Formulas of First-Order Logic in Distributive Normal Form

Karen Nelte

September 1997

A thesis prepared for the degree of Master of Science
Supervisors: Professor Chris Brink and Dr Ilkka Kieseppä

Department of Mathematics and Applied Mathematics
University of Cape Town

Abstract

It was shown by Jaakko Hintikka that every formula of first-order logic can be written as a disjunction of formulas called *constituents*. Such a disjunction is called a *distributive normal form* of the formula. It is a generalization of the disjunctive normal form for propositional logic. However, there are some significant differences between these two normal forms, caused chiefly by the impossibility of defining the constituents in such a way that they are all consistent. Distributive normal forms and some of their properties are studied. For example, the size of distributive normal forms is examined, and although we can't determine exactly how many constituents (of each form) are consistent, it is shown that the vast majority are inconsistent. Hintikka's definition of *trivial inconsistency* is studied, and a new definition of trivial inconsistency is given in terms of a necessary condition for the consistency of a constituent which is stronger than the condition which Hintikka used in his definition of trivial inconsistency. An error in Hintikka's attempted proof of the *completeness theorem of the theory of distributive normal forms* is pointed out, and a similar completeness theorem is proved using the new definition of trivial inconsistency.

Acknowledgements

I wish to thank my supervisors Professor Chris Brink and Dr Ilkka Kieseppä for their guidance and encouragement in the preparation of this thesis. I also thank Dr Peter Jipsen for the helpful discussions on logic, Professor Roger Maddux for the idea that distributive normal forms could be a thesis topic, and Dr Hubertus Bargenda for the assurance that maths at UCT could be what I thought maths ought to be. I thank FACCS-lab for the pleasant environment and use of computing facilities. I am grateful to the Foundation for Research Development for financial assistance during my postgraduate studies, and to the Department of Mathematics and Applied Mathematics for the tutorship. I also thank my parents for their support and encouragement.

Contents

Introduction

The aim of this thesis is to present the subject of the distributive normal form for classical first-order logic in a way that is both precise and easy to read. The existence of this normal form was proved by Hintikka [1953]. In Hintikka [1965a], it was described using slightly different terminology and some of its important properties were studied. In particular, an attempt was made to show that a particular disproof method, based on the distributive normal form, was a complete disproof method for first-order logic. But the attempted proof contained an error (see p. 114). Hintikka has used distributive normal forms as a technical basis for his discussion of a number of topics such as information, analysis, synthesis, induction and deduction. These subjects are discussed in Hintikka [1973d], Saarinen *et. al.* [1979], and Bogdan [1987].

The disjunctive normal form for classical propositional logic is well-known. It can be conveniently described by reference to *possible worlds*. A possible world for a language is a possible state that can be described by that language, given by the truth values of the formulas (for propositional logic) or sentences (for first-order logic). Thus, for a propositional language, a possible world may be specified by a (propositional) valuation. If there are finitely many propositional variables $p_1, \ldots, p_n$, then there are finitely many possible worlds, and it is easy to find formulas which describe these worlds. We simply need to say whether each of the variables p_i is true or false, so the formulas of the form

$$(\pm)p_1 \wedge \ldots \wedge (\pm)p_n$$

where each occurrence of $(\pm)$ may be replaced by a negation sign $\neg$ or by nothing, describe the possible worlds. Now, any formula is the disjunction of the formulas describing the possible worlds in which it is true. This disjunction is the (full) disjunctive normal form of the formula.

The *distributive normal form* for first-order logic is a generalization of the propositional disjunctive normal form. It may also be described with reference to possible worlds. However, the situation here is more complex because there are infinitely many possible worlds, and each requires an infinite number of formulas to describe it. Thus, it is not possible to find a finite set of formulas such that each formula can be expressed as some disjunction of elements of that set. This brings us to the notion of *depth*. There are different ways in which the depth of a formula can be defined (see chapter 2). For now, it may be considered as the maximum number of quantifiers whose scopes overlap. If we consider the sentences of a fixed depth, we have a situation similar to the propositional case. Where there are formulas which are descriptions of possible worlds (as in propositional logic), they are the strongest (most informative) formulas. In first-order logic there are not any strongest sentences, but there are strongest sentences (called *constituents*) of each depth. And each sentence of some depth can be expressed as a disjunction of constituents of that depth. The situation for formulas with free variables is similar.

Since distributive normal forms are syntactically rather complex, I use chapter 2 to introduce them by way of an example. For this example, I have chosen a language which is as simple as possible, and consider only the case of sentences, for which I try to give an intuitive description of what constituents are and why they give a normal form for first-order sentences.

Chapter 1 gives the particular definitions of general logical concepts that I use. It also discusses the propositional disjunctive normal form, and two generalizations of it, to the minimal modal logic, and to monadic first-order logic, as preparation for the definitions of constituents for full first-order logic and proof of the existence of distributive normal form, in chapter 3.

Chapter 4 is in a sense the main part of this thesis. It examines the size of distributive normal forms, finds a lower bound for the fraction of inconsistent constituents, and shows how the existence of an algorithm for determining whether or not an arbitrary constituent is consistent would imply that first-order logic was decidable. Then some sufficient, though not necessary, decidable conditions for the inconsistency of a constituent are considered, and it is shown how they can be used to prove a *completeness theorem of the theory of distributive normal forms*. Though it is a slightly different result to the one Hintikka tried to prove in Hintikka [1965a], it still provides a way to prove the completeness of first-order logic.

Chapter 5 shows some examples of how distributive normal forms have been used in inductive logic and verisimilitude.

Finally some remaining open questions are briefly mentioned.

Because it is often necessary to know the precise definition of a term or symbol in contexts far from where it is defined, an index and symbol table are provided.

Chapter 1

Logical background

This chapter gives the basic definitions for classical propositional and first-order logic, and discusses the propositional disjunctive normal form which is generalized to *distributive normal form* for first-order logic in later chapters. It also considers the normal forms for the minimal modal logic and monadic first-order logic, both of which can be seen as generalizations of the propositional disjunctive normal form which introduce some of the ideas that will be used for full first-order logic in a somewhat simpler setting.

Much of the terminology and notation used in the literature for various logical concepts is not standard. I have tried to use notation and terminology that at least is somewhat uniform for the different logics discussed. In this chapter I give the particular forms of the definitions that I will use, and also indicate some alternative definitions found in the literature. Most results stated here are well-known and proofs can be found in the various books cited.

In writing this chapter, I have consulted the following books: Barwise [1977], Boolos and Jeffrey [1974], Bell and Machover [1977], Bell and Slomson [1971], Church [1956], Chang and Lee [1973], Fitting [1990], Hilbert and Ackermann [1950], Hamilton [1988], Hughes and Cresswell [1968], Hughes and Londey [1965], Hodel [1995], Mendelson [1987], Robbin [1969], Rogers [1971], Shoenfield [1967], Smullyan [1968], van Benthem [1983].

1.1 Propositional logic

Propositional syntax

A *propositional language* may be defined in various different ways by taking different sets of logical connectives as primitive. I use the following definition.

A *propositional language* $\mathcal{L}$ contains the following symbols

- logical connectives: $\neg$ $\vee$
- a non-zero countable number of propositional variables, the set of which is denoted $\mathbf{Var}_{\mathcal{L}}$
- parentheses: ( )

Different propositional languages have different sets of propositional variables.

The *formulas over a subset* $\mathbf{P}$ *of* $\mathbf{Var}_{\mathcal{L}}$ are defined as follows.

- An element of $\mathbf{P}$ is a formula over $\mathbf{P}$.
- If F and G are formulas over $\mathbf{P}$ then $\neg F$ and $(F \vee G)$ are formulas over $\mathbf{P}$, which are called the *negation* of F and the *disjunction* of F and G respectively.

The parentheses are used to record the way in which the formula was formed. When writing a disjunction on its own (i.e. not as part of another formula), then its parentheses may be omitted.

The *formulas of a propositional language* $\mathcal{L}$ are the formulas over $\mathbf{Var}_{\mathcal{L}}$. The set of formulas of $\mathcal{L}$ is denoted $\mathbf{Form}_{\mathcal{L}}$. Where we don't need to specify the language we are dealing with, we may write **Var** for the set of propositional variables and **Form** for the set of formulas.

To give abbreviations for certain formulas, the additional logical connectives $\wedge$, $\rightarrow$, $\leftrightarrow$ are defined by

- $(F \wedge G)$ abbreviates $\neg(\neg F \vee \neg G)$
- $(F \rightarrow G)$ abbreviates $\neg F \vee G$
- $(F \leftrightarrow G)$ abbreviates $(F \rightarrow G) \wedge (G \rightarrow F)$.

Again, when writing these formulas on their own, their parentheses may be omitted. The formulas $F \wedge G$, $F \rightarrow G$ and $F \leftrightarrow G$ are called repectively the *conjunction*, *conditional* and *biconditional* of F and G.

Propositional semantics

Consider $\mathbf{2}$ as the set $\{0, 1\}$. A *valuation* is a function $v : \mathbf{Form} \rightarrow \mathbf{2}$ which satisfies

- $v(\neg F) = \begin{cases} 0 & \text{if } v(F) = 1 \\ 1 & \text{if } v(F) = 0 \end{cases}$

- $v(F \vee G) = \begin{cases} 0 & \text{if } v(F) = 0 \text{ and } v(G) = 0 \\ 1 & \text{if } v(F) = 1 \text{ or } v(G) = 1. \end{cases}$

A valuation v : **Form** $\to$ **2** is uniquely determined by the values it takes on **Var**. For a function w : **Var** $\to$ **2**, the valuation v : **Form** $\to$ **2** which satisfies $v(p_i) = w(p_i)$ for each $p_i \in$ **Var** is called the *extension* of w to **Form**. Thus a valuation could also be defined as a function w : **Var** $\to$ **2**. (This is shown, for example, by Smullyan [1968] (p. 10–11).)

If $v(F) = 1$, F is called *true under* v and if $v(F) = 0$, F is called *false under* v.

The above definition of valuation is simply writing out explicitly the values obtained by considering **2** as the 2-element Boolean algebra which has 1 as the top and 0 as the bottom, and interpreting $\neg$ as complement and $\vee$ as join.

If F is a conjunction, conditional, or biconditional, the value of $v(F)$ is determined by the above definition by rewriting the formula that has been defined as an abbreviation in full. For example,

$$v(F \wedge G) = v(\neg(\neg F \vee \neg G)) = \begin{cases} 0 & \text{if } v(F) = 0 \text{ or } v(G) = 0 \\ 1 & \text{if } v(F) = 1 \text{ and } v(G) = 1. \end{cases}$$

Thus conjunction is interpreted as meet in the Boolean algebra **2**.

