\texttt{some}\ A\ C}$$

This too has the stipulated syllogistic form, but it is *invalid*: one may easily choose *A*, *B*, *C* so as to make the premises true but the conclusion false.

[1] This is the so-called *first figure*—three more figures are obtained by permuting *AB* or *BC* in the premises. We are simplifying Aristotle's mode of presenting the syllogisms, but not distorting it. Observe in particular that Aristotle was the first to use *variables* in logic, and thus to introduce the idea of an *inference scheme*.

[2] Apparently it was Apuleios of Madaura (2nd century AD) who first introduced this diagrammatic representation.

A syllogism is a particular instantiation of a syllogistic scheme:

All Greeks are sailors
No sailors are scared
———————————
No Greeks are scared

is a valid syllogism, instantiating the scheme in (1.1), and

All whales are mammals
Some mammals have fins
———————————
Some whales have fins

is an invalid one instantiating (1.2).

It was perfectly clear to Aristotle (though not always to his followers) that the *actual* truth or falsity of its premises or conclusion is irrelevant to the (in)validity of a syllogism—except that no valid inference can have true premises and a false conclusion. What matters is whether it is *possible* that the premises are true and the conclusion false: if it is possible, the syllogism is invalid; if not, it is valid. In particular, a valid syllogism can have false premises. For example, in the first (valid) syllogism above, all three statements involved are false, whereas in the second (invalid) one they are all true.

As to the logical relations depicted in the square of opposition, there is a small but crucial difference between the classical and the modern version of this square. Since this point has often been misunderstood, and since it serves to illustrate an important issue in the semantics of quantification, it is worthwhile to get clear about it.

The classical square of opposition, i.e. the square as it appears in the work of Aristotle (though he did not use the diagram) and in most subsequent work up to the advent of modern logic in the late nineteenth century, is as in Fig. 1.1. The **A** and **E**

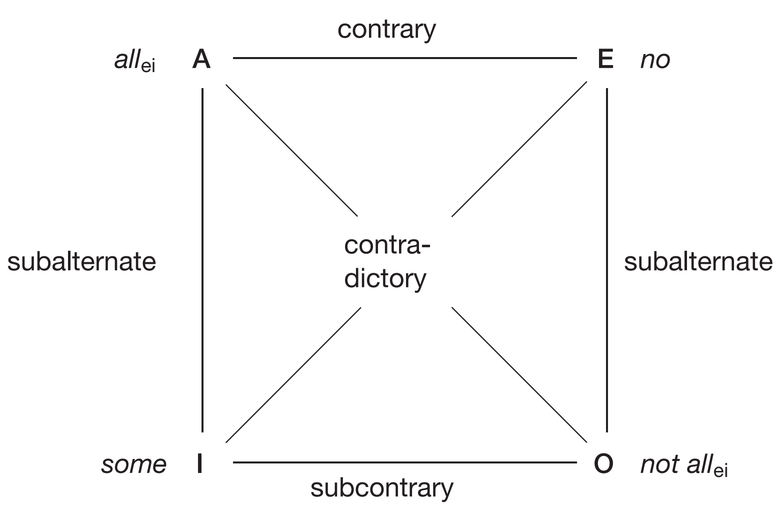

Figure 1.1 The classical square

quantifiers are called *universal*, whereas the **I** and **O** quantifiers are *particular*. Also, the **A** and **I** quantifiers are called *affirmative*, and the **E** and **O** quantifiers *negative*. Now, the important point is that the quantifier in the **A** position is what we have here called all_{ei}, i.e. the quantifier *all* with *existential import*. So $all_{ei}(A, B)$ in effect means that all *A*s are *B and there are some A*s. This is explicit with many medieval authors, but also clearly implicit in Aristotle's work: for example, in the fact that he (and almost everyone else doing syllogistics before the age of modern logic) considered the following scheme as valid:

$$(1.3)\quad \frac{\begin{array}{l}\texttt{all}\ AB\\ \texttt{all}\ BC\end{array}}{\texttt{some}\ AC}$$

The logical relations in the classical square are as follows: Diagonals connect *contradictory* propositions, i.e. propositions that cannot have the same truth value. The **A** and **E** propositions are *contrary*: they cannot both be true (note that this too presupposes that the **A** quantifier has existential import). The **I** and **O** propositions are *subcontrary*: they cannot both be false. Finally, the **I** proposition is *subalternate* to the **A** proposition: it cannot be false if the **A** proposition is true, in other words, it is *implied* by the **A** proposition (again presupposing the latter quantifier has existential import). Similarly for the **O** and **E** propositions.

In addition, the convertibility—or as we shall say, the *symmetry*—of the **I** and **E** positions, i.e. the fact that *no A's are B* implies that *no B's are A*, and similarly for *some*, was also taken by Aristotle and his followers to belong to the basic logical facts about the square of opposition. We discuss symmetry in Chapter 6.1.

Notice that it follows that the **A** and **O** propositions are negations of each other (similarly for the **I** and **E** propositions). Thus, the quantifier at the **O** position means that either something in *A* is not in *B*, or *there is nothing in A*. So the **O** proposition is *true* when *A* is empty; i.e. contrary to the modern usage, the quantifier *not all* does not have existential import. (*Q* has existential import if $Q(A, B)$ implies that *A* is non-empty.)

It seems, however, that during the late nineteenth and twentieth centuries this fact was often forgotten, and consequently it was thought that the logical laws described by the classical square of opposition were actually inconsistent. For example, people wondered how *no*(A, B) could imply *not all*(A, B). And of course it doesn't—with the modern understanding of *not all*(A, B). But it does imply $not\ all_{ei}(A, B)$—either there is nothing in *A*, in which case $not\ all_{ei}(A, B)$ holds, or there is something in *A*, which then cannot be in *B* by assumption, so again $not\ all_{ei}(A, B)$ holds—and that is all the classical square claims. Furthermore, it is often supposed that the problem arose from insufficient clarity about *empty terms*, i.e. expressions denoting the empty set.[3]

[3] For a recent statement of this view, see Spade 2002: 17. Another example is Kneale and Kneale 1962: 55–60. Consequently, it is also often assumed that the problems go away if one restricts attention to non-empty terms. But actually one has to disallow their complements, i.e. universal terms, as well, which seems less palatable. In any case, neither restriction is motivated; see n. 4.

For a detailed argument that most of this later discussion simply rests on a mistaken interpretation of the classical square, we refer to Parsons 2004. The upshot is that, apparently, neither Aristotle nor (with a few exceptions) medieval philosophers disallowed empty terms, and some medieval philosophers explicitly endorsed them.[4] And as long as one remembers that the **O** quantifier is the negation of the quantifier all_{ei}, nothing is wrong with the logic of the classical square of opposition.

A totally different issue, however, is which interpretation of words like `all` and `every` is 'correct' or, rather, most adequate for linguistic and logical purposes. Nowadays, *all* is used without existential import, and the modern square of opposition is as in Fig. 1.2.

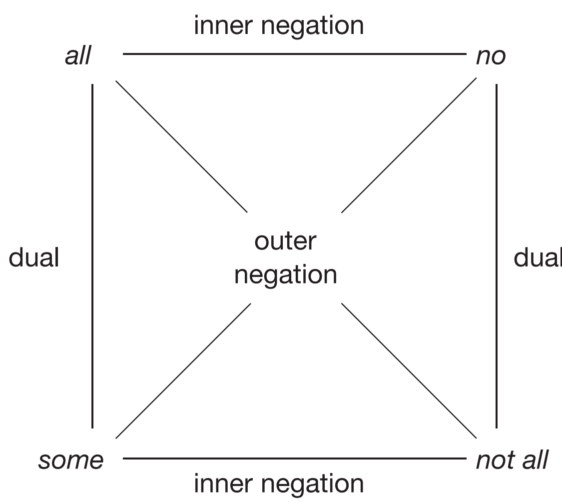

Figure 1.2 The modern square

Parsons (2004) appears to think this square is impoverished and less interesting, but we disagree on this point. The main virtue of the modern square is that it depicts three important forms of *negation* that appear in natural (and logical) languages.

three forms of negation

As in the classical square, the diagonals indicate 'contradictory' negation, or as we shall say, *outer negation*. When Q_i and Q_j are at the ends of a diagonal, the proposition Q_i `As are B` is simply the negation of Q_j `As are B`; i.e. it is equivalent to `It is not the case that` Q_j `As are B`. This propositional negation 'lifts' to the (outer) negation of a quantifier, and we can write $Q_i = \neg Q_j$ (and hence $Q_j = \neg\neg Q_j = \neg Q_i$).

[4] In Paul of Venice's *Logica Magna* (*c*. 1400), he gives

(i) Some man who is a donkey is not a donkey.

as an example of a sentence which is *true* since the subject term is empty; see Parsons 2004: sect. 5. So he allows empty terms, and confirms the interpretation of the **O** quantifier given above.

A horizontal line between Q_i and Q_j now stands for what we shall call *inner negation*: here Q_i As are B is equivalent to Q_j As are not B, which can be thought of as applying the inner negation $Q_j\neg$ to the denotations of A and B.

Finally, a vertical line in the square indicates that the respective quantifiers are each other's *duals*, where the dual of Q_i is the outer negation of its inner negation (or vice versa): $Q_i^{\mathrm{d}} = \neg(Q_i\neg) = (\neg Q_i)\neg = \neg Q_i\neg$

The modern square is *closed* under these forms of negation: applying any number of these operations to a quantifier in the square will not lead outside it. For example, $(no^{\mathrm{d}})\neg = \neg no\neg\neg = \neg no = some$. (It is not closed under other Boolean operations; e.g. the quantifier *some but not all*, which is the conjunction of *some* and its inner negation, is not in the square.)