Given a function v : **Var** $\to$ **2**, we can give a definition which is equivalent to the one above for the truth values of all formulas by the relation $\models$ defined by

- $v \models p_i$ iff $v(p_i) = 1$ if p_i is a propositional variable
- $v \models \neg F$ iff $v \not\models F$
- $v \models F \vee G$ iff $v \models F$ or $v \models G$.

Now, for each formula F, $v \models F$ iff F is true under the extension of v to **Form**. This form of definition of the truth values of formulas is used later for the modal and first-order cases.

It is also useful to define a *partial valuation* on a subset **P** of **Var** to be a function from the set of formulas over **P** to **2** which satisfies the same conditions given above for a valuation, or equivalently as a function v : **P** $\to$ **2** extended to a function from the set of formulas over **P** to **2** as for a valuation.

If F is true under v then v is said to *satisfy* F. A valuation *satisfies* a set of formulas $\mathbf{X}$ if it satisfies each element of $\mathbf{X}$. A formula or set of formulas is *satisfiable* if it is satisfied by some valuation. A formula is a *tautology* (also called a *propositionally valid* formula) if it is satisfied by all valuations. That F is a tautology is denoted $\models F$.

A formula F is a *logical consequence* of a formula or set of formulas $\mathbf{X}$ if every valuation which satisfies $\mathbf{X}$ also satisfies F. There is no standard symbol used to denote logical consequence. I use $\mathbf{X} \Longrightarrow F$ to denote that F is a logical consequence of $\mathbf{X}$. That F is a logical consequence of the empty set is denoted $\Longrightarrow F$, which is equivalent to F being a tautology, that is $\models F$ iff $\Longrightarrow F$.

Two formulas are *logically equivalent* if they are satisfied by the same valuations. I use $F \Longleftrightarrow G$ to denote that F and G are logically equivalent. By these definitions $F \Longleftrightarrow G$ iff ($F \Longrightarrow G$ and $G \Longrightarrow F$).

$\Longleftrightarrow$ is an equivalence relation on the set of formulas and for many purposes it is not necessary to distinguish between equivalent formulas. Since for any formulas F, G, H and $* \in \{\wedge, \vee\}$, $F * G \Longleftrightarrow G * F$ and $(F * G) * H \Longleftrightarrow F * (G * H)$, we can write $F * G * H$ for the formula with parentheses in either position. Also, for any finite non-empty set of formulas $\mathbf{X}$,

- $\bigvee \mathbf{X}$ abbreviates $X_1 \vee \ldots \vee X_n$
- $\bigwedge \mathbf{X}$ abbreviates $X_1 \wedge \ldots \wedge X_n$

where $X_1, \ldots, X_n$ are the elements of $\mathbf{X}$ in any order. $\bigvee\{X_i \mid i \in I\}$ may also be written as $\bigvee_{i \in I} X_i$, and $\bigwedge\{X_i \mid i \in I\}$ as $\bigwedge_{i \in I} X_i$. Also, since all unsatisfiable formulas are logically equivalent, we use $\bigvee \emptyset$ as an abbreviation for any unsatisfiable formula, and since all valid formulas are logically equivalent, we use $\bigwedge \emptyset$ as an abbreviation for any valid formula. This agrees with the interpretation of disjunction as join and conjunction as meet in **2**, and will allow us to state the existence of conjunctive and disjunctive normal forms for all formulas without having to consider valid or unsatisfiable formulas separately.

The following notation will give us a convenient way of defining certain sets of formulas.

I use the symbol = instead of the word "abbreviates" in defining abbreviations (or names) for formulas, and also between formulas where one has been defined as an abbreviation of the other. The context should make clear how it is being used.

If F is a formula, then the formulas of the form $(\pm)F$ are F and $\neg F$. For any finite non-empty set of formulas $\mathbf{X} = \{X_1, \ldots, X_n\}$, we will call the formulas of the form

$$(\pm)X_1 \wedge \ldots \wedge (\pm)X_n$$

the *basic conjunctions generated by* $\mathbf{X}$. In chapter 3, a number of lists of formulas are defined as those basic conjunctions which are generated by some set of formulas. For some particular generating sets, the basic conjunctions generated will be called *primitive conjunctions*. These definitions can be stated just in terms of sets of formulas. But it is convenient to have the basic conjunctions generated by some set listed in some order (if the formulas in the generating set are given in some order). So for a definition of the form

$$Z_i = (\pm)X_1 \wedge \ldots \wedge (\pm)X_n$$

we take $i = 1$ to give the formula with all its conjuncts unnegated and proceed in normal counting order (the truth values of the conjuncts change more often the further to the right they are) until $i = 2^n$ gives the formula with all its conjuncts negated. This order will be called *standard order*. For example, a definition like

$$H_i = (\pm)F \wedge (\pm)G$$

gives $H_1 = F \wedge G$, $H_2 = F \wedge \neg G$, $H_3 = \neg F \wedge G$, $H_4 = \neg F \wedge \neg G$.

For a set of formulas $\mathbf{X} = \{X_1, \ldots, X_n\}$, for each valuation $v : \mathbf{Form} \rightarrow \mathbf{2}$, exactly one of the basic conjunctions generated by $\mathbf{X}$ is true under v: Let F be the basic conjunction generated by $\mathbf{X}$ such that for each $i \in \{1, \ldots, n\}$, if $v \models X_i$ then X_i is a conjunct of F, and if $v \not\models X_i$ then $\neg X_i$ is a conjunct of F. Then $v \models F$. And if G is any basic conjunction generated by $\mathbf{X}$ other than F, then there is at least one i for which one of the following conditions holds.

- X_i is a conjunct of F and $\neg X_i$ is a conjunct of G, in which case $v \models X_i$ so $v \not\models \neg X_i$ so $v \not\models G$.
- $\neg X_i$ is a conjunct of F and X_i is a conjunct of G, in which case $v \models \neg X_i$ so $v \not\models X_i$ so $v \not\models G$.

Thus $v \not\models G$.

If $\mathbf{X}$ is a finite set of formulas, then for each element X_i of $\mathbf{X}$, there is at least one disjunction of basic conjunctions generated by $\mathbf{X}$ to which X_i is logically equivalent. The disjuncts may be taken to be all the basic conjunctions generated by $\mathbf{X}$ in which X_i occurs unnegated. If any formula in $\mathbf{X}$ other than X_i is valid or unsatisfiable then some of these disjuncts will be unsatisfiable, so the expression of X_i as a disjunction of basic conjunctions generated by $\mathbf{X}$ will not be unique. (This is expressing X_i in propositional disjunctive normal form, as defined in the next section, with the elements of $\mathbf{X}$ for the propositional variables.)

Any disjunctions (or conjunctions) which have the same disjuncts (or conjuncts), though the order in which they occur or the number of times they are repeated may differ, will be called *notational variations* of each other. For example, $F \wedge F \wedge G$ and $G \wedge F$ are notational variations of each other. Whenever a formula is said to be unique, formulas which are notational variations of each other are considered as the same.

Metatheory

We have just seen examples of how the *syntax* and *semantics* of a logic may be defined. The *syntax* consists of the definition of a formal language which is given by saying what symbols it contains and which strings of these symbols are to be *formulas*. The definition must be such that we can effectively determine whether or not any given string of symbols is a formula of the language. The *semantics* consists of the specification of some mathematical structure (e.g. valuations for propositional logic, and models and valuations for first-order logic) together with a definition of how formulas may be interpreted as either *true* or *false* in such a structure, and also the definition of some general concepts such as *logical consequence* and *validity* which are defined in terms of conditions on all of the structures in which the formulas are interpreted.

A *proof method* for a logic is a way of formally deriving certain formulas from sets of formulas, based on the syntax of the logic. To say that a derivation is *formal* means that the only steps it may contain are those which are specified by some precisely defined set of rules (and we must be able to effectively determine whether or not any proposed step is an instance of one of these rules). To say that it is based on syntax means that the rules are in terms of uninterpreted symbols, including those of the language defined by the syntax. Those formulas that can be derived from the empty set are called *theorems*. Many kinds of proof methods have been developed, for example: (Hilbert) axiom systems, natural deduction, tableaux, resolution, (Gentzen) sequent calculi. (Fitting [1990] gives proof methods of all these kinds.) In all of them the derivations have a finite length or size, so a derivation from any set of formulas can use at most finitely many elements of that set.

A concept is called *syntactic* if its definition is based on syntax and *semantic* if its definition is based on semantics. For example, a proof method is a syntactic concept. There are a number of concepts, such as decidability and consistency (defined below), for which there are both syntactic and semantic versions of the concept, where the syntactic version is relative to some proof method and is equivalent to the semantic version for those proof methods which are complete (defined below).

A proof method is called *sound* if any formula that can be derived from some set of formulas is a logical consequence of them. A proof method is *complete* if it is sound and every formula that is a logical consequence of some set of formulas can be derived from them. (Completeness is often defined so as not to include soundness, but Kneebone [1963] (p. 70), for example, defines completeness so that a proof method must be sound to be complete. And both Barwise [1977] (p. 35) and Chang and Keisler [1973] (p. 8, 32) include soundness in the statement of the completeness theorems.) There are also weaker notions of soundness and completeness where a proof method is called *weakly sound* if every theorem is valid, and *weakly complete* if it is weakly sound and every valid formula is a theorem.

A logic is called *complete* if there is some complete proof method for it.

Propositional logic has complete proof methods of all of the kinds mentioned above (proved e.g. in Fitting [1990]), and thus is complete.

A *decision procedure* for a property is a mechanical test, which, applied to any object of the appropriate kind, after some finite number of steps correctly identifies whether or not that object has the property. For a logic, a *decision procedure* is such a test for determining whether or not a formula is valid. A logic is called *decidable* if it has a decision procedure. This use of "decidable" to mean that validity is decidable is found in Hilbert and Ackermann [1950] (p. 112-113) and Barwise [1977] (p. 16). A proof method may be called decidable if there is an effective method of determining whether or not any formula is a theorem (e.g. Hamilton [1988] (p. 43) and Rogers [1971] (p. 215) use "decidable" in this sense). For a complete proof method these two notions of decidability are equivalent.

If v is a valuation and F a formula, then the value of $v(F)$ depends only on the propositional variables occurring in F, of which there are a finite number. So to determine whether or not any formula F is a tautology, we need only consider those partial valuations v over the set of propositional variables in F, of which there are 2^n where n is the number of propositional variables in F. Then F is a tautology iff for each such v it is the case that $v(F) = 1$. This decision procedure for propositional logic is essentially the one often referred to as the *truth-table* method. This method can also be used to determine whether or not a formula F is a logical consequence of a finite set of formulas $\mathbf{X}$ since $\mathbf{X} \Longrightarrow F$ iff $\Longrightarrow \bigwedge \mathbf{X} \to F$ (this is immediate from definition and is the version of the semantic deduction theorem as stated, for example, by Chang and Lee [1973] (p. 16)) iff $\models \bigwedge \mathbf{X} \to F$.

A formula or set of formulas $\mathbf{X}$ is often called *consistent* (relative to some proof method) if there is no formula such that both it and its negation can be derived from $\mathbf{X}$. For any complete proof method for propositional logic, this notion of consistency is equivalent to satisfiability:

Let $\mathbf{X}$ be a set of formulas. If $\mathbf{X}$ is not consistent then there is a formula G such that both G and $\neg G$ can be derived from $\mathbf{X}$. By soundness, $\mathbf{X} \Longrightarrow G$ and $\mathbf{X} \Longrightarrow \neg G$. So, by the definition of logical consequence, if there is a valuation v such that $v \models \mathbf{X}$, then both $v \models G$ and $v \models \neg G$ which can not be, so there is no valuation that satisfies $\mathbf{X}$, so $\mathbf{X}$ is not satisfiable. Conversely, if $\mathbf{X}$ is not satisfiable then let G be any formula. Then $\mathbf{X} \Longrightarrow G$ (and similarly $\mathbf{X} \Longrightarrow \neg G$) since the condition for logical consequence holds vacuously because there is no valuation v such that $v \models \mathbf{X}$. By completeness, G and $\neg G$ can both be derived from $\mathbf{X}$. So $\mathbf{X}$ is inconsistent.

Consistent is sometimes used to mean satisfiable. For example, Barwise [1977] (p. 25) defines "consistent" like this. We could speak of *semantic consistency* and *syntactic consistency*, then as we have just seen, for a complete proof method they are equivalent. And in the literature on distributive normal forms, a *consistent* formula means one that is satisfiable, so I will also use "consistent" in this sense.

1.2 Propositional disjunctive normal form

There are a number of different approaches to proving the existence of disjunctive normal form. A very intuitive approach is the use of truth-tables, which can be formalized by using partial valuations (defined in the previous section). I use an approach which is more similar to the one used later for the first-order case.

A *propositional constituent over a finite non-empty subset* $\{p_i\}_{i\in I}$ *of the propositional variables* is a formula of the form $\bigwedge_{i\in I}(\pm)p_i$.

Lemma 1.1

1. *For each finite non-empty subset* $\mathbf{P}$ *of the propositional variables, for each valuation, exactly one propositional constituent over* $\mathbf{P}$ *is true.*

2. *Each propositional constituent is satisfied by some valuation.*

PROOF

1. The propositional constituents over $\mathbf{P}$ are the basic conjunctions generated by some set of formulas, so for each valuation, exactly one of them is true (shown on p. 7).

2. Given some propositional constituent $C = \bigwedge_{i\in I}(\pm)p_i$, define a valuation v by

$$v(p_i) = \begin{cases} 0 & \text{if } i \in I \text{ and } \neg p_i \text{ is a conjunct of } C \\ 1 & \text{if } i \in I \text{ and } p_i \text{ is a conjunct of } C \\ 1 & \text{if } i \notin I. \end{cases}$$

   Then $v \models C$. □

Theorem 1.2 *For any propositional language, for every formula* F*, for every finite superset* $\mathbf{P}$ *of the variables in* F*, there is a unique disjunction of propositional constituents over* $\mathbf{P}$ *to which* F *is equivalent.*

PROOF We first show the existence of the required disjunction. Let F be a formula and $\mathbf{P}$ a finite superset of the variables in F and $\{C_j\}_{j\in J}$ the set of propositional constituents over $\mathbf{P}$. We use induction on the formation of the formula. If F is a propositional variable p_i, then $F \iff \bigvee_{j\in K} C_j$ where $\{C_j\}_{j\in K}$ is the set of propositional constituents over $\mathbf{P}$ that contain p_i as a conjunct. If F is $G \vee H$ and $G \iff \bigvee_{j\in K} C_j$ and $H \iff \bigvee_{j\in L} C_j$ for some $K, L \subseteq J$, then $F \iff \bigvee_{j\in K\cup L} C_j$. If F is $\neg G$ and $G \iff \bigvee_{j\in K} C_j$ for some $K \subseteq J$, then $F \iff \bigvee_{j\in J\setminus K} C_j$. All of the above cases are immediate from lemma 1.1, part 1 and the definition of a valuation.

To show uniqueness, if F_1 and F_2 are each disjunctions of some of the C_j's and F_1 includes some disjunct C_i that F_2 doesn't, then there is some valuation v such that $v \models C_i$ by lemma 1.1, part 2, so $v \models F_1$, and for all $j \neq i$, $v \not\models C_j$ by lemma 1.1, part 1, so $v \not\models F_2$. So $F_1 \not\Leftrightarrow F_2$. □

For a formula F and a finite superset $\mathbf{P}$ of the variables in F, by theorem 1.2 there is a unique disjunction of propositional constituents over $\mathbf{P}$ to which F is equivalent. This disjunction is called the *full disjunctive normal form relative to* $\mathbf{P}$ of F. Any disjunction of propositional constituents, not necessarily all over the same subset of variables, to which F is equivalent is called a *disjunctive normal form* of F.

As a special case of the above theorem, for a language with only finitely many propositional variables, every formula has a full disjunctive normal form relative to the set of all variables.

Although we already have a decision procedure for propositional logic, for comparison with later cases we note that converting a formula to disjunctive normal form provides another decision procedure.

Theorem 1.3 *A formula whose variables are contained in* $\mathbf{P}$ *is (propositionally) valid iff its full disjunctive normal form relative to* $\mathbf{P}$ *contains all the constituents over* $\mathbf{P}$.

Proof The disjunction all constituents over $\mathbf{P}$ is valid by lemma 1.1 (part 1). Any disjunction of constituents over $\mathbf{P}$ which doesn't contain them all is not valid by lemma 1.1 (parts 1 and 2). □

Theorem 1.2 provides an algorithm for converting a formula to disjunctive normal form, which theorem 1.3 then provides a means of testing for validity.

1.3 Modal logic

Modal syntax

There are many different kinds of modal logics. I consider just the simplest because it provides a stepping-stone between the propositional and first-order cases. In particular, it introduces the form of first-order constituents in a setting which is considerably simpler than the first-order case, with the result that all modal constituents are consistent (unlike the first-order case).

A modal language is an extension of a propositional language. The *modal languages* we will consider are those which contain the following symbols

- logical connectives: $\neg$ $\vee$
- modal operator: $\Diamond$
- a non-zero countable number of propositional variables
- parentheses: ( )

The set of propositional variables of a modal language $\mathcal{L}$ will be denoted $\mathbf{Var}_{\mathcal{L}}$ (or **Var** if it is not necessary to specify the language) as in the propositional case.

The *formulas over a subset* **P** *of* **Var** are defined as follows.

- An element of **P** is a formula over **P**.
- If A and B are formulas over **P** then $\neg A$ and $(A \vee B)$ and $\Diamond A$ are formulas over **P**.

The *formulas of a modal language* $\mathcal{L}$ are the formulas over $\mathbf{Var}_{\mathcal{L}}$, the set of which is denoted $\mathbf{Form}_{\mathcal{L}}$ (or **Form**).

The additional logical connectives are defined as for the propositional case. We use the same conventions regarding when parentheses may be omitted in writing formulas as in the propositional case.

Another modal operator is defined as a means of abbreviation:

- $\Box A$ abbreviates $\neg\Diamond\neg A$.

Modal semantics

A *frame* is a pair $\mathcal{F} = \langle W, R \rangle$ where

- W is a set, called the *universe* of $\mathcal{F}$, the elements of which are called *worlds*. (They should not be confused with the possible worlds described in chapter 2, though there are some similarities between them.)
- $R \subseteq W^2$ is called the *accessibility relation* of $\mathcal{F}$.

A *model* for a modal language $\mathcal{L}$ is a triple $\mathcal{M} = \langle W, R, V \rangle$ where

- $\langle W, R \rangle$ is a frame
- $V \subseteq W \times \mathbf{Var}_{\mathcal{L}}$ is called a *valuation on* W.

A model $\langle W, R, V\rangle$ is said to be a model over the frame $\langle W, R\rangle$. A frame $\langle W, R\rangle$ is said to have any property that the relation R has, such as reflexivity.

A *submodel* of a model $\langle W, R, V\rangle$ is a model $\langle W_i, R_i, V_i\rangle$ such that $W_i \subseteq W$, R_i is R restricted to W_i, and V_i is V restricted to W_i. A set of submodels $\{\langle W_i, R_i, V_i\rangle \mid i \in I\}$ of a model is said to be *non-overlapping* if for each distinct $i, j \in I$, $W_i \cap W_j = \emptyset$.

For a model $\mathcal{M} = \langle W, R, V\rangle$, the valuation $V \subseteq W \times \mathbf{Var}$ determines the truth values of all the propositional variables (and thus of all formulas not containing modal operators) at each element of W. The truth values of formulas containing modal operators depend also on the relation R. The full definition of the truth values is given by the relation $\models$ defined by

- $\mathcal{M}, w \models p$ iff $\langle w, p\rangle \in V$ for $p \in \mathbf{Var}$
- $\mathcal{M}, w \models \neg A$ iff $\mathcal{M}, w \not\models A$
- $\mathcal{M}, w \models A \vee B$ iff $\mathcal{M}, w \models A$ or $\mathcal{M}, w \models B$
- $\mathcal{M}, w \models \Diamond A$ iff there is some $v \in W$ with $\langle w, v\rangle \in R$ and $\mathcal{M}, v \models A$.

If $\mathcal{M}, w \models A$ then we say that A *is true at* w *in* $\mathcal{M}$. If $\mathcal{M}, w \models X$ for all $X \in \mathbf{X}$ then we write $\mathcal{M}, w \models \mathbf{X}$.

A formula A is *valid on a model* $\mathcal{M}$ if, for all worlds w in $\mathcal{M}$, it is the case that $\mathcal{M}, w \models A$ (denoted $\mathcal{M} \models_m A$). A formula A is *valid on a frame* $\mathcal{F}$ if, for all models $\mathcal{M}$ over $\mathcal{F}$, A is valid on $\mathcal{M}$ (denoted $\mathcal{F} \models_f A$).

For this simplest modal language, different modal logics are obtained by considering different classes of frames.

I will be considering the *minimal modal logic* **K** as defined in van Benthem [1983].

A formula is **K**-*valid* if it is valid on all frames (or equivalently, if it is valid on all models).