As we will see later on (Chapter 3.2.3), all three forms of negation have natural manifestations in real languages. Moreover, this notion of a square of opposition applies to quantifiers other than those in Aristotle's square; indeed, any quantifier generates a square. That is not true of the classical square: only outer negation is present there, but not the other two forms.[5]

The issue at stake here is not whether empty terms should be allowed or not, but whether all and every have existential import or not. Do they have existential import? Everyone is familiar with the fact that it is usually odd to affirm that every *A* is *B* when one knows there are no *A*s. The dispute (and the confusion) between the two squares of opposition nicely puts the finger on the meaning of every. In modern terms, the main question is whether the existential import that is often felt with uses of every belongs to the meaning—the truth conditions—or rather is a presupposition or a Gricean implicature. We return to this issue in Chapter 4.2.1. However, we have already given here one kind of argument for choosing the modern interpretation without existential import: in this way the meaning of every fits nicely with the

[5] Though the difference between the classical Aristotelian square and the modern version might at first seem small—*all* instead of all_{ei}, and similarly for *not all*—the principled differences are huge. First, whereas outer negation is presented in both squares, neither inner negation nor dual is contained in the classical square. For example, the dual of the quantifier all_{ei} is the quantifier which holds of *A* and *B* iff *either* some *A* is *B or A* is empty. The latter is rather unnatural, and may not even be a quantirelation in the sense of Ch. 0.2.3.

Second, the logical relations depicted in the two diagrams are quite different. For example, the relation holding between two contrary propositions, that they cannot both be true (but can both be false), amounts to one implying the outer negation of the other (but not being implied by that negation). That relation neither entails nor is entailed by one of the propositions being the inner negation of the other.

Third, the classical square is not generated by any of its members. To make this statement precise, let us define *a classical square* as consisting of four quantifiers arranged as in Fig. 1.1 and with the same logical relations—contradictories, contraries, subcontraries, and subalternates—holding between the respective positions. Then each position will determine the quantifier at the diagonally opposed position, i.e. its outer negation, but *not* the quantifiers at the other two positions. For example, the reader may easily verify that

(ii) For every k and every $n \geq k$, [**A**: *at least n*; **E**: *fewer than k*; **I**: *at least k*; **O**: *fewer than n*] is a classical square.

other three quantifiers in the modern square of opposition, and thus with the three kinds of negation that any semantics of natural languages has to account for anyway. In short, *logical coherence* speaks in favor of the modern interpretation (see n. 5).

We have dwelt at some length on the logic of the square of opposition, but we also noted that the inferences which were the main focus of Aristotle's attention were the syllogisms. There are 256 syllogistic schemes. Aristotle characterized which ones of these are valid, not by enumeration but by an axiomatic method whereby all valid ones—and no others—were deducible from just two syllogisms. Apart from being the first example of a deductive system, and of a metalogical investigation of such a system, it was an impressive contribution to the logico-semantic analysis of quantification.

However, it must be emphasized that Aristotle's analysis, even when facts about the square of opposition are added, does not exhaust the meaning of these four quantifier expressions,[6] since there are many valid inference schemes involving them which do not have these forms: for example,

John knows every professor
Mary is a professor

John knows Mary

No student knows every professor
Some student knows every assistant

Some professor is not an assistant

These inferences go beyond syllogistic form in at least the following ways: (i) names of individuals play an essential role; (ii) there are not only names of properties (such as adjectives, nouns, intransitive verbs) as in the syllogisms, but also names of binary relations (such as transitive verbs); (iii) quantification can be iterated (occur in both the subject and the object of a transitive verb, for example). While none of these features may seem, at least in hindsight, like a formidable challenge, it is certainly much harder than in the syllogistic case to describe the logical structures needed to account for the validity of inferences of these kinds. At the same time, it is rather clear that a logic which cannot handle such inferences will not be able to render, say, the structure of proofs in elementary geometry (like Euclid's), or, for that matter, much of everyday reasoning.

The failure to realize the limitations of the syllogistic form, together with the undisputed authority of Aristotle among medieval philosophers and onward, is part of the explanation why logic led a rather stagnant and unproductive existence after Aristotle, all the way up to the late nineteenth century. Only when the syllogistics is extended to modern predicate logic do we in fact get a full set of inference schemes which, in a

[6] Aristotle didn't claim it did; in fact, he was well aware that there are other forms of logically valid reasoning. It is not really known why he attached such importance to the syllogisms. Perhaps he was (justly) amazed by the fact that he could apply the axiomatic method from geometry—deriving valid syllogisms from others—to objects that were not mathematical but linguistic.

precise sense, capture *all* valid inferences pertaining to the four quantifiers that Aristotle studied.[7]

1.1.1.1 Proof-theoretic and model-theoretic semantics

Approaching meaning via inference patterns is characteristic of a *proof-theoretic* perspective on semantics. The idea is that an expression's meaning is contained in a specific set of inference rules involving that expression.[8] In the case of our four quantifiers, however, it seems that whenever a certain system of such rules is proposed—such as the syllogisms—we can always ask if these schemes are *correct*, and if they are exhaustive or *complete*. Since we understand these questions, and likewise what reasonable answers would be, one may wonder if there isn't some other more primary sense in which we know the meaning of these expressions. But, however this (thorny) issue in the philosophy of language is resolved, it is clear that in the case of Aristotelian quantifiers there is indeed a different and more direct way in which their meaning can be given. Interestingly, that way is also clearly present, at least in retrospect, in the syllogistics.

For, on reflection, it is clear that each of these four quantifier expressions stands for a particular binary relation between properties, or, looking at the matter more extensionally, a *binary relation between sets of individuals*. When A, B, C are arbitrary sets, these relations can be given in standard set- theoretic notation as follows:[9]

$$\begin{aligned} &\textit{all}\,(A,B) \iff A \subseteq B \\ &\textit{some}\,(A,B) \iff A \cap B \neq \emptyset \\ &\textit{no}\,(A,B) \iff A \cap B = \emptyset \\ &\textit{not all}\,(A,B) \iff A - B \neq \emptyset \end{aligned}$$ [10]

[7] That is, all valid inferences using only the machinery of predicate logic—this is Gödel's *completeness theorem* for that logic.

[8] In particular, in intuitionistic logic, the explanations of what is a *proof* of a proposition of a certain form are taken to embody the meaning of those propositions. Thus the main semantic concept is an epistemic one. Truth is seen as a derived concept, amounting to the existence of a proof. See Ranta 1994 for an approach to semantics along these lines, within the framework of constructive type theory, and Sundholm 1989 for a similar attempt to deal with generalized quantifiers. Applying these ideas to natural languages, one obtains the *verificationist* semantics which represents one strand in modern philosophy of language; see Dummett 1975; Prawitz 1987.

These approaches place proof prior to truth in the order of semantic explanation. By contrast, many uses of proof theory are perfectly compatible with the reverse order. Proof systems for first-order logic enrich our understanding of that logic, and can be seen to be adequate by Gödel's completeness theorem. For generalized quantifiers, a classic example is Keisler 1970, which deals with an axiomatization of the quantifier *there are uncountably many*; again, the main objective is to prove a completeness theorem.

[9] We are here, as in most of this book, using a standard relational notation. Instead of $\textit{all}\,(A,B)$, one could equally well write $(A,B) \in \textit{all}$, or, using characteristic functions, $\textit{all}\,(A,B) = 1$ (= True).

[10] As we saw, the relations in effect considered by Aristotle were *some*, *no*, and

$$\begin{aligned} &\textit{all}_{\mathrm{ei}}\,(A,B) \iff \emptyset \neq A \subseteq B \\ &\textit{not-all}_{\mathrm{ei}}\,(A,B) \iff A - B \neq \emptyset \text{ or } A = \emptyset \end{aligned}$$

So, for example,

(1.4) All Greeks are sailors.

simply *means* that the set of Greeks stands in the inclusion relation to the set of sailors.

Such a formulation is consonant with a *model-theoretic* perspective on meaning, where truth and reference are basic concepts, rather than epistemic ones such as the concept of proof. In this book we consistently apply a model-theoretic approach to quantification. This is not to say that we think proof-theoretic considerations are irrelevant. For example, one case where such considerations are important concerns the notion of a *logical constant*, as we will argue in Chapter 9.4. However, our approach to the meaning of quantifier expressions is always in terms of the model-theoretic objects that these expressions denote.

1.1.1.2 Quantifier expressions and their denotations

In this connection, let us note something that has been implicit in the above. Quantifier *expressions* are syntactic objects, different in different languages. On the present account, some such expressions 'stand for', or 'denote', or have as their 'extensions', particular relations between sets. So, for example, the English `no` and the Swedish `ingen` both denote the relation that holds between two sets if and only if they are disjoint. There is nothing language-dependent about these relations. But, of course, to talk about them, we need to use language: a *meta-language* containing some set-theoretic terminology and, in the present book, English. In this meta-language, we sometimes use the handy convention that an English quantifier expression, in italics, names the corresponding relation between sets. Thus, *no* as defined above is the disjointness relation, and hence the relation denoted by both the Swedish expression `ingen` and the English expression `no`.

We note that on the present (Aristotle-inspired) analysis, each of the main words in a sentence like (1.4) has an extension; it denotes a set-theoretic object. Just as `sailor` denotes the set of sailors and `Greek` the set of Greeks, so `all` denotes the inclusion relation.

In this book we mainly use the term *denote* for the relation between a word and a corresponding model-theoretic object (its extension or denotation), with the idea that this is a modern and fairly neutral term. So we can maintain, for example, that practically everyone agrees that predicates like `sailor` and `Greek` have sets as denotations, *even though* (a) sets are a modern invention; (b) medieval philosophers, who thought a lot about the relations between words and the world, rarely talked about denotation; (c) nominalists of various ilks—medieval or later—deny that reference to anything other than individuals occurs, and sometimes also the existence of abstract objects like sets. The point is that even if you claim, as a nominalist like Ockham did, that `sailor` only refers to individual sailors, it does refer to all and only sailors, so implicitly at least, the denotation, i.e. the set of sailors, plays a role in the semantics.[11]

[11] Likewise, the criticism of so-called name theories of meaning—ridiculed by Ryle (the 'Fido'–Fido theory) for resting on a category mistake (Ryle 1949)—does not have much force

For quantifier expressions, on the other hand, one cannot say that anyone before Frege (section 1.2.4 below) thought of them as denoting. Their status in this respect is at best unclear, and some logicians explicitly held that it was incoherent to think of them as denoting anything: for example, Russell (who did use the term 'denote'). One plausible reason for this is that there seemed to be no good candidates for the denotation of these expressions, so they had to be dealt with in another way. However, we already saw that Aristotle's account of the four quantifiers mentioned so far does point to one such candidate:

Quantifier expressions denote relations between sets of individuals.

While it may be anachronistic to attribute that idea to Aristotle himself, it certainly is consistent with his approach and his focus on the syllogistic schemes. And, as it turns out, this simple idea resolves the problems encountered by earlier logicians, and provides the foundation of a coherent and fruitful account of quantification.

1.1.2 The Middle Ages

Medieval logicians and philosophers devoted much effort to the semantics of quantified statements, essentially restricted to syllogistic form. For a modern light introduction to (late) medieval logic and semantics we refer the reader to Spade 2002. Here we shall recall just one important distinction that was standard in those days—between words that have an independent meaning and words that don't. The former were called *categorematic*, the latter *syncategorematic*.

1.1.2.1 Categorematic and syncategorematic terms

The following quote is illustrative:[12]

> Categorematic terms have a definite and certain signification, e.g. this name 'man' signifies all men, and this name 'animal' all animals, and this name 'whiteness' all whitenesses. But syncategorematic terms, such as are 'all', 'no', 'some', 'whole', 'besides', 'only', 'in so far as' and such-like, do not have a definite and certain signification, nor do they signify anything distinct from what is signified by the categoremata. ... Hence this syncategoremata 'all' has no definite significance, but when attached to 'man' makes it stand or suppose for all men .... And the same is to be held proportionately for the others, ... though distinct functions are exercised by distinct syncatgoremata .... (William of Ockham, *Summa Logicae*, i, *c*. 1320; quoted from Bocheński 1970: 157–8)

against our use of "denote". Model-theoretic semantics takes no serious metaphysical stand on *naming*. And if one can associate extensions, such as sets, functions, or other abstract objects, with linguistic expressions in a systematic way, the value of such an enterprise is to be measured by its fruitfulness, its power of prediction and explanation, etc., rather than on a priori grounds.

[12] We are grateful to Wilfrid Hodges for substantial advice on medieval semantics, and in particular on what Ockham is really saying in this passage.

The word "signify" is a standard medieval term for the relation between signs and the things they are signs of. That *x signifies y* means roughly that *x* establishes an understanding of *y* in the mind of the speaker or hearer (Spade 2002: ch. 3, sect. E). Signification is thus a kind of psychologico-causal relation. The other main term was *supposit*, or "suppose" (for). The quote illustrates the interaction between the two: a word supposits for something only in a given linguistic context (and may supposit for another thing in another context), whereas signification is more absolute and context-independent (Spade 2002: ch. 8, Sect. A).[13] Now the first sentence of the quote should not, given Ockham's nominalist persuasion, be taken to mean that nouns signify any universal or abstract objects, but rather that each man is signified by `man`, etc. But the status of universals is not the issue here—the point is that quantifier expressions and other syncategoremata do not establish the understanding of anything "definite and certain" in our minds.

Ockham is not saying that a word like `all` doesn't stand for anything, but that it doesn't have the signification relation to anything. However, when combined with a noun, it *makes the noun stand* (supposit) *for* something. Syncategorematic words never signify or supposit for anything; instead, they have a systematic effect on what other words stand for.

This general contrast is fairly clear. Less clear is what it is that expressions like `all`, `some`, `no` make their adjacent nouns—or, alternatively, the respective *noun phrases*[14] —stand for. Taking "stand for" now in the weak sense of our "denote",[15] it would seem that `all men` again denotes the set of men (in Ockham's case, it stands for each man), and we might similarly assume that `some man` denotes a particular man. But what, then, would `no man` denote? It could hardly be the empty set, for then there would be no difference between `no man` and `no dog`. As we show in a postscript to this chapter (section 1.4), trying to make quantified noun phrases stand for individuals or sets of individuals is in fact an impossible task.

[13] Basic idea: In

(a) Every man is an animal.

`man` supposits for each man if you are a nominalist (and for something more abstract, like a set, if you are a realist); this is usually called *personal* supposition. In

(b) Man is a species.

it supposits for a universal concept (which in Ockham's case is a thing in the mind); *simple* supposition. And in

(c) Man has three letters

the word `man` supposits for itself; *material* supposition. The context may be a preceding quantifier expression, or something else about the form of the sentence, or a restricted universe of quantification (a *context set* in the sense to be explained in sect. 1.3.4 below).

[14] It is not clear that Ockham took noun phrases to be meaningful linguistic units, and in general medieval logicians did not have a notion of complex constituents of sentences (although they did think of sentences as complex objects). But presumably it is no great distortion of the facts to assume that what Ockham thinks the word `man`, when preceded by `all`, stands for is what he would have thought the noun phrase `all men`—had he had that notion—stands for.

[15] The closest medieval correspondent would seem to be the way realists took the relation of personal supposition (see n. 13).

The next quote too begins with an attempt to explain the semantic function of the Aristotelian quantifier expressions.

> The universal sign is that by which it is signified that the universal term to which it is adjoined stands copulatively for its suppositum (*per modum copulationis*) . . . The particular sign is that by which it is signified that a universal term stands disjunctively for all its supposita. . . . Hence it is requisite and necessary for the truth of this: 'some man runs', that it be true of some (definite) man to say that he runs, i.e. that one of the singular (propositions) is true which is a part of the disjunctive (proposition): 'Socrates (runs) or Plato runs, and so of each', since it is sufficient for the truth of a disjunctive that one of its parts be true. (Albert of Saxony, *Logica Albertucii Perutilis Logica*, iii (Venice, 1522; quoted from Bocheński 1970: 234))

The latter part of the quote, on the other hand, is a way of stating the truth conditions for quantified sentences, in terms of (long) disjunctions and conjunctions. Whatever one may think of the proposed definition, it brings out the important point that it is perfectly possible to state adequate truth conditions for quantified *sentences* without assuming that the quantifier expressions themselves denote anything. Indeed, this is how the Tarskian truth definition is usually formulated in current textbooks in first-order logic. But although medieval logicians may have taken the position—and, for example, Russell (see below) certainly did—that it is *necessary* to proceed in this way, generalized quantifier theory shows that it is in fact possible to treat quantifier expressions as denoting. Such an approach, which, as we have seen, is quite in line with Aristotle's syllogistics, has definite advantages: it identifies important syntactic and semantic categories, and it conforms better to the Principle of Compositionality, according to which the meaning of a complex expression is determined by the meanings of its parts and the mode of composition. More about this later.

1.1.2.2 Modern variants

What is the modern counterpart of the medieval distinction between categorematic and syncategorematic terms? It might be tempting to define *logic* as the study of the syncategorematic terms. But such an identification should be resisted, we think. What makes a term merit the attribute logical (or, for that matter, constant) is one thing, having to do with particular features of its semantics; we return to the issue of logicality in Chapter 9. However, the fact that a word or morpheme does not have independent semantic status is quite another thing, which applies to many expressions other than those traditionally seen as logical.[16]

A possibly related linguistic notion is that of a *grammatical morpheme*. Such a word may not belong in a dictionary at all, or if it does, there could be a description of its phonological and syntactic features (e.g. valence features: which other words it combines with), but not directly of its meaning. The meaning arises, as it were, via a

[16] When Spade suggests (2002: 116) that, from a modern point of view, categorematic terms are those that get interpreted in models, whereas the syncategorematic ones get their semantic role from the corresponding clauses in the truth definition, he comes close to identifying syncategorematicity with logicality: the standard modern idea is that non-logical constants are interpreted in models, but not logical ones.

grammatical rule; hence the name. Typical English examples might be the progressive verb ending `-ing`, the word `it` in sentences such as

(1.5) It is hard to know what Ockham meant.

and the infinitive particle `to` (the first occurrence below):

(1.6) To lie to your teacher is bad.

But, as these examples also indicate, the words that the medievals thought of as syncategorematic, such as `every`, `no`, `besides`, `only`, `whole`, and the copula `is`, are in general *not* grammatical morphemes: you do find them in dictionaries, along with attempts at explaining their meaning.[17]

It seems that there are really two basic aspects of syncategorematic terms in the medieval sense, or two senses in which the meaning of such terms is "not independent". One is the lack of ability to produce clear ideas in our minds. The other is that they are functions or *operators*: they need an *argument* to act on. Given the argument, the combined phrase gets a definite meaning (or, as the medievals preferred to put it, the argument itself gets a possibly new definite meaning).