As examples of other modal logics: A formula is **T**-*valid* if it is valid on all reflexive frames. A formula is **S4**-*valid* if it is valid on all reflexive transitive frames.

A formula is called **K**-*satisfiable* (or **K**-*consistent*) if there is some world w in some model $\mathcal{M}$ such that A is true at w in $\mathcal{M}$.

There are a number of different notions of consequence for a modal logic. I mention two:

A frame-based notion of consequence for **K** is defined by: for $\mathbf{X}$ a formula or set of formulas and A a formula, $\mathbf{X} \Longrightarrow_{\mathbf{K},f} A$ iff for all frames $\mathcal{F}$ such that $\mathcal{F} \models_f \mathbf{X}$, also $\mathcal{F} \models_f A$. A corresponding model-based notion of consequence is defined by $\mathbf{X} \Longrightarrow_{\mathbf{K},m} A$ iff for all models $\mathcal{M}$ such that $\mathcal{M} \models_m \mathbf{X}$, also $\mathcal{M} \models_m A$. If $\mathbf{X} \Longrightarrow_{\mathbf{K},m} A$, then A is called a **K**-*consequence* of $\mathbf{X}$. If $\mathbf{X} \Longrightarrow_{\mathbf{K},m} A$ then $\mathbf{X} \Longrightarrow_{\mathbf{K},f} A$, but not conversely. For example, if p is a propositional variable, then $p \Longrightarrow_{\mathbf{K},f} (p \wedge \neg p)$, but $p \not\Longrightarrow_{\mathbf{K},m} (p \wedge \neg p)$.

Corresponding to each of these notions of consequence is a notion of equivalence where two formulas are equivalent if each implies the other: $A \Longleftrightarrow_{\mathbf{K},f} B$ iff $A \Longrightarrow_{\mathbf{K},f} B$ and $B \Longrightarrow_{\mathbf{K},f} A$; $A \Longleftrightarrow_{\mathbf{K},m} B$ iff $A \Longrightarrow_{\mathbf{K},m} B$ and $B \Longrightarrow_{\mathbf{K},m} A$. If $A \Longleftrightarrow_{\mathbf{K},m} B$ then A and B are said to be **K**-*equivalent.*

If for all worlds w in all models $\mathcal{M}$, $\mathcal{M}, w \models A$ iff $\mathcal{M}, w \models B$, then A and B are called *equivalent* (denoted $A \Longleftrightarrow B$). If two formulas are equivalent then they are **K**-equivalent, but not conversely. For example, if p is a propositional variable, then $p \Longleftrightarrow_{\mathbf{K},m} (\Box p \wedge p)$, but $p \not\Longleftrightarrow (\Box p \wedge p)$.

Metatheory

Van Benthem [1983] proves the following completeness theorem for **K**: There is a proof method such that for every set of formulas $\mathbf{X}$ and formula A, $\mathbf{X} \Longrightarrow_{\mathbf{K},m} A$ iff A can be derived from $\mathbf{X}$. And there is no complete proof method with the frame-based notion of consequence.

K is decidable (also shown in *ibid.*).

1.4 Modal disjunctive normal form

The following definitions are from Fine [1975], with terminology that agrees with that later used for the first-order case.

The *(modal) depth* (also called *degree*) of a formula is the length of the longest sequence of nested $\Diamond$'s it contains: where $d(A)$ denotes the depth of A, $d(p) = 0$ if p is a propositional variable; $d(\neg A) = d(A)$; $d(A \vee B) = \max\{d(A), d(B)\}$; $d(\Diamond A) = d(A) + 1$.

For each fixed finite subset $p_1, \ldots, p_n$ of the variables, a *modal constituent of depth d over* $p_1, \ldots, p_n$ is denoted C_i^d and defined by

$$\begin{aligned} C_i^0 &= (\pm)p_1 \wedge \ldots \wedge (\pm)p_n \\ C_i^d &= (\pm)p_1 \wedge \ldots \wedge (\pm)p_n \wedge \\ &\quad\ (\pm)\Diamond C_1^{d-1} \wedge \ldots \wedge (\pm)\Diamond C_q^{d-1} \end{aligned}$$

where $C_1^{d-1}, \ldots, C_q^{d-1}$ are all the modal constituents of depth $d-1$ over the variables $p_1, \ldots, p_n$, and i is the index by which the constituents are listed.

So for $d > 0$, a modal constituent C_i^d is of the form $C_j^0 \wedge (\pm)\Diamond C_1^{d-1} \wedge \ldots \wedge (\pm)\Diamond C_q^{d-1}$ for some depth-0 constituent C_j^0.

Lemma 1.4

1. *For each finite subset $p_1, \ldots, p_n$ of the propositional variables, for each model, at each world in the model exactly one modal constituent of depth d over $p_1, \ldots, p_n$ is true for each d.*

2. *Each modal constituent is true at some world in some model.*

PROOF

1. Each set of modal constituents of some depth over some set of variables is the set of formulas of the form $(\pm)X_1 \wedge \ldots \wedge (\pm)X_m$ for some formulas $X_1, \ldots, X_m$. For any world w in a model $\mathcal{M}$, the truth values of these constituents depend on the values of $X_1, \ldots, X_m$ as in the propositional case, thus exactly one such constituent is true.

2. We use induction on the depth of the constituent to prove that each constituent is true at some world w in some model $\mathcal{M} = \langle W, R, V \rangle$ such that there is no world v in W with $\langle v, w \rangle \in R$. Each constituent of depth 0 is true in a model containing only one world with a valuation chosen as in the propositional case. Assume that for some d, each of the constituents C_j^{d-1} $(j = 1, \ldots, q)$ is true at some world w_j in some model $\mathcal{M}_j = \langle W_j, R_j, V_j \rangle$ where there is no element v in W_j with $\langle v, w_j \rangle \in R_j$. For a depth-$d$ constituent $C_i^d = C_k^0 \wedge (\pm)\Diamond C_1^{d-1} \wedge \ldots \wedge (\pm)\Diamond C_q^{d-1}$ we construct a model $\mathcal{M} = \langle W, R, V \rangle$ as follows: for each j such that $\Diamond C_j^{d-1}$ occurs unnegated in C_i^d, let $\mathcal{M}_j$ be the model that exists by hypothesis. Let $\mathcal{M}$ be the model which contains each of these models $\mathcal{M}_j$ as non-overlapping submodels, and also contains one additional world w with $\langle w, w_j \rangle \in R$ for each of these j's, and V is defined at w according to C_k^0. Then, since by part 1 above only one formula of the form C_j^{d-1} is true at each world and each formula C_j^{d-1} is consistent, there are no dependencies between the formulas $\Diamond C_j^{d-1}$ where $j = 1, \ldots, q$, so C_i^d is true at w in $\mathcal{M}$. □

Lemma 1.5 *If $\mathbf{X}$ is a finite set of formulas, then each element X_i of $\mathbf{X}$ is equivalent to some disjunction of basic conjunctions generated by $\mathbf{X}$.*

PROOF As in the propositional case, this is just expressing X_i in the propositional disjunctive normal form with the elements of $\mathbf{X}$ for the propositional variables: If $\mathbf{X} = \{X_1, \ldots, X_n\}$ then for any world in any model the values of the basic conjunctions generated by $\mathbf{X}$ depend only on the values of $X_1, \ldots, X_n$, and are determined as in the propositional case, so X_i is equivalent to the disjunction of all the formulas of the form $(\pm)X_1 \wedge \ldots \wedge X_i \wedge \ldots \wedge (\pm)X_n$ (the basic conjunctions generated by $\mathbf{X}$ in which X_i occurs unnegated). □

Theorem 1.6 *Every modal formula of depth $\leq d$ is equivalent to a unique disjunction of modal constituents of depth d over any finite superset of the variables contained in the formula.*

PROOF Throughout this proof the constituents are over an arbitrary fixed finite subset of the variables and the formulas are those whose variables are contained in this set. We use induction on the depth d. For depth 0 we just have the propositional case which is proved by theorem 1.2. Assume that for some $d \geq 1$, any formula of depth $< d$ is equivalent to a disjunction of modal constituents of depth $d-1$. Let $\{C_j^d\}_{j\in J}$ be the set of constituents of depth d. Let A be any formula of depth $\leq d$. We use induction on the formation of the formula. If A is a variable p_i then the formulas of the form C_j^d are the basic conjunctions generated by a set which contains p_i, so A is equivalent to a disjunction of formulas of the form C_j^d by lemma 1.5. If A is $B \vee C$ and $B \Longleftrightarrow \bigvee_{j\in K} C_j^d$ and $C \Longleftrightarrow \bigvee_{j\in L} C_j^d$ then $A \Longleftrightarrow \bigvee_{j\in K\cup L} C_i^d$. If A is $\neg B$ and $B \Longleftrightarrow \bigvee_{j\in K} C_j^d$ then $A \Longleftrightarrow \bigvee_{j\in J\setminus K} C_j^d$. If A is $\Diamond B$ then B is of depth $< d$ so by hypothesis B is equivalent to some disjunction $\bigvee_{i\in I} C_i^{d-1}$ of constituents of depth $d-1$, so A is equivalent to $\bigvee_{i\in I} \Diamond C_i^{d-1}$ since $\Diamond$ distributes over disjunction. Each $\Diamond C_i^{d-1}$ is equivalent to the disjunction of all the C_j^d's in which $\Diamond C_i^{d-1}$ occurs unnegated (by lemma 1.5). Thus A is equivalent to a disjunction of formulas of the form C_j^d.

Uniqueness follows from the **K**-consistency of constituents as in the propositional case (replacing "some valuation" with "some world in some model"). □

For a modal formula F of depth $\leq d$ and a finite superset $\mathbf{P}$ of the variables in F, the disjunction of modal constituents of depth d over $\mathbf{P}$ to which F is equivalent is called the *disjunctive normal form of depth d relative to* $\mathbf{P}$ of F. If the set of variables to which a disjunctive normal form is relative is not mentioned, it can be taken to be any fixed finite superset of the variables in the formula.

As a consequence of the above theorem, every modal formula of depth $\leq d$ is **K**-equivalent to a disjunction of modal constituents of depth d. But I don't know if there is a *unique* disjunction of modal constituents of depth d to which it is **K**-equivalent.

Fine [1975] also shows that the modal logics **D**, **T**, **K4** have similar normal forms.

Theorem 1.7 *A modal formula A of depth d is* **K***-valid iff its disjunctive normal form at depth d contains all constituents of depth d.*

PROOF The disjunction of all constituents of depth d is **K**-valid by lemma 1.4 (part 1). Any disjunction of depth-d constituents which doesn't contain them all is not **K**-valid by lemma 1.4 (parts 1 and 2). □

Theorem 1.7 provides us with a decision procedure for **K**: The proof of theorem 1.6 provides an algorithm for converting any formula to a disjunctive normal form which is **K**-valid iff the original formula is **K**-valid. The **K**-validity of the disjunctive normal form is then tested by theorem 1.7.