These two aspects are partly orthogonal, it appears to us, and from a modern viewpoint it is the latter one which is most interesting. What you think can be presented clearly to the mind obviously depends crucially on your theory of the mind, and it is not at all evident that only predicate expressions have this property. That something is an operator requiring an argument, on the other hand, seems like a much more robust notion, at least in most of the cases that the medievals looked at. Today we can easily agree that their syncategoremata are indeed operators. But this does not entail that we do not have clear ideas of *how* these operators work.

A further modern issue is how you *present* the workings of these operators. In particular, can you see the quantifier expressions as *denoting*, or even "signifying", given operators? The conceptually simplest way to do this is to take the Aristotelian hint that they denote relations between sets. These are (second-order) relations; as operators, they map sets of individuals (predicate denotations) to sets of sets of individuals (noun phrase denotations). So the denotations are fairly abstract, and have only become standard with the modern application of generalized quantifier theory to natural language semantics. (It is worth recalling that irrational numbers and infinite sets were once anything but clear ideas in people's minds, however unproblematic the concepts are to modern minds.)

The alternative is to describe the operators contextually: instead of saying what a word like `every` denotes, you give uniform truth conditions for *sentences* beginning with `every`. This is the standard procedure in logic. The net result is the same, but now you need one clause for each quantifier expression, whereas with the other

[17] An extensionally better match is the linguistic distinction between *open* and *closed classes*: words belonging to closed classes often qualify as syncategorematic. But the idea behind closed classes—that they contain a small number of words to which new ones are not easily added—seems very different from the idea of not having signification.

approach you have only one general clause for quantifiers. We will see later how all of this works in detail.

Summing up, we would say that the medieval idea of syncategorematic terms as operators was quite viable, but that with (generalized) quantifiers, truth functions, etc., one also has the option of seeing such expressions as standing for "definite and certain" ideas, contrary to what the medievals (understandably) thought.

1.2 QUANTIFIERS IN EARLY PREDICATE LOGIC

Predicate logic was invented at the end of the nineteenth century. A number of philosophers and mathematicians had similar ideas, partly independently of each other, at about the same time, but pride of place goes without a doubt to Gottlob Frege. Nevertheless, it is interesting to see the shape which these new ideas about quantification took with some other logicians too, notably Peirce, Peano, and Russell.[18]

One crucial addition in the new logic was variable-binding: the idea of variables that could be bound by certain operators, in this case the universal and existential quantifiers. The idea came from mathematics—it is not something that can be directly 'copied' from natural languages—and it took some time to crystallize.[19]

[18] As to the history of the English word "quantifier", we can do no better than quote the following from Hodges' comments (pers. comm.):

William Hamilton of Edinburgh claimed to have 'minted' the words 'quantify' and 'quantification', presumably in lectures in Edinburgh around 1840. De Morgan (1847), Appendix, 312, confirms Hamilton's 'minting'. However, Hamilton's usage is that to quantify is to *interpret* a phrase as saying something about quantity, whether or not it does on the surface. Thus he talks of 'quantification of the predicate', meaning a particular theory that the predicate contains its own quantification. The modern usage seems to begin with De Morgan himself, who in De Morgan 1862 says a good deal about 'quantifying words', but then quietly shortens this to 'quantifiers'. Thus 'We are to take in both `all` and `some-not-all` as quantifiers.'

[19] In the notation for integrals, like

(a) $\int_a^b f(x)dx$

which was introduced by Leibniz in the late seventeenth century, the variable x cannot be assigned values; it is not free but bound. Likewise, algebraic equalities like

(b) $x + (y + z) = (x + y) + z$

have implicit universal quantifiers at the front. But here you seemingly can replace the variables by constants—by universal instantiation. Seeing, as Frege did, that (a) and (b) use the same mechanism, and describing that mechanism in a precise way, was no small feat.

As to natural languages, pronouns sometimes resemble bound variables, for example, in

(c) Every girl who meets him realizes that she will soon forget him.

Here `she` can be seen as a variable bound by `every girl`. `Him`, on the other hand, is deictic and somewhat similar to a *free* variable. But this analogy between pronouns and individual variables in logic was noted only well into the twentieth century, perhaps for the first time by Quine. (Quine

In a way, variable-binding belongs to the *syntax* of predicate logic, though it of course engenders the *semantic* task of explaining the truth conditions of sentences with bound variables (an explanation which took even longer to become standard among logicians). In this respect, too, Frege was unique. His explanation, though perfectly correct and precise, did not take the form that was given fifty years later by Tarski's truth definition. Instead, he treated—no doubt because of his strong views about compositionality—the quantifier symbols as categorematic, standing for certain second-order objects. This is closely related to the idea of quantifiers as relations between sets that we have traced back to Aristotle, though in Frege's case combined with a much more expressive formal language than the syllogistics and containing the mechanism of variable-binding. Nothing even remotely similar can be found with the other early predicate logicians, so let us begin with them.

1.2.1 Peirce

Peirce in fact designed two systems of predicate logical notation. One was two-dimensional and diagrammatic, employing so-called existential graphs. What the other one looked like is indicated in the following quote:

[T]he whole expression of the proposition consist[s] of two parts, a pure Boolean expression referring to an individual and a Quantifying part saying what individual this is. Thus, if k means 'he is king' and h, 'he is happy', the Boolean

$$(\overline{k} + h)$$

means that the individual spoken of is either not a king or is happy. Now, applying the quantification, we may write

$$Any(\overline{k} + h)$$

to mean that this is true of any individual in the (limited) universe. . . .

In order to render this notation as iconical as possible we may use Σ for *some*, suggesting a sum, and Π for *all*, suggesting a product. Thus $\Sigma_i x_i$ means that x is true of some one of the individuals denoted by i or

$$\Sigma_i x_i = x_i + x_j + x_k + \text{etc.}$$

In the same way, $\Pi_i x_i$ means that x is true of all these individuals, or

$$\Pi_i x_i = x_i x_j x_k, \text{ etc.}$$

. . . It is to be remarked that $\Sigma_i x_i$ and $\Pi_i x_i$ are only *similar* to a sum and a product; they are not strictly of that nature, because the individuals of the universe may be innumerable. (Peirce 1885; quoted from Bocheński 1970: 349)

Thus k, h, x are formulas here. Note Peirce's use of English pronouns as (bindable) variables in the informal explanation. Quantified sentences are divided into a (Boolean) formula and a quantifier symbol, similarly to the modern notation. The

1950 employed pronouns to explain to logic students the use of variables in logic. Quine 1960 used variables to explain how pronouns work in natural language.)

quantifier symbols are chosen in a way to indicate their meaning (in terms of conjunction and disjunction), but Peirce does not yet have quite the notation for bound variables. In $\Sigma_i x_i$ it would seem that i is a variable whose occurrences in the formula x get bound by the quantification, but in $x_i + x_j + x_k + \ldots$ the i, j, k look more like names of individuals. One sees that a formally correct expression of these ideas is still a non-trivial matter.

Note also that Peirce (in contrast with Frege; see below) seems to allow that the (discourse?) *universe* may vary from occasion to occasion.

1.2.2 Peano

One feature of standard predicate logic is that the same variables that occur free in formulas can get bound by quantification. It appears that the first to introduce this idea was Peano:

> If the propositions a, b, contain undetermined beings, such as $x, y, \ldots$, i.e. if there are relationships among the beings themselves, then $a \supset_{x,y,\ldots} b$ signifies: whatever $x, y, \ldots$, may be, b is deduced from the proposition a. (Peano 1889; quoted from Bocheński 1970: 350)

Here we have a consistent use of variable-binding, albeit only for universally quantified conditionals.

1.2.3 Russell

Bertrand Russell devoted much thought to the notion of *denotation*. In his early philosophy he subscribed to an almost Meinongian view:[20] all expressions of a certain form had to denote *something*, and it was the logician's task to say what these denotations were. Here is a quote from Russell (1903: 59):

> In the case of a class a which has a finite number of terms [members]—say, $a_1, a_2, a_3, \ldots a_n$, we can illustrate these various notions as follows:
>
> (1) *All a's* denotes a_1 and a_2 and ... and a_n.
>
> (2) *Every a* denotes a_1 and denotes a_2 and ... and denotes a_n.
>
> (3) *Any a* denotes a_1 or a_2 or ... or a_n, where *or* has the meaning that it is irrelevant which we take.
>
> (4) *An a* denotes a_1 or a_2 or ... or a_n, where *or* has the meaning that no one in particular must be taken, just as in *all a's* we must not take any one in particular.
>
> (5) *Some a* denotes a_1 or denotes a_2 or ... or denotes a_n, where it is not irrelevant which is taken, but on the contrary some one particular a must be taken.

[20] Alexius Meinong was an Austrian philosopher who took the theory of intentionality (directedness) of his teacher Franz Brentano to extreme consequences in his theory of objects (*Gegenstandstheorie*). Every thought or judgment has an object, no matter if it is a thought about Mont Blanc or a golden mountain or a round square; the only difference in the latter two cases is that these objects don't *exist* (the last one is even impossible), but they are objects no less, according to Meinong. Russell later became one of Meinong's most severe critics. In recent years Meinong's philosophy has had a certain revival; see e.g. Parsons 1980.

This is a bold attempt to explain the denotation of (what we would now call) certain quantified noun phrases; nevertheless, it is clear that the account is beset by problems similar to those in the medieval tradition.[21] It is a useful reminder of how hard this problem really was.

Later on, Russell explained the quantifiers in terms of a propositional function's being 'always true', 'sometimes true', etc., with a syntax using the notion of 'real' (free) versus 'apparent' (bound) variables.

Russell's modern view of denotation begins with his famous paper 'On denoting' (Russell 1905). In it, he still talks of "denoting phrases", but now emphatically denies that they have any meaning "in isolation". Likewise, they are not assigned any denotation.[22] Instead, Russell uses the tools of modern logic to rewrite sentences that *appear* to have a denoting phrase as, say, the subject—in the *real* logical form, according to Russell, no such phrases are left.

The reasons for Russell's change of view about denotation were chiefly logical: he found that the earlier position was incoherent and could not be consistently upheld. His main occupation was with definite descriptions and the problems arising when the purported described object did not exist. But the translation into logical form disposed of other quantified noun phrases as well (containing `every`, `some`, etc.), and the problem of their denotation disappeared. The divergence between surface form and logical form was a crucial discovery for Russell, with far-reaching consequences for logic, epistemology, and philosophy of language.[23] Though his arguments were quite forceful, it would seem that later developments in formal semantics, and in particular the theory of (generalized) quantifiers, have seriously undermined them. Roughly, this theory provides a logical form which does treat the offending expressions as denoting, and which thus brings out a closer structural similarity between surface and logical structure. One may debate which logical form is the correct one (to the extent that this question makes sense), but one can no longer claim that no

[21] For example, what is the difference between (1) and (2)? It seems that *all a's* denotes the set $\{a_1, \ldots, a_n\}$, whereas *every a* denotes—ambiguously—each one of the a_i. But if the latter were the case, one ought to be able to have an utterance of "Every a is blue" mean that, say, a_2 is blue, which clearly is impossible. A charitable interpretation may note that Russell is on to a linguistic insight that "every" is *distributive* in some way that "*all*" is not, but clearly he has not managed to express it adequately.

Similar comments can be made about the difference that Russell tries to establish between *some* and the indefinite article *a*, but here his intuitions seem to run counter to the prevailing view: i.e. that the indefinite article can be used to talk about a particular individual, whereas *some* usually involves a mere existence claim.

[22] Except that successful definite descriptions, i.e. descriptions that manage to single out a unique object, are sometimes said to denote that object. But there is nothing similar for other quantified noun phrases in Russell's paper.

[23] For example, Russell's epistemology is built on it. The main problem for Russell was how we can know anything, and talk, about objects that we are not *acquainted* with, given that he allowed very few objects that we are acquainted with (sense data, ourselves, perhaps some universals, but not physical objects or other people). On Russell's new theory, it is not necessary to know these objects directly. For the descriptions we use to be successful, it suffices to know *that* certain sentences about objects are true; and these sentences ultimately need not contain reference to anything with which we are not acquainted.

precise logical form which treats quantifier expressions or noun phrases as denoting is available. Indeed, the basic concepts needed for such a treatment were provided already by Frege.

1.2.4 Frege

Already in *Begriffsschrift* (1879), Frege was clear about the syntax as well as the semantics of quantifiers. But his two-dimensional logical notation did not survive, so below we use a modernized variant.[24]

First-level (*n*-ary) *functions* take (*n*) objects as arguments and yield an object as value. *Second-level* functions take first-level functions as arguments, and so on; values are always objects.[25] Frege was the first to use the trick—now standard in type theory[26]—of reducing predicates (*concepts*) to functions: an *n*-ary first-level predicate is a function from *n* objects to the *truth values* the True and the False (or 1 and 0), and similarly for higher-level predicates.

For example, from the sentence

(1.7) John is the father of Mary.

we can obtain the two unary first-level predicates designated by the expressions

ξ is the father of Mary

and

John is the father of η

as well as the binary first-level predicate designated by

(1.8) ξ is the father of η

(In modern predicate logic we would write *father-of* (ξ,*Mary*), *father-of* (*John*,η), and *father-of* (ξ, η), instead.) Here we have abstracted away one or two of the proper nouns in (1.7). We can also abstract away `is the father of`, obtaining the unary second-level predicate denoted by

(1.9) Ψ(John, Mary)

where Ψ stands for any binary first-level predicate. ('Mixed' predicates, like $\Psi(\xi$, Mary), were not allowed by Frege.) For example, (1.8) denotes the function

24 Frege's full-blown theory of quantification and higher-level functions and concepts was presented in *Grundgesetze* (Frege 1893); see esp. §§22–3.

25 The dichotomy function/object is fundamental with Frege. Functions, as seen above, are 'unsaturated'; they have holes or places, marked with variables in the notation, that can be filled. The filler can be object or (in the case of e.g. second-level functions) again a function, but the result (value) is always an object. Objects have by definition no places to saturate. Ordinary physical objects, numbers, extensions (sets), linguistic expressions, and the two truth values are examples of Fregean objects.

26 A type theorist might not call this a trick at all, claiming instead that functions are more fundamental mathematical objects than sets or relations.

which sends a pair of objects to the True if the first object is the father of the second, and all other pairs of objects to the False. Equivalently, we can say that (1.8) denotes the relation *father of*.[27] (1.9) can be interpreted as the set of all binary relations that hold between the individuals John and Mary.[28]

Now suppose that

(1.10) $A(\xi)$

is a syntactic name of a unary first-level predicate. According to Frege, the (object) variable ξ does not belong to the name; it just marks a place, and we could as well write

$$A(\cdot)$$

The sentence

(1.11) $\forall x A(x)$

is obtained, according to Frege, by 'inserting' the name (1.10) into the second-level predicate name

(1.12) $\forall x \Psi(x)$

(1.12) is a primitive name denoting the *universal quantifier*, i.e. the unary second-level predicate which is true (gives the value the True) for precisely those first-level (unary) predicates which are true of every object. (Again the (first-level) variable Ψ in (1.12) is just a place-holder.) So (1.11) denotes the *value* of the universal quantifier (1.12) applied to the predicate (denoted by) (1.10), i.e. a truth value. The result is that

$$\forall x A(x) \text{ is true iff } A(\xi) \text{ is true for any object } \xi$$

Other quantifiers can be given as primitive, or defined in terms of the universal quantifier and propositional operators. For example,

$$\neg\forall x \neg\Psi(x)$$

is the existential quantifier, and

$$\forall x(\Psi(x) \rightarrow \Phi(x))$$

[27] Expressions like `father of` are not phrases or constituents in any linguistic sense. Rather, `father` is usually thought to require a prepositional phrase (PP) complement (so `father` is a lexical item, and `father of Mary` but not `father of` is a phrase). But Frege should not be taken to make any syntactic claims here. His point is the semantic one that you can obtain a function by abstracting any element out of any phrase.

[28] These formulations slur over an important (Fregean) distinction. All the parametric expressions above are unsaturated, and denote concepts (functions), not objects. Sets and relations, on the other hand, are the *extensions* of concepts, and they are objects, according to Frege, who has a special notation for the extension (he calls it *Wertverlauf*, "course-of-values") of a concept. One may compare this to modern λ notation, where the parametric expressions are just formulas with free variables, and λ abstracts like $\lambda\xi,\eta\,$*father-of*$\,(\xi,\eta)$ denote the corresponding relations.

is the binary quantifier (second-level predicate) *all*—one of the four Aristotelian quantifiers. Summarizing, we note in particular

- Frege's clear distinction between names (formulas, terms) and their denotations;
- the distinction between free and bound variables (Frege used different letters, whereas nowadays we usually use the same), and that quantifier symbols are variable-binding operators;
- the fact that quantifier symbols are not syncategorematic, but denote well-defined entities, *quantifiers*, i.e. second-order (second-level) relations.[29]

1.3 TRUTH AND MODELS

Tarski (1935) was the first to formulate a formal *theory of truth*, thus making the notion acceptable to mathematicians, and, more importantly, enabling a clear distinction between simple truth and notions such as validity and provability. In this definition, logical symbols such as quantifiers and propositional connectives are explained contextually (see Chapter 1.1.2.2); they are not assigned independent interpretations; instead, truth conditions are given for each corresponding *sentence* form. Since the number of logical symbols needed for the mathematical purposes for which the logic was originally intended (such as formalizing set theory or arithmetic) is quite small, and there is one clause in the truth definition for each symbol, the truth definition is compact and easy to manage. This contextual or sentential format of the truth definition is still standard practice; see Chapter 2.2.

1.3.1 Absolute and relative truth

In one important respect Tarski's original truth definition is half-way between Frege's conception and a modern one. Frege's notion of truth is *absolute*: all symbols (except variables and various punctuation symbols) have a given meaning, and the universe of quantification is always the same: the class of *all* objects. The modern notion, on the other hand, is that of *truth in a structure* or *model*. Tarski too, in 1935, considers only symbols with a fixed interpretation, and although he mentions the possibility of relativizing truth to an arbitrary domain,[30] he does not really formulate the notion of truth in a structure until his model-theoretic work in the 1950s. The switch from an

[29] We should observe, however, that although Frege had the resources to deal with arbitrary quantifiers, the only primitive one in his *Begriffsschrift* was the universal quantifier ∀. Indeed, with Russell he strongly criticized traditional logic for its adherence to subject-predicate form, rightly pointing out that such adherence had hampered the development of an adequate logical formalism. He did not note that a formalization of natural language sentences which *preserves* this form is possible via systematic employment of quantifiers. That insight came about 100 years later, with the work of Richard Montague (1974).

[30] In connection with "present day ... work in the methodology of the deductive sciences (in particular ... the Göttingen school grouped around Hilbert)" Tarski 1935: 199 in Woodger's translation).

absolute to a relative notion of truth is in itself quite interesting; see Hodges 1986 for an illuminating discussion.