1.5 First-order logic

First-order syntax

As with propositional languages, first-order languages may be defined in various different ways. I use the following definition.

A *first-order language* $\mathcal{L}$ contains the following symbols:

- logical connectives: $\neg$ $\vee$
- existential quantifier: $\exists$
- variables: x_i for each $i \in \mathbb{N}$, with the set $\{x_i\}_{i\in\mathbb{N}}$ denoted **Var**
- a countable number of constant symbols, the set of which is denoted $\mathbf{Const}_{\mathcal{L}}$
- a countable number of function symbols, each of which is associated with some element of $\mathbb{N}$ called its *arity*
- a non-zero countable number of predicate symbols, each of which is associated with some element of $\mathbb{N}$ called its *arity*
- punctuation: ( ) ,

The constant, function and predicate symbols vary from language to language. A function or predicate symbol of arity n is called *n-ary*. Instead of 1-ary we may say *unary*, and instead of 2-ary *binary*.

The *terms* of a language are defined by:

- Constant symbols and variables are terms.
- If $t_1, \ldots, t_n$ are terms and f is an n-ary function symbol, then $f(t_1, \ldots, t_n)$ is a term.

The set of terms of a language $\mathcal{L}$ is denoted $\mathbf{Term}_{\mathcal{L}}$.

The *atomic formulas over a subset* **S** *of* $\mathbf{Term}_{\mathcal{L}}$ are defined by:

- If $s_1, \ldots, s_n \in \mathbf{S}$ and P is an n-ary predicate symbol, then $P(s_1, \ldots, s_n)$ is an atomic formula over **S**.

The *formulas* of a first-order language $\mathcal{L}$ are defined by:

- An atomic formula over $\mathbf{Term}_{\mathcal{L}}$ is a formula.
- If F, G are formulas and x_i is a variable, then $\neg F$, $(F \vee G)$, $(\exists x_i)F$ are formulas.

The set of formulas of $\mathcal{L}$ is denoted $\mathbf{Form}_{\mathcal{L}}$.

The additional logical connectives are defined as for the propositional case. We use the same conventions regarding when parentheses may be omitted in writing formulas as in the propositional case. Where it is not necessary to specify the language we may write **Term** or **Form**.

The universal quantifier is defined as an abbreviation:

- $(\forall x_i)F$ abbreviates $\neg(\exists x_i)\neg F$.

If $(\exists x_i)F$ is a formula, then F is called the *scope* of $(\exists x_i)$. Any occurrence of a variable x_i in the scope of a quantifier over x_i or immediately preceded by a quantifier is called a *bound* occurrence of the variable. Any occurrence of a variable which is not bound is *free*. A formula with no free variables is called a *sentence*.

First-order semantics

A *model* for a first-order language $\mathcal{L}$ is a pair $\mathcal{M} = \langle D, I\rangle$ where

- D is a set, called the *domain* (or *universe*) of $\mathcal{M}$. Elements of D are called *individuals* of $\mathcal{M}$.
- I is a function called an *interpretation* that maps
  - every constant symbol of $\mathcal{L}$ to some element of D,
  - for each $n \in \mathbb{N}$, every n-ary function symbol of $\mathcal{L}$ to some n-ary operation on D,
  - for each $n \in \mathbb{N}$, every n-ary predicate symbol of $\mathcal{L}$ to some n-ary relation on D.

For a model $\langle D, I\rangle$ for a language $\mathcal{L}$, a *valuation* v is a function from $\mathbf{Term}_{\mathcal{L}}$ to D which satisfies

- if c is a constant symbol, $v(c) = I(c)$
- if f is an n-ary function symbol and $t_1, \ldots, t_n$ are terms,
  $v(f(t_1, \ldots, t_n)) = I(f)(v(t_1), \ldots, v(t_n))$.

A valuation is uniquely determined by the values it takes on **Var**, and can also be considered as a function $\mathbf{Var} \to D$ extended to a function $\mathbf{Term}_{\mathcal{L}} \to D$ by the above condition considered as a definition. In giving definitions of particular valuations it is sufficient to give the values they take on **Var**.

If $v : \mathbf{Var} \to D$ is a valuation, $a \in D$, and $x \in \mathbf{Var}$, then $v(a/x)$ denotes the valuation $v' : \mathbf{Var} \to D$ given by

$$v'(y) = \begin{cases} v(y) & \text{if } y \neq x \\ a & \text{if } y = x. \end{cases}$$

A model $\mathcal{M} = \langle D, I \rangle$ and valuation v for $\mathcal{M}$ together determine the truth values of all formulas by the relation $\models$ defined by

- $\mathcal{M}, v \models P(t_1, \ldots, t_n)$ iff $\langle v(t_1), \ldots, v(t_n) \rangle \in I(P)$
- $\mathcal{M}, v \models \neg F$ iff $\mathcal{M}, v \not\models F$
- $\mathcal{M}, v \models F \vee G$ iff $\mathcal{M}, v \models F$ or $\mathcal{M}, v \models G$
- $\mathcal{M}, v \models (\exists x_i)F$ iff there is some $a \in D$ such that $\mathcal{M}, v(a/x_i) \models F$.

If $\mathcal{M}, v \models F$ then we say that F *is true in* $\mathcal{M}$ *under* v.

For a model $\mathcal{M}$ and a formula F, for each valuation v for $\mathcal{M}$, whether or not $\mathcal{M}, v \models F$ depends only on the values to which v maps the free variables of F. In particular, if F has no free variables, then either $\mathcal{M}, v \models F$ for all valuations v for $\mathcal{M}$ or for no valuations v for $\mathcal{M}$.

If $\mathcal{M}, v \models F$ for all valuations v for the model $\mathcal{M}$, we write $\mathcal{M} \models F$ and F is called *true in the model* $\mathcal{M}$, and $\mathcal{M}$ is called a *model of* F. If $\mathbf{X}$ is a set of formulas and $\mathcal{M} \models X_i$ for each $X_i \in \mathbf{X}$, then $\mathcal{M}$ is called a *model of* $\mathbf{X}$, and we write $\mathcal{M} \models \mathbf{X}$. If $\mathcal{M} \models F$ for all models $\mathcal{M}$, we write $\models F$, and F is called *valid* (or *logically valid*).

There are two different notions of satisfiability for first-order logic. One definition (found e.g. in Shoenfield [1967]) requires that a formula be true in some model (i.e. for every valuation for that model) to be called satisfiable. The definition I use (found in Fitting [1990], Smullyan [1968] and Bell and Machover [1977]) only requires the formula to be true for some valuation for some model: A formula F is called *satisfiable* if there is some model $\mathcal{M}$ and some valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models F$. A set of formulas $\mathbf{X}$ is *satisfiable* if there is some model $\mathcal{M}$ and some valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models X_i$ for each $X_i \in \mathbf{X}$. To show that these two notions of satisfiability are different, the formula $((\exists x)\neg P(x, x)) \wedge P(x, y)$ is satisfiable if it is not required to hold for all the valuations that map x and y to the same element, which is the second sense of satisfiable, but it is not satisfiable in the first sense.

As in the propositional case, *consistent* will be used as a synonym for satisfiable. The relation between this sense of consistency and the proof-theoretic sense is discussed in the next section.

There is a notion of tautology for first-order logic which is different from that of a (first-order) valid formula. A formula of a first-order language $\mathcal{L}$ is called a *tautology* if it is a propositional tautology in the propositional language whose variables are the atomic formulas and formulas that start with quantifiers of $\mathcal{L}$. This definition, which is equivalent to one found in Smullyan [1968], seems to be related to Hintikka's discussion of tautology in Hintikka [1973b].

There are also two different notions of logical consequence for first-order logic. One (found e.g. in Shoenfield [1967]) is defined by: a formula F is a logical consequence of a formula or set of formulas $\mathbf{X}$ iff every model of $\mathbf{X}$ is also a model of F. The definition I use (found e.g. in Bell and Machover [1977]) is: F is a *logical consequence* of $\mathbf{X}$ iff for every model $\mathcal{M}$ and valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models \mathbf{X}$, also $\mathcal{M}, v \models F$, which I denote by $\mathbf{X} \Longrightarrow F$. For formulas F and G, this definition gives $F \Longrightarrow G$ iff $\models F \to G$. This is immediate from the definitions: Writing "$(\forall \mathcal{M}, \forall v)$" for "for every model $\mathcal{M}$ and every valuation v for $\mathcal{M}$", we get

$$
\begin{aligned}
F \Longrightarrow G \quad & \text{iff} \quad (\forall \mathcal{M}, \forall v) \text{ if } \mathcal{M}, v \models F \text{ then } M, v \models G \\
& \text{iff} \quad (\forall \mathcal{M}, \forall v) \mathcal{M}, v \not\models F \text{ or } \mathcal{M}, v \models G \\
& \text{iff} \quad (\forall \mathcal{M}, \forall v) \mathcal{M}, v \models \neg F \text{ or } \mathcal{M}, v \models G \\
& \text{iff} \quad (\forall \mathcal{M}, \forall v) \mathcal{M}, v \models \neg F \vee G \\
& \text{iff} \quad (\forall \mathcal{M}, \forall v) \mathcal{M}, v \models F \to G \qquad \text{iff} \ \models F \to G.
\end{aligned}
$$

If $F \Longrightarrow G$ then we may say F *implies* G.

By a similar proof to the one above (and the fact that for any model and valuation there is some formula that is true), a formula G is valid iff for all formulas F, $F \Longrightarrow G$.

Corresponding to each notion of logical consequence is a notion of logical equivalence, where two formulas are equivalent if each is a consequence of the other. For the notion of logical consequence which I have denoted $\Longrightarrow$, the corresponding notion of logical equivalence of formulas F and G is that for every model $\mathcal{M}$ and valuation v for $\mathcal{M}$, it is the case that $\mathcal{M}, v \models F$ iff $\mathcal{M}, v \models G$, which I denote by $F \Longleftrightarrow G$. This notion of logical equivalence is also referred to simply as *equivalence*. If two formulas are equivalent, we can also say that one *can be expressed* (or *written*) *as* the other, or that they have *the same meaning*.

Since conjunction and disjunction are commutative and associative relative to the first-order notion of logical equivalence just defined, we can use the same abbreviations for finite conjunctions and disjunctions as defined for the propositional case.

The following notation is introduced to make the formulas used in the study of distributive normal forms easier to write.