The model-theoretic notion of truth is relative to two things: an interpretation of the non-logical symbols and a universe of quantification. Let us consider these in turn.

1.3.2 Uninterpreted symbols

We are by now accustomed to thinking of logical languages as containing, in addition to *logical* symbols, whose meaning is explained contextually, such as $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\exists$, $\forall$, etc., also uninterpreted, or *non-logical*, symbols, which receive a meaning, or rather an extension, by an *interpretation*.

The point to note in the present context is that the notion of an interpretation applies nicely to quantification in natural languages, keeping in mind that an interpretation (in this technical sense) *assigns extensions* to certain words. For it is characteristic of most quantifier expressions that only the extensions of their 'arguments' are relevant. That is, in contrast with many other kinds of phrases in natural languages, *most quantifier phrases are extensional*. Furthermore, phrases that provide arguments to quantifiers, like nouns or adjectives or verb phrases, do have sets as denotations, and it is often natural to see these denotations as depending on what the facts are. The extensions of `dog` and `blue` are certain sets; under different circumstances they would have been other sets. It makes sense, then, to have a model or interpretation assign the appropriate extensions to them. Quantifier phrases, on the other hand, do not appear to depend on the facts in this way, so their interpretation is *constant*.

The above description is preliminary. It is a main theme of this book to emphasize and make precise these characteristic features of quantifiers in natural languages: constancy, extensionality (with a few notable exceptions), and the topic neutrality which contrasts them with, say, verbs, adjectives, and nouns. And it is largely these features that make model theory a suitable instrument for a general account of quantification.

Methodological digression. We are claiming that a first-order framework allows us to get at the meaning of quantifier expressions, but we are *not* making the same claim for, say, nouns. Here the claim is only that it provides all the needed denotations—but it may provide many more (e.g. various uncountable sets), and it says nothing about how these denotations are constrained. For example, nothing in predicate logic prevents us from interpreting `bachelor` in such a way that it is disjoint from the interpretation of `man`.

But this is not a problem; rather, it illustrates a general feature of all modeling.[31] We are modeling a linguistic phenomenon—quantification. The model should

[31] Where "model" is here used in the sense in which one can build a model of an aircraft to test its aerodynamical properties, not in the logical sense which is otherwise used in this book: viz. that of a mathematical structure or interpretation.

reflect the relevant features of that phenomenon, but may disregard other features of language. Abstracting away from certain features of the world, it allows us, if successful, to get a clearer view of others, by means of systematizations, explanations, even predictions that would otherwise not be available. By way of an analogy, to build a model that allows you to find out if a certain ship will float—before putting it to sea—you need to model certain of its aspects, like density, but not others, like the material of its hull—provided, of course, that the material doesn't have effects on floatability which you didn't foresee. That it doesn't is part of the claim that the model is successful. *End of digression.*

Incidentally, the treatment of individual constants in predicate logic fits well with the use of proper names in natural languages. Here the reason is not that the denotation of a name depends on what the world is like, but, first, that the main purpose of a name is to denote an individual, and second, that although languages in general contain a large number of names, most names denote many different individuals. Thus, in these respects they resemble uninterpreted symbols in a formal language. Selecting a denotation for a name is a way of fixing linguistic content for a certain occasion, and it accords with important aspects of how names are actually used, though, again, it does not constitute—and is not intended to constitute—anything like a full-scale analysis of this use.

The application of model theory allows for three degrees of fixity or variation in analyzing meanings. Certain meanings are completely fixed, given by the semantic rules for truth in arbitrary models. Others are relatively fixed, given by a model's choice of universe and interpretation function. The remaining ones are hardly fixed at all, given only by assignments to free variables. This machinery provides, we claim, just the analytic tool for an account of quantification in general.

1.3.3 Universes

The second relativization of truth is to a *universe*. In contrast with Frege's approach, where the universe is always the totality of all objects, such relativization is now a standard feature of formal mathematical languages. Also, as we shall see, it is a fundamental characteristic of natural (as opposed to logical) languages. In most common circumstances, one does not talk about (quantify over) *everything* (but see section 1.3.5 below.) For example, one restricts attention to a *universe of discourse*. This universe need not be referred to or marked in any way by syntactic features of the sentences used. It can be wholly implicit, provided by the relevant context of the discourse.

Suppose, for example, we want to talk about the structure of natural numbers with addition, multiplication, etc. To say that adding 0 to a number does not increase it, we write, in mathematics,

$$\forall x(x + 0 = x)$$

Only the context can insure that this sentence is about natural numbers and not, say, real numbers.

Of course, if one wants, the universe can be made explicit. Suppose we are in some bigger set-theoretic universe and want to express the same fact about natural numbers. We introduce a one-place predicate N for the natural numbers, and write instead:

$$\forall x(N(x) \to (x + 0 = x))$$

Now quantification is over all sets (including numbers), but the form of the sentence in fact restricts it to natural numbers.

The technical term for such restriction in logic is *relativization*. Any sentence φ written in first-order logic has a relativized version $\varphi^{(P)}$, for any one-place predicate P (not occurring in φ), obtained by restricting the quantifiers $\forall$ and $\exists$ to P.[32] Then $\varphi^{(P)}$ is true in a structure $\mathcal{M}$ if and only if φ is true in the structure $\mathcal{M} \restriction P$ obtained by 'cutting down' $\mathcal{M}$ to the subuniverse determined by P.

Now, it is a remarkable fact about most natural languages that relativization is, as it were, built into them, by means of determiners and other quantirelations. An indication of this is that the relativized versions are more easily expressed than the unrelativized ones. Compare the rather awkward

(1.13) a. For every thing (in the universe), if it is a number then it . . .
b. There is some thing (in the universe) which is a number and it . . .

which use English renderings of the unrelativized $\forall$ and $\exists$, with the much simpler but equivalent

(1.14) a. Every number . . .
b. Some number . . .

where the relativized versions are used, by means of the determiners `every` and `some`. Indeed, determiners have the semantic effect of restricting quantification to the subset denoted by the noun in the corresponding nominal phrase. (In addition, they have the syntactic effect of accomplishing *binding* in a way which is often more elegant than in formal languages, without the overt use—at least in simpler cases—of pronouns or variables.)

Quantirelations' function of restricting quantification is a fundamental fact about natural language. It will be amply discussed in Chapter 4.5. For the moment, however, our point is another. One might think, at first blush, that the mechanism of relativization would render unnecessary the use of universes in natural discourse. In principle, at least for quantification by means of determiners, if these restrict quantification to their noun arguments anyway, couldn't one assume once and for all a constant discourse universe consisting of absolutely everything? However, that is very far from the way in which natural languages actually work. Limited discourse universes that depend on context are the rule rather than the exception.

Here is an obvious example. If one claims on a certain occasion that

(1.15) All professors came to the party.

[32] Essentially, this means replacing occurrences of $\forall x \ldots$ and $\exists y \ldots$ by $\forall x(P(x) \to \ldots)$ and $\exists y(P(y) \wedge \ldots)$, respectively.

it is understood that one is not talking about all professors in the world, but perhaps only those in, say, the department where the party took place. So the universe of discourse might be the set of people either working or studying in that department. Getting this universe right matters crucially to the truth or falsity of the sentence, and thus to what is being claimed by that use of it. But none of the words in it, nor its grammatical form, carries this information. The universe is provided from somewhere else.

In principle, it would be possible to obtain the correct interpretation of (1.15) by making the denotation of `professor`, rather than the discourse universe, depend on context, so that on this occasion it denoted the set of professors in that particular department. We note in the next subsection that mechanisms of this kind are indeed operative in natural discourse. But although such mechanisms could take care of (1.15), we shall also see that in general they do not eliminate the need for context-dependent discourse universes.

We conclude that the notion of truth in a structure or model with a corresponding universe is well suited to natural language semantics. Should we go further and simply *identify* the discourse universe with the universe of the model? This is partly a methodological question, but we feel there is no need for such a stipulation. Often enough it makes sense to identify them, and a handy convention might be that if nothing else is said, the two are identical. But sometimes the discourse universe might be taken to be a subset of the universe of the model, or of some other set built from that universe.

Think of a model—a universe plus an interpretation—as representing what we want to keep fixed at a certain point. Other semantically relevant things may be allowed to vary. For example, in logic, assignments (of values to variables) do not belong to the model. In language, various sorts of context may provide crucial information, such as the universe of discourse or other means of restricting domains of quantification (see the next subsection), or the reference of certain pronouns, or, to take a different kind of example, the 'thresholds' of certain quantifier expressions like `most`.[33] A discourse universe is presumably something fairly constant —perhaps it should stay the same during a whole piece of discourse—but nothing in principle prevents one from accounting for several pieces of discourse within one model.

1.3.4 Context sets

At this point it is appropriate to mention that the mechanisms for restricting quantification in natural languages are in fact more subtle than has been hinted at so far. On an occasion of uttering (1.15), the restriction on `all professors` *could* be the whole discourse universe. But in (1.16), `most children` is presumably restricted to English children, whereas `several pen pals` is *not* so restricted. The discourse universe must contain people from all over the world. It must also contain countries, since they are quantified over as well here.

[33] How many *A*s must be *B* in order for `Most As are B` to be true? Sometimes any number more than half seems enough, but other times a larger percentage is required.

(1.16) The English love to write letters. Most children have several pen pals in many countries.

Thus, we have contextually given restrictions on quantified phrases that cannot be accounted for by means of discourse universes. Such 'local universes' were called *context sets* in Westerståhl 1985*a*.