For first-order formulas we use the same definition of *basic conjunctions generated* by a finite non-empty set $\mathbf{X}$ of formulas as in the propositional case. We also introduce the notation $\Delta\mathbf{X}$ for this set. So if $\mathbf{X} = \{X_1, \ldots, X_n\}$, then $\Delta\mathbf{X} = \{(\pm)X_1 \wedge \ldots \wedge (\pm)X_n\}$. (The Δ notation is from Rantala [1987], though I have defined it differently.)

Sets of formulas of the form $\Delta\mathbf{X}$ have the following similar property to the propositional case: For each model $\mathcal{M}$ and each valuation v for $\mathcal{M}$ exactly one element of $\Delta\mathbf{X}$ is true. It is the formula F such that for each $X_i \in \mathbf{X}$, if $\mathcal{M}, v \models X_i$ then X_i is a conjunct of F and if $\mathcal{M}, v \not\models X_i$ then $\neg X_i$ is a conjunct of F. The proof is as for the propositional case.

Also, as for the propositional case, if $\mathbf{X}$ is a finite set of first-order formulas, then for each element X_i of $\mathbf{X}$, there is at least one disjunction of basic conjunctions generated by $\mathbf{X}$ to which X_i is logically equivalent.

For sets $\mathbf{A}$ and $\mathbf{B}$ of formulas, the set $\bigwedge(\mathbf{A}, \mathbf{B}) = \{A \wedge B \mid A \in \mathbf{A}, B \in \mathbf{B}\}$.

For a set $\mathbf{A}$ of formulas and a variable x, the set $\exists x(\mathbf{A}) = \{(\exists x)A \mid A \in \mathbf{A}\}$.

Metatheory

As for propositional logic, there are notions of completeness and weak completeness for first-order logic. In addition, the two definitions of logical consequence give different notions of completeness. Weak completeness is not affected by the definition of logical consequence, and is proved for example in Fitting [1990]: There is a proof method such that for every formula F, $\models F$ iff F is a theorem. Although the definition of weak completeness does not involve logical consequence, the following condition holds for weakly complete proof methods: For any formulas F and G, since $F \Longrightarrow G$ iff $\models F \rightarrow G$ (shown on p. 20), we have $F \rightarrow G$ is a theorem iff $F \Longrightarrow G$.

The equivalence of satisfiability and the proof-theoretic sense of consistency that we have seen for propositional logic may be generalized to first-order logic. We get two different generalizations using different definitions of satisfiability and logical consequence. To state these results, I introduce the following terminology and notation: A formula F is called *satisfiable*$_1$ if there is some model $\mathcal{M}$ such that $\mathcal{M} \models F$. For a formula or set of formulas $\mathbf{X}$ and a formula F, $\mathbf{X} \Longrightarrow_1 F$ iff for all models $\mathcal{M}$ such that $\mathcal{M} \models \mathbf{X}$, also $\mathcal{M} \models F$. A proof method is called *complete* if (F can be derived from $\mathbf{X}$ iff $\mathbf{X} \Longrightarrow F$), and *complete*$_1$ if (F can be derived from $\mathbf{X}$ iff $\mathbf{X} \Longrightarrow_1 F$). For any proof method, a formula or set of formulas $\mathbf{X}$ is called *proof-theoretically consistent* if there is no formula such that both it and its negation can be derived from $\mathbf{X}$.

Bell and Machover [1977] define a proof method for first-order logic (p. 108–109) which is proved to be complete (p. 109–121). Rogers [1971] (p. 43–68) defines a proof method and shows it to be complete$_1$. So, there are proof methods for first-order logic which are complete, and there are other proof methods (not equivalent to the complete ones) which are complete$_1$.

For any complete proof method for first-order logic, a set of formulas is proof-theoretically consistent iff it is satisfiable:

Let $\mathbf{X}$ be a set of formulas. If $\mathbf{X}$ is not proof-theoretically consistent then there is a formula F such that both F and $\neg F$ can be derived from $\mathbf{X}$. By soundness $\mathbf{X} \Longrightarrow F$ and $\mathbf{X} \Longrightarrow \neg F$. So, by the definition of $\Longrightarrow$, if there is some model $\mathcal{M}$ and some valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models \mathbf{X}$, then both $\mathcal{M}, v \models F$ and $\mathcal{M}, v \models \neg F$ which can not be, so $\mathbf{X}$ is not satisfiable. Conversely, if $\mathbf{X}$ is not satisfiable then let F be any formula. Then $\mathbf{X} \Longrightarrow F$ (and similarly $\mathbf{X} \Longrightarrow \neg F$) since there is no model $\mathcal{M}$ and valuation v such that $\mathcal{M}, v \models \mathbf{X}$, so the condition for logical consequence holds vacuously. By completeness, both F and $\neg F$ can be derived from $\mathbf{X}$, so $\mathbf{X}$ is not proof-theoretically consistent.

For any complete$_1$ proof method for first-order logic, a set of formulas is proof-theoretically consistent iff it is satisfiable$_1$:

Let $\mathbf{X}$ be a set of formulas. If $\mathbf{X}$ is not proof-theoretically consistent then there is a formula F such that both F and $\neg F$ can be derived from $\mathbf{X}$. By sound$_1$ness $\mathbf{X} \Longrightarrow_1 F$ and $\mathbf{X} \Longrightarrow_1 \neg F$. So, by the definition of $\Longrightarrow_1$, if there is some model $\mathcal{M}$ such that $\mathcal{M} \models \mathbf{X}$, then both $\mathcal{M} \models F$ and $\mathcal{M} \models \neg F$ which can not be, so $\mathbf{X}$ is not satisfiable$_1$. Conversely, if $\mathbf{X}$ is not satisfiable$_1$ then let F be any formula. Then $\mathbf{X} \Longrightarrow_1 F$ (and similarly $\mathbf{X} \Longrightarrow_1 \neg F$) since there is no model $\mathcal{M}$ such that $\mathcal{M} \models \mathbf{X}$, so the condition for $\Longrightarrow_1$ holds vacuously. By complete$_1$ness, both F and $\neg F$ can be derived from $\mathbf{X}$, so $\mathbf{X}$ is not proof-theoretically consistent.

In the following chapters, I use *consistent* as a synonym for satisfiable.

There are first-order languages for which the logic is not decidable. If the language contains

- at least one n-ary predicate symbol for some $n > 1$, or
- at least one n-ary function symbol for some $n > 1$, or
- more than one unary function symbol

then the logic is undecidable. If it contains only unary predicate symbols and constants, but no function symbols, then the logic is decidable (Hodel [1995], p. 208–209).

Even if there is no algorithm which can decide for all formulas of the language whether or not they are valid, there are some subsets of formulas for which there are algorithms which can decide whether or not any formula in the set is valid. Such sets of formulas are called *decidable*.

1.6 Monadic first-order disjunctive normal form

A first-order language which contains only unary predicate symbols and no function or constant symbols is called a *monadic (first-order) language*.

Let x be a variable which will remain fixed for this discussion. For each fixed finite subset $P_1, \ldots, P_n$ of predicate symbols, a *(monadic) attributive constituent relative to* $P_1, \ldots, P_n$ is denoted $A_i(x)$ and defined by

$$A_i(x) = (\pm)P_1(x) \wedge \ldots \wedge (\pm)P_n(x)$$

where $i = 1, \ldots, 2^n$ is the index by which the attributive constituents are listed. A *(monadic) constituent relative to* $P_1, \ldots, P_n$ is denoted C_i and defined by

$$C_i = (\pm)(\exists x)A_1(x) \wedge \ldots \wedge (\pm)(\exists x)A_q(x)$$

where $A_1(x), \ldots, A_q(x)$ are all the attributive constituents relative to $P_1, \ldots, P_n$, and $i = 1, \ldots, 2^q$ is the index by which the constituents are listed.

The terminology of *attributive constituent* and *constituent* has been chosen because of what these formulas describe. An attributive constituent describes a kind of individual by saying what properties it has. A constituent then says what kinds of individuals exist. This idea is generalized to the whole of first-order logic in the next two chapters.

Lemma 1.8

1. *For each finite subset* $P_1, \ldots, P_n$ *of predicate symbols, for each model, for each valuation, exactly one monadic constituent relative to* $P_1, \ldots, P_n$ *is true.*

2. *Each monadic constituent is consistent.*

Proof

1. The monadic constituents relative to some set of predicate symbols are the basic conjunctions generated by some set of formulas, so exactly one such constituent is true for any given model and valuation (shown on p. 21).

2. Each attributive constituent is consistent: If $A_i(x) = (\pm)P_1(x) \wedge \ldots \wedge (\pm)P_n(x)$, let $\mathcal{M}$ be a model whose domain contains one element e and whose interpretation I is defined by: for each $j \in \{1, \ldots, n\}$, if $P_j(x)$ is a conjunct of $A_i(x)$ then $I(P_j) = \{\langle e \rangle\}$

and if $\neg P_j(x)$ is a conjunct of $A_i(x)$ then $I(P_j) = \emptyset$ (and to complete the definition, if P is any predicate symbol not in $\{P_1, \ldots, P_n\}$ let $I(P) = \emptyset$). Let v be the valuation that maps x to e. Then $\mathcal{M}, v \models A_i(x)$. For any model $\mathcal{M}$, for each element a of the domain, $\mathcal{M}, v(a/x) \models A_i(x)$ for exactly one i because the formulas of the form $A_i(x)$ are the basic conjunctions generated by some set of formulas, so exactly one of them is true for a given model and valuation. (This can be expressed informally as: each a satisfies exactly one $A_i(x)$.) And each $A_i(x)$ is satisfiable. So the formulas of the form $(\exists x)A_i(x)$ can each be made true or false independently of the others by the existence or lack of it of a suitable individual a. □

The *depth* of a formula will be defined in full later (in chapter 2). The cases we need now are: A formula of depth 0 is one which does not contain any quantifiers. A formula of depth 1 is one which contains at least one quantifier, but no quantifier in the scope of another quantifier.