This phenomenon is by no means unusual. Consider

(1.17) The philosophy students this year are quite sociable. Almost all freshmen already know several graduate students in other departments.

In (1.17) it would in principle be possible to cook up a discourse universe—consisting of philosophy freshmen, graduate students in other departments, the departments themselves, but nothing else—so that restricting quantification to it yields the right truth conditions. But in other cases such a strategy will not work, and anyway it would be completely artificial. It is in fact quite natural to use contextually given subuniverses limited to a particular occurrence of a noun phrase.[34]

While some version of context sets is undoubtedly used, there has been a recent discussion of whether the phenomenon is semantic or pragmatic. A semantic account presumably has to represent context sets by some sort of parameters or variables.[35] This was proposed in Westerståhl 1985*a* and argued for at length by Stanley and Szabo (2000), who gave additional support for the idea from *binding* phenomena, as in

(1.18) Whenever John shows up, most people tend to leave.

Here `most people` has to be restricted in a way that depends on the occasions of John showing up—a fixed context set will not do. This looks like an even stronger case for context set variables than examples like (1.16). Still, while presumably everyone agrees that binding occurs *somewhere* within an utterance of (1.18), pragmatic accounts of this effect have also been proposed.

One such account, propounded by Kent Bach (1994, 2000) among others, denies any contextual restriction at all in the semantics of (1.15) and (1.18), for example, which are taken to express the implausible propositions that all professors in the universe came to the party, and that whenever John shows up, most people in the universe tend to leave. However, pragmatic factors are said to insure that these propositions are not the ones *communicated* by utterances of (1.15) and (1.18); so the contextual restrictions show up only in the communicative situation. We cannot enter into this discussion here or do justice to the various arguments that have been put forward, but we do feel that an account like Bach's fails to capture the intuitively clear

[34] In the examples here, we have let a linguistic context provide the relevant context sets. This is for expository purposes; one can easily imagine the linguistic context replaced by a non-linguistic one.

[35] There is also the alternative of thinking of sentences requiring contextual restriction as elliptical or (surface) incomplete, so that the sentence 'really' used in an utterance of (1.15) would be, say, `All professors in John's department came to the party`. This idea seems empirically unmotivated to us; see Stanley and Szabo 2000 for discussion.

difference between standard contextual restriction, on the one hand, as in the above sentences, and sentences like

(1.19) Everyone was fighting.

on the other hand, uttered, say, during a description of a barroom brawl. Here the speaker may be well aware that perhaps not *everyone* in the bar participated in the fight—some bystanders might have been huddling under tables—and so she is well aware that the sentence may be literally false. Nevertheless, she succeeds in communicating another content, perhaps that *a lot of* people were fighting. But this case of using a literally false sentence to communicate something by hyperbole is quite different, it seems, from (1.15) and (1.18), where the supposed literal content is not even remotely entertained by the speaker.

The issue, however, is complex, and concerns syntactic and semantic form, as well as where the line between semantics and pragmatics should be drawn. For a recent overview of this debate, and a particular instance (very different from Bach's) of the pragmatic position in it, see Recanati 2004. As indicated, we ourselves tend to favor semantic accounts whenever they are feasible.

Another question is whether context sets belong to the determiner or to the noun in a quantified phrase. In Westerståhl 1985*a* the former view was taken, but Stanley and Szabo (2000) and Stanley (2002) claim that context sets rather restrict the head noun. Thus, on a given occasion, an utterance of a certain noun which normally denotes a set A will actually denote $A \cap X$ instead, where X is a context set. It may also be that both mechanisms occur. Linguists often argue that the definite article `the` has a given slot for contextual restriction, for example, in

(1.20) Every toothpaste tube was missing the cap.

The idea is that the meaning of `the` leads one to look for a context set (whose intersection with the set denoted by the noun has just one element in the singular case), a sort of 'familiarity' requirement. Other determiners may have a similar feature, whereas yet others lack it.

We note, however, that even in the cases where the determiner might have a context set parameter, that parameter has a different role from the other two arguments, which we will call the *restriction* and the *scope*, i.e. the first and the second arguments of the Aristotelian quantifiers in section 1.1.1. As we have already said and will make even clearer later (Chapter 4 onwards), it is most natural to regard determiners as denoting *binary* relations between sets. The possibility that some determiners may have context set parameters does not alter this fact. We recognize that to account for binding phenomena as in (1.18) and (1.20), some context mechanism must be made explicit. But for describing the meanings and stating the properties of determiners in the most natural and efficient way, these mechanisms can for most purposes be left implicit. For this reason, we will usually ignore context sets in this book.[36]

[36] A further issue is if a *set* parameter is always sufficient. One can argue that an intensional property is sometimes required. Stanley and Szabo (2000) give examples; moreover, they argue

1.3.5 Quantifying over everything?

A common idea nowadays is that Russell's paradox has shown that there cannot be a totality of everything—contrary to what Frege thought—and that quantification over everything is for this reason an incoherent notion. But this conclusion is by no means straightforward. Set theorists standardly grant that the paradox shows there can be no *set* of all sets, but call that totality a (proper) *class* instead, and happily quantify over it in the object language of set theory (at least those who think that language has an intended interpretation). Model theorists, who usually require models to have sets as universes, happily talk about "all models", again quantifying over the same totality. Is such quantifying over all sets, or over everything, just loose and imprecise talk that must be abandoned?

Williamson (2003) makes a forceful case that quantification over everything is not only possible but a necessary and fundamental feature of language. He argues that the opposite position, 'generality-relativism', is self-defeating: *it* cannot be coherently stated, he claims, and moreover cannot provide adequate accounts of natural laws, universal kind statements, and, notably, truth and meaning. To avoid a threatening Russell-like paradox in semantics, which he formulates in terms of interpretations rather than sets, his solution is basically to give up the idea that interpretations are *things*, and hence can be quantified over like other things, i.e. with first-order quantification. Instead, the semantics of an object language must be done in an irreducibly *second-order* way.

Others are not convinced and continue to claim that quantification over everything is incoherent. For example, Glanzberg (2004) uses the Russellian reasoning to argue that for every large domain, in particular every domain purporting to contain everything, there are in fact things falling outside the domain. This is a version of what Dummett (1991, 1993) has called *indefinite extensibility*, again with the conclusion that it is impossible to quantify over everything.

We shall not go into the philosophical dispute here. But the debate touches on linguistic issues too, and we make some brief comments about these.

First, as we said above, it is a fact about natural languages—in contrast with standard logical languages—that they employ *restricted* quantification, in the sense that they have a built-in slot for a noun that restricts the universe of quantification. Even `everything` decomposes into the determiner `every` and the noun `thing` going into the restriction slot. Williamson stresses that this does *not* entail that one cannot quantify over everything: simply treat `thing` as a logical noun that denotes the totality of everything. But it does indicate, we think, that restricted quantification is primary in natural language: in other words, that quantification over a limited domain is the rule rather than the exception.

that even in an extensional treatment, domain restriction parameters should instead have the form $f(i)$, where i is an individual and f a function from individuals to context sets, in order to handle quantified contexts such as (1.18).

Second, even if it were the case that one cannot quantify over everything, there is nothing that cannot be quantified over with restricted quantification, since, trivially, every thing is included in some (small) domain.

These points are small. The next point is more important, and one on which we are in sympathy with Williamson: although restricted in the above sense, natural language quantification is over everything in that the meaning of determiners like `every`, `at least five`, `most`, etc. does not involve any particular domains. We do not have to learn separately the meaning of `at least five cows`, `at least five colors`, `at least five integers`, `at least five electrons`, etc. The meaning of `at least five` and the respective nouns suffice. *There are no restrictions on the eligible domains* (except that they contain things that can be counted). Similarly, there are no restrictions on the implicit discourse universes or context sets that can be attached to quantified sentences.

This is why we interpret determiners as *operators* that *with each domain* associate a binary relation between arbitrary subsets of that domain.[37] It cannot be emphasized enough that determiners have *global* meanings in this sense. As we will see, this has important consequences for the semantics of determiners and other quantifier-denoting expressions, consequences that are sometimes obscured in the literature, because linguists often argue within a *local* perspective, where a universe is fixed by context. But certain features of quantifiers are simply not visible from such a local perspective. This point will return again and again in later chapters; see Chapter 3.1.1 for further discussion.

In this book, we follow the standard model-theoretic tradition of treating universes as *sets*. In one sense, this is less general than required; (proper) classes can be universes too, and perhaps even the totality of everything. However, our purpose is an account of the meaning of quantifier expressions. An essential part of that meaning is that these expressions have built-in slots for domains and for subsets of those domains. But, equally essentially, nothing is presupposed about the nature of the domains. Therefore, the philosophical and logical problems about very large domains are in a sense irrelevant to the meaning of quantifier expressions. You have to explain how these expressions behave on *any* domain. What domains really exist is a question for metaphysics, but not for the meaning of determiners and similar expressions.

To take an example, suppose you state, as we did above, that every thing is an element of some small domain. As Williamson forcefully points out, if there is no totality of everything, then the intended meaning of that claim simply doesn't get through! So, in that case, you are either not making any claim at all, or making a different

[37] So instead of saying that `every` denotes the inclusion relation as in sect. 1.1.1.2 above, we will say that *on each universe M* it denotes the inclusion relation *over M*. If there really were a totality of everything, we could say instead that it denotes the inclusion relation over that totality; but we don't want to prejudge that issue, and in any case we have seen that the formulation in terms of universes is congenial to both natural and mathematical language.

If one thinks of quantifiers instead as operators on noun arguments, as medieval philosophers like Ockham can be interpreted as recommending (sect. 1.1.2), the present treatment means that one more argument is added to those operators: viz. an argument for the universe.

claim from the one you wished to make. But—and this is our main point—the eventual resolution of the issue of whether you do or not tells us nothing new about the meaning of `every`. That meaning is already described in a fully adequate way by saying that it denotes a quantifier that with any domain associates the inclusion relation over that domain. What domains there are is a separate matter, and this is why we are content to use standard model-theoretic semantics for the purposes of this book.

1.4 POSTSCRIPT: WHY QUANTIFIERS CANNOT DENOTE INDIVIDUALS OR SETS OF INDIVIDUALS

There is a simple argument, familiar from elementary courses in first-order logic, that quantified noun phrases such as `every woman` and `some man` cannot denote individuals. It hinges on the observation that noun phrases that may denote individuals, such as proper names, exhibit certain semantic behavior in sentences. Specifically,

(1.21) John is brave or John is not brave.

can't be false, and

(1.22) John is brave and John is not brave.

can't be true. This behavior is necessitated by the fact that the set of brave individuals is the complement of the set of individuals who are not brave. In contrast to (1.21) and (1.22), the sentence

(1.23) Every woman is brave or every woman is not brave.

clearly can be false, and

(1.24) Some man is brave and some man is not brave.

clearly can be true. So `every woman` and `some man` obviously do not behave semantically or logically as they would have to if they denoted individuals.

To demonstrate that quantified noun phrases like `every woman` and `some man` can't denote *sets* of individuals requires a more sophisticated argument, which we present below. As we will see, the argument shows *a fortiori* that in a compositional semantics, quantified noun phrases can't denote individuals. The demonstration is based on ideas in Suppes 1976.

Let us consider how many distinct things are denoted by quantified noun phrases. To make the argument, it suffices to consider only two type $\langle 1, 1\rangle$ quantifiers. We'll take *every* and *some*, the relations denoted by `every` and `some`. The argument starts from the straightforward observations that

(1.25) `Every A is B` is true iff $A \subseteq B$

(1.26) `Some A is B` is true iff $A \cap B \neq \emptyset$

where A is the set denoted by A, and B the set denoted by B. These truth conditions on the sentences impose a fundamental requirement on the semantic adequacy of whatever every A and some A are taken to denote.

We introduce the notation [[every A]] for whatever every A does denote. Similarly for [[some A]] and some A. The truth value of a sentence of the form every A is B is produced by a rule that composes the denotation [[every A]] with B, the denotation of is B, and similarly for some A is B, [[some A]] and B. So if the semantics is compositional, there must be a relation R that holds in both cases between the denotations of the quantified noun phrase and the predicate for the sentence to be true. Then the truth conditions (1.25) and (1.26) translate into the following fundamental requirements on the semantic adequacy of the denotations [[every A]] and [[some A]].

(1.27) $R([[\texttt{every A}]], B)$ iff $A \subseteq B$

(1.28) $R([[\texttt{some A}]], B)$ iff $A \cap B \neq \emptyset$

We will use these semantic adequacy requirements to determine how many distinct denotations are required for all noun phrases of the forms every A and some A. As will be seen when we have done this, more denotations are needed than there are sets of individuals. Therefore, whatever the denotations of these quantified noun phrases are, they cannot be sets of individuals. *A fortiori*, they cannot be individuals, of which even fewer exist than sets of individuals.

To facilitate counting, let us define operators q_{every} and q_{some} such that

$$q_{\text{every}}(A) = [[\texttt{every A}]]$$

and

$$q_{\text{some}}(A) = [[\texttt{some A}]]$$

where A is the denotation of A. No assumption is made about the *values* of these operators, i.e. about what sort of things every A and some A might denote. We first show that the operators must be one-to-one.

Lemma 1
(1.27) *implies that* q_{every} *is 1–1, and* (1.28) *implies that* q_{some} *is 1–1.*

Proof. Suppose (1.27) holds, and $q_{\text{every}}(A_1) = q_{\text{every}}(A_2)$. Then we have:

$$\begin{aligned} A_1 \subseteq A_1 \quad &\Longrightarrow \quad R(q_{\text{every}}(A_1), A_1) \\ &\Longrightarrow \quad R(q_{\text{every}}(A_2), A_1) \\ &\Longrightarrow \quad A_2 \subseteq A_1 \end{aligned}$$

Thus, $A_2 \subseteq A_1$. By an exactly symmetric argument, $A_2 \subseteq A_2$ implies $A_1 \subseteq A_2$. Thus $A_1 = A_2$, and q_{every} is one-to-one.

Suppose (1.28) holds and $q_{\text{some}}(A_1) = q_{\text{some}}(A_2)$. Then:

$$\begin{aligned} A_1 \subseteq A_1 \quad &\Longrightarrow \quad A_1 \cap \overline{A_1} = \emptyset \\ &\Longrightarrow \quad \text{not } R(q_{\text{some}}(A_1), \overline{A_1}) \\ &\Longrightarrow \quad \text{not } R(q_{\text{some}}(A_2), \overline{A_1}) \\ &\Longrightarrow \quad A_2 \cap \overline{A_1} = \emptyset \\ &\Longrightarrow \quad A_2 \subseteq A_1 \end{aligned}$$

So again, $A_2 \subseteq A_1$, and by a symmetric argument, $A_1 \subseteq A_2$, i.e. $A_1 = A_2$, and q_{some} is one-to-one. □

Next we show that the ranges of q_{every} and q_{some} intersect only in the images of singleton sets; that is, if $[[$`every` $\mathtt{A}_1]] = [[$`some` $\mathtt{A}_2]]$, then $\mathtt{A}_1$ and $\mathtt{A}_2$ denote the same set of one individual.

Lemma 2
(1.27) *and* (1.28) *jointly imply that if* $q_{every}(A_1) = q_{some}(A_2)$, *then* $A_1 = A_2 = \{a\}$ *for some individual a.*

Proof. Suppose (1.27) and (1.28) hold, and $q_{\text{every}}(A_1) = q_{\text{some}}(A_2)$. Then, for every set B,

$$\begin{aligned} A_1 \subseteq B \quad &\Longleftrightarrow \quad R(q_{\text{every}}(A_1), B) \\ &\Longleftrightarrow \quad R(q_{\text{some}}(A_2), B) \\ &\Longleftrightarrow \quad A_2 \cap B \neq \emptyset \end{aligned}$$

With $B = A_1$, we obtain $A_2 \cap A_1 \neq \emptyset$, since $A_1 \subseteq A_1$. Let a be a member of $A_2 \cap A_1$. With $B = \{a\}$, we obtain $A_1 \subseteq \{a\}$, since $A_2 \cap \{a\} \neq \emptyset$. Thus $A_1 = \{a\}$, since $a \in A_1$. And since $A_1 \not\subseteq \overline{\{a\}}$, with $B = \overline{\{a\}}$, we obtain $A_2 \cap \overline{\{a\}} = \emptyset$. Thus $A_2 \subseteq \{a\}$ and so $A_2 = \{a\}$. □

What we set out to establish now follows from these two lemmas. Consider any finite universe of discourse M with m individuals as its elements. Then exactly 2^m sets of individuals exist. q_{every} maps each of these sets to something, mapping different sets to different things. Thus 2^m different values are needed for $q_{\text{every}}(A)$ as A ranges over all the subsets of M. Likewise, q_{some} maps each of the 2^m subsets of M to different things, requiring 2^m distinct values for $q_{\text{some}}(A)$ as A ranges over all subsets of M. The key point of the argument is this: At most m of the values they require can be shared by q_{every} and q_{some}. Only when $A_1 = A_2 = \{a\}$ for some $a \in M$, can $q_{\text{every}}(A_1)$ be identical to $q_{\text{some}}(A_2)$. Thus q_{every} and q_{some} between them require at least $2^m + 2^m - m = 2^{m+1} - m$ distinct values. However, there are only 2^m sets of individuals in the universe, and $2^m < 2^{m+1} - m$. Therefore the denotations of `every A` and `some A` cannot be sets of individuals, as there are not enough such sets to meet the semantic adequacy requirements (1.27) and (1.28). *A fortiori*, the denotations of `every A` and `some A` cannot be individuals, as even fewer individuals exist than sets of individuals.[38]

Note that this argument does not depend on the language containing lots of quantifier-denoting expressions. We chose two of the most simple and familiar determiners, `every` and `some`, and, by calculating the number of denotations, of whatever kind, needed jointly for `every A` and `some A` as *A* varies over arbitrary subsets of the universe, showed that this number exceeds the number of such subsets, let alone the number of individuals. The only assumptions required to reach this conclusion are (a) that every subset of a finite universe can in principle be denoted by some expression, and (b) that the semantics is compositional, so that the semantic values of `every A` and `some A`, and of sentences containing them, are given by the *same* rule.

Recall that in section 1.1.2.1 above we said, in connection with Ockham's account of syncategorematic terms, that one might *try* to let `every A` denote the set *A* (the denotation of `A`), and let `some A` denote some individual *a* in *A* (or perhaps $\{a\}$), but that problems were bound to appear with the denotation of `no A`. We now see that not even the seemingly simple suggestion for `every` and `some` works out, if the semantics is to be compositional.

Indeed, treating `every A` and `some A` as denoting *sets of sets* of individuals—the proposal discussed in the next chapter—is the simplest semantically adequate choice among those available.

[38] Even if some denotations were individuals, and other denotations sets of individuals, only $2^m + m$ denotations would be available; and when $m > 2$, this number is less than the $2^{m+1} - m$ distinct denotations that are needed.