Lemma 1.9 *Every monadic formula F of depth 0 whose only variable is x is equivalent to a disjunction of monadic attributive constituents relative to any finite superset (of predicate symbols) of the predicate symbols that occur in F.*

PROOF A formula of depth 0 whose only variable is x is a Boolean combination of the formulas that generate the attributive constituents relative to any finite superset of the predicate symbols in the formula, and thus has a propositional disjunctive normal form which is a disjunction of these attributive constituents. □

Lemma 1.10 *Every sentence of a monadic first-order language is equivalent to a sentence of depth 1 whose only variable is x.*

PROOF Every sentence of depth 1 contains only one variable, which is bound, so by a change of bound variable if necessary is equivalent to a sentence of depth 1 whose only variable is x. Thus we need only show that each sentence is equivalent to one of depth 1. This is a special case of the result that each formula F of a monadic first-order language is equivalent to some formula with the same free variables as F in which no variable occurs in the scope of a quantifier over another variable (for which I will say there is no *irrelevant quantification*), which we prove by induction on the formation of the formula. If F is an atomic formula then F itself is the required equivalent formula. If F is $G \vee H$ and G', H' are equivalent to G, H respectively and have the same free variables as them and have no irrelevant quantification, then $G' \vee H'$ is equivalent to F and has the same free variables as F and has no irrelevant quantification. If F is $\neg G$ and G' is equivalent to G and has the same free variables as it and has no irrelevant quantification, then $\neg G'$ is equivalent to F and has the same free variables as it and no irrelevant quantification. If F is $(\exists y)G$ and G' is equivalent to G and has the same free variables as it and no irrelevant quantification, then transform G' propositionally, by using De Morgan's law as many times as is necessary, to get a formula H in which each negation sign occurs directly before an atomic formula or a quantifier. H has no irrelevant quantification (because G' did not), and H is of a form which I will call *intermediate*, defined by: An *intermediate* formula is a disjunction or conjunction where each disjunct or conjunct is an intermediate formula, or an atomic formula or negation of one, or a formula starting with a quantifier in which all atomic formulas contain the variable of quantification or a negation of such a formula. We use induction on the formation of intermediate formulas to show that for all intermediate formulas H, $(\exists y)H$ is equivalent to a formula without irrelevant quantification and with the same free variables as $(\exists y)H$. If H is an atomic formula or negation of one, $(\pm)A(z)$, then if z is y then $(\exists y)H$ itself has no irrelevant quantification and thus is the required formula. If z is not y then $(\exists y)H$ is equivalent to H and H has no irrelevant quantification and has the same free variables as $(\exists y)H$. If H is a formula starting with a quantifier in which all atomic formulas contain the variable of quantification or a negation of such a formula, $(\pm)(\exists z)I$, then $(\exists y)H$ is equivalent to H because the only variable which occurs in H is z. If H is $I_1 \vee I_2$ or $I_1 \wedge I_2$ then $(\exists y)H$ is equivalent to $J_1 \vee J_2$ or $J_1 \wedge J_2$ respectively, where for each $i \in \{1,2\}$, J_i is $(\exists y)I_i$ if y occurs in I_i and otherwise is I_i. For each i, if J_i is I_i then it doesn't have irrelevant quantification, here let K_i be J_i. If J_i is $(\exists y)I_i$ then by the inductive hypothesis J_i can be written as some formula K_i without irrelevant quantification and without changing the free variables. Then $(\exists y)H$ is equivalent to $K_1 \vee K_2$ or $K_1 \wedge K_2$ respectively. □

Theorem 1.11 *Every sentence of a monadic first-order language is equivalent to a unique disjunction of monadic constituents relative to any finite superset of the predicate symbols that occur in the formula.*

PROOF By lemma 1.10, we need only consider sentences of depth 1 whose only variable is x. Let F be such a sentence. We use induction on the formation of the sentence. The basic case is when F is $(\exists x)G$ for some formula G of depth 0 whose only variable is x, so by lemma 1.9 G is equivalent to some disjunction of attributive constituents $\bigvee_{j\in J} A_j$, so $(\exists x)G \iff (\exists x)\bigvee_{j\in J} A_j \iff \bigvee_{j\in J}(\exists x)A_j$ and each $(\exists x)A_j$ is equivalent to the disjunction of all the constituents C_i such that $(\exists x)A_j$ occurs unnegated in C_i. So F is equivalent to some disjunction of constituents. If F is $G \vee H$ and $G \iff \bigvee_{i\in I} C_i$ and $H \iff \bigvee_{i\in J} C_i$ then $F \iff \bigvee_{i\in I\cup J} C_i$. If F is $\neg G$ and $G \iff \bigvee_{i\in I} C_i$ then $F \iff \bigvee_{i\notin I} C_i$ by lemma 1.8 (proof is as for corollary 3.2 part 1). Uniqueness follows from the consistency of constituents as in the propositional case. □

For a monadic formula F and a finite superset $\mathbf{P}$ of the predicate symbols in F, the disjunction of monadic constituents relative to $\mathbf{P}$ to which F is equivalent is called the *disjunctive normal form (or distributive normal form) relative to* $\mathbf{P}$ of F.

Theorem 1.12 *A monadic formula F is valid iff its disjunctive normal form relative to any (or, equivalently, to all) finite superset(s) $\mathbf{P}$ of predicates containing those in F contains all the constituents relative to $\mathbf{P}$.*

PROOF The disjunction of all constituents relative to $\mathbf{P}$ is valid by lemma 1.8 (part 1). Any disjunction of constituents relative to $\mathbf{P}$ which doesn't contain them all is not valid by lemma 1.8 (parts 1 and 2). □

As for the propositional and modal cases, the above theorem provides a decision procedure for monadic first-order logic: The proof of theorem 1.11 provides an algorithm for converting a formula to disjunctive normal form, which is then tested for validity by theorem 1.12. The decidability of monadic first-order logic is shown by a different method in Boolos and Jeffrey [1974].

1.7 First-order logic with identity

Syntactically, first-order logic with identity is just a special case of first-order logic, but the semantics is defined using only a restricted set of models.

A *first-order language with identity* (also called a *first-order language with equality*) is a first-order language which contains the binary predicate symbol $=$. And we will assume that no first-order language contains the symbol $=$ unless it is meant as the identity symbol of a language with identity.

Terms and formulas are defined as for any first-order language. If s, t are terms, then the atomic formula $=(s,t)$ may also be written as $s = t$, and $\neg(s = t)$ may be written as $s \neq t$.

A *normal model* for a first-order language with identity is a (first-order) model $\mathcal{M} = \langle D, I \rangle$ such that $I(=)$ is the identity relation on D.

Valuations and the relation $\models$ are defined as for any (first-order) models. If F is true in all normal models, we write $\models_= F$, and F is called *valid with identity*. If there is some normal model $\mathcal{M}$ and some valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models F$, then F is *satisfiable* (or *consistent*) *with identity*.

For formulas F and G, there is the following notion of consequence: $F \Longrightarrow_= G$ iff for every normal model $\mathcal{M}$ and valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models F$, also $\mathcal{M}, v \models G$. If $F \Longrightarrow G$ then $F \Longrightarrow_= G$, but not conversely. We call F and G *equivalent with identity* and write $F \Longleftrightarrow_= G$ if $F \Longrightarrow_= G$ and $G \Longrightarrow_= F$.

For formulas F and G, $F \Longrightarrow_= G$ iff $\models_= F \to G$. This is proved just like the case for languages without identity, as on p. 20, except letting "$(\forall \mathcal{M}, \forall v)$" be short for "for every normal model $\mathcal{M}$ and every valuation v for $\mathcal{M}$".

For languages with identity, the definition of *notational variation* is extended so that $s = t$ and $t = s$ are notational variations of each other. So, for example, saying that $s = t$ is a conjunct of some formula means that either $s = t$ or $t = s$ is a conjunct.

First-order logic with identity is weakly complete: There is a proof method for which any formula F is valid with identity iff F is a theorem. Fitting [1990] gives both resolution and tableau proof methods that are weakly complete. So, since $F \Longrightarrow_= G$ iff $\models_= F \to G$, for any weakly complete proof method, $F \to G$ is a theorem iff $F \Longrightarrow_= G$.

There are first-order languages with identity for which the logic is not decidable (Hodel [1995]).

Chapter 2

An introductory example

This chapter first introduces some concepts that will be needed to define distributive normal forms and to explain intuitively what the formulas that occur in them represent, and then considers some of these formulas for a particular language.

Unfortunately the term "possible world" has been used with a variety of different meanings. I describe only the sense in which I use it. A *possible world* is a possible situation or state of affairs. For example, for a first-order language, a possible world may be specified by a set of individuals together with the values of the atomic formulas over these individuals. Such a world may be represented by the usual models of first-order logic, or by state-descriptions (defined below), or by any mathematical structure that gives sufficient information to determine the values of all sentences. We can think of a possible world as that which determines the values of all the sentences which are not truth-functionally determined. A possible world is an informal concept, and it is the representations of the possible worlds, and not the worlds themselves, that are defined precisely as some sort of mathematical entities.

Carnap [1956] (p. 9) defines *state-descriptions* for languages whose constant symbols correspond exactly to the elements of a universe. A state-description is a set of formulas of such a first-order language whose elements are the atomic formulas over the constant symbols, each either negated or not. State-descriptions fully describe possible worlds, but have the disadvantages that the language must contain names for each member of the universe, and that when the universe in infinite, so are they. It seems that any exhaustive description (in terms of formulas of some first-order language) of a possible world will have these same disadvantages. This is expressed by Hintikka [1973c] (p. 9) as "It is only natural that we cannot give an exhaustive description of a universe without knowing all its members, or of an infinite universe without using an infinity of sentences."

Two approaches to this problem are suggested in Hintikka [1973c] (p. 9–11, 18). One involves using sets of formulas that only give partial descriptions of possible worlds. This leads to the idea of *model sets* (defined in chapter 4). They achieve partial description by requiring consistency, but not requiring that they give the values of all the atomic formulas, as a state-description must. Where partial descriptions of possible worlds are sufficient, model sets are very useful. For example, construction of model sets provides a complete method of proof for first-order logic (Hintikka [1955], p. 37–45). However, they have the disadvantage that they are not mutually exclusive descriptions. Any type of description which is a partial description just by leaving out some information, as is the case for model sets, will have this disadvantage.

The other approach is an attempt to provide descriptions, which, while not full descriptions of possible worlds, nevertheless are mutually exclusive. Hintikka calls such descriptions 'descriptions of possible *kinds* of worlds'. They are achieved by giving as exhaustive descriptions of possible worlds as can be done with certain restricted resources of expression. In particular, the descriptions should be independent of the names of individuals of the universe, and the number of individuals that may be considered in their relation to each other should be limited. Without naming individuals, they can still be considered through the bound variables of quantification, and described in relation to other individuals by the predicates they satisfy. Thus, the existence, or lack of it, of certain *kinds* of individuals may be asserted, and a possible kind of world described. Restricting the number of individuals that may be considered at a time causes the descriptions to, in some sense, be partial. However, to satisfy the requirement of being as exhaustive as possible, they must say whether each kind of individual that can be described within the given limits exists or not. Hence such descriptions are mutually exclusive. A constituent is a certain generalization of this kind of description that may contain free variables.

Constituents differ from the other descriptions and partial descriptions of possible worlds that have been mentioned in that they are part of the language, without having to add extra symbols to it, and so, because they are finite, each constituent can be represented as a single formula of the language. But, any partial description of a possible world can be thought of as a description of a possible kind of world, so, for example, any sentence can be said to describe a possible kind of world. Constituents do not only describe kinds of worlds in this weakest possible sense of kind of world. We later define a notion of *depth* and show that (see lemma 3.1): For any fixed depth (and fixed set of free variables), exactly one constituent is true for each possible world (and valuation). In this sense constituents are the strongest (most informative) formulas at a fixed depth.

It was shown by Hintikka that every first-order formula can be written as a disjunction of constituents, called a *distributive normal form* of the formula. This corresponds with the idea that a sentence describes the set of possible worlds in which it is true. To deal with an infinite set containing *all* possible worlds may be difficult. In this case we have the slightly modified view that a sentence describes the set of possible kinds of worlds in which it is true. Each possible kind of world is described by a constituent, so a disjunction of constituents describes a set of possible kinds of worlds.

A sequence of constituents, each of which implies the previous one, may be thought of as a sequence of descriptions of a possible world (or model), each of which includes all the information of the previous one. Such a sequence can be said to describe a possible world at deeper and deeper *depths*. The constituents that may occur in the distributive normal form of a formula must be, in some sense, just as complex as the formula, which is that they must be at least as deep as the formula. Hintikka [1973a] (p. 141) gives an intuitive explanation of this depth as the maximum number of individuals that need to be considered in their relation to each other. For any formula, individuals may be referred to by free variables or constants, and in addition, for each quantifier we need to consider one individual. So a first approximation to the maximum number of individuals that need to be considered at one time is the sum of the number of free variables and constants, and the maximum number of quantifiers whose scopes overlap. This sum is called the *degree* of the formula. The maximum number of quantifiers whose scopes overlap is called the *depth* of a formula by Hintikka (though he further refines the definition of depth), and is called the *quantifier rank* by Rantala [1987] (p. 44–45). The quantifier rank of a formula F is denoted $qr(F)$ and defined by: $qr(F) = 0$ if F is atomic; $qr(\neg F) = qr(F)$; $qr(F \vee G) = \max\{qr(F), qr(G)\}$; $qr((\exists x)F) = qr(F)+1$. This definition of quantifier rank does not always give the number of individuals that need to be considered in their relation to each other that are introduced by quantifiers, since the nesting of quantifiers is not a sufficient condition for the individuals that they introduce to be related. Hintikka [1973d] introduces the notion of *connected* (p. 142) or *related* (p. 19) variables and quantifiers to reflect the idea of individuals being related to each other. The depth of a formula can then be defined as the maximum number of nested and connected/related quantifiers in it. It is not clear that this definition of depth fully captures the idea of the number of related individuals that need to be considered, as is pointed out by Rantala [1987] (p. 72–73, Note 1). However, this concept is not needed for the definition of distributive normal form. We will use as our notion of depth the quantifier rank defined above.

The idea of 'the maximum number of individuals that need to be considered in their relation to each other' for a formula is somewhat imprecise, but it seems to me to mean the smallest quantifier rank of the formulas that are equivalent to the given one. Since there is not an effective way to determine this, we can not use it as a definition if we want to be able to determine the depths of formulas. A definition of depth that is a refinement of quantifier rank in effect gives a way of reducing a formula to one with possibly smaller quantifier rank. Although we won't necessarily get the smallest possible quantifier rank, we can define the depth of a formula as the smallest quantifier rank of an equivalent formula that can be obtained by any method which always terminates after a finite number of steps. The definition as the maximum number of nested and connected quantifiers is an example of this. If we happen to think of some other condition which can be mechanically checked and gives the quantifier rank of an equivalent formula, we could include it in the definition of depth, so the definition is somewhat arbitrary.

of that every not it clear. quantifier rank. quantifier rank.)

For a sentence which expresses a law governing some phenomenon, Hintikka considers the depth of the sentence to be also the *interactional depth* of the law. By *interactional depth* he means "the number of different entities involved in the laws governing certain phenomena" (Hintikka [1973e], p. 313). So, for example, Newton's law of gravitation has an interactional depth of 2 where the entities or individuals are taken to be point masses.

I now consider the first-order language which has only one binary predicate symbol P and no function or constant symbols. Constituents are very large and as a result difficult to write down, so any example I use must be one of the simplest languages. The reason I have chosen the very simplest (non-monadic) is that I later (in chapter 4) consider some limits on the usefulness of constituents and distributive normal forms which are imposed by their size. For this I return to this example to find certain lower bounds.

The way constituents are constructed is by first defining certain formulas called *attributive constituents* with the intention that each describes some *kind of individual* (it actually turns out that only some of them do), and then a constituent says for each kind of individual whether or not it exists, and thus describes a kind of world. A kind of individual is described by its relation to other individuals, given by the predicates they satisfy. We can't just list all possible kinds of individuals since there are infinitely many, and formulas have finite length. So the descriptions of kinds of individuals are arranged into finite sets such that every possible individual is of exactly one of the kinds in each of these sets. This is achieved by limiting the number of other individuals that a kind of individual may be described in relation to, which is a limit on the depth of the formulas involved.

To form attributive constituents, we start by arranging the atomic formulas in *levels*. The purpose of these levels is to keep track of which variables are used. The atomic formulas of level 1 are those over x_1, and for each greater level k the atomic formulas are those over $x_1, \ldots, x_k$ that are not on any previous level. For our particular example, we have:

Level 1 : $P(x_1, x_1)$

Level 2 : $P(x_1, x_2),\ P(x_2, x_1),\ P(x_2, x_2)$

Level 3 : $P(x_1, x_3),\ P(x_3, x_1),\ P(x_2, x_3),\ P(x_3, x_2),\ P(x_3, x_3)$

$\vdots$

We next form *primitive conjunctions* at each level. A primitive conjunction of level k has as conjuncts, for each atomic formula of level k, either it or its negation. So, introducing a notation to be explained shortly, the primitive conjunctions are:

$$\text{Level 1 :}\quad \begin{aligned} P(x_1, x_1) &= \gamma_1^0(x_1) \\ \neg P(x_1, x_1) &= \gamma_2^0(x_1) \end{aligned}$$

$$
\begin{aligned}
\text{Level 2}: \quad & P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2) &= \gamma_1^0(x_1,x_2)\\
& P(x_1,x_2) \wedge P(x_2,x_1) \wedge \neg P(x_2,x_2) &= \gamma_2^0(x_1,x_2)\\
& P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge P(x_2,x_2) &= \gamma_3^0(x_1,x_2)\\
& P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge \neg P(x_2,x_2) &= \gamma_4^0(x_1,x_2)\\
& \neg P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2) &= \gamma_5^0(x_1,x_2)\\
& \neg P(x_1,x_2) \wedge P(x_2,x_1) \wedge \neg P(x_2,x_2) &= \gamma_6^0(x_1,x_2)\\
& \neg P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge P(x_2,x_2) &= \gamma_7^0(x_1,x_2)\\
& \neg P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge \neg P(x_2,x_2) &= \gamma_8^0(x_1,x_2)
\end{aligned}
$$

$\vdots$

The order in which the above formulas are listed illustrates the standard order for formulas defined using the notation $(\pm)$, and using this notation the primitive conjunctions are the formulas of the form:

Level 1 : $(\pm)P(x_1, x_1)$

Level 2 : $(\pm)P(x_1, x_2) \wedge (\pm)P(x_2, x_1) \wedge (\pm)P(x_2, x_2)$

Level 3 : $(\pm)P(x_1, x_3) \wedge (\pm)P(x_3, x_1) \wedge (\pm)P(x_2, x_3) \wedge (\pm)P(x_3, x_2) \wedge (\pm)P(x_3, x_3)$

$\vdots$

The primitive conjunctions of each level k can be seen to describe some individual x_k relative to the individuals $x_1, \ldots, x_{k-1}$ as fully as is possible with only depth-0 formulas. Thus primitive conjunctions are the attributive constituents of depth 0. We use $\gamma_i^d(x_1, \ldots, x_k)$ to denote an attributive constituent of depth d and level k, which explains the notation for the primitive conjunctions above.

An attributive constituent of depth 1 and level k describes an individual x_k relative to $x_1, \ldots, x_{k-1}$ by its depth 0 and depth 1 properties. The depth-1 attributive constituents are:

$$
\begin{aligned}
\text{Level 1}: \quad \gamma^1(x_1) \;=\; & (\pm)P(x_1,x_1)\wedge\\
& (\pm)(\exists x_2)\gamma_1^0(x_1,x_2) \wedge \ldots \wedge (\pm)(\exists x_2)\gamma_8^0(x_1,x_2)
\end{aligned}
$$

$$
\begin{aligned}
\text{Level 2}: \quad \gamma^1(x_1,x_2) \;=\; & (\pm)P(x_1,x_2) \wedge (\pm)P(x_2,x_1) \wedge (\pm)P(x_2,x_2)\wedge\\
& (\pm)(\exists x_3)\gamma_1^0(x_1,x_2,x_3) \wedge \ldots \wedge (\pm)(\exists x_3)\gamma_{32}^0(x_1,x_2,x_3)
\end{aligned}
$$

$$
\begin{aligned}
\text{Level 3}: \quad \gamma^1(x_1,x_2,x_3) \;=\;& (\pm)P(x_1,x_3) \wedge (\pm)P(x_3,x_1) \wedge (\pm)P(x_2,x_3)\wedge \\
& (\pm)P(x_3,x_2) \wedge (\pm)P(x_3,x_3)\wedge \\
& (\pm)(\exists x_4)\gamma_1^0(x_1,\ldots,x_4) \wedge \ldots \wedge (\pm)(\exists x_4)\gamma_{128}^0(x_1,\ldots,x_4)
\end{aligned}
$$

$\vdots$

These formulas can also be written as

$$
\begin{aligned}
\text{Level 1}: \quad \gamma^1(x_1) \;=\;& \gamma_i^0(x_1)\wedge \\
& (\pm)(\exists x_2)\gamma_1^0(x_1,x_2) \wedge \ldots \wedge (\pm)(\exists x_2)\gamma_8^0(x_1,x_2)
\end{aligned}
$$

$$
\begin{aligned}
\text{Level 2}: \quad \gamma^1(x_1,x_2) \;=\;& \gamma_i^0(x_1,x_2)\wedge \\
& (\pm)(\exists x_3)\gamma_1^0(x_1,x_2,x_3) \wedge \ldots \wedge (\pm)(\exists x_3)\gamma_{32}^0(x_1,x_2,x_3)
\end{aligned}
$$

$$
\begin{aligned}
\text{Level 3}: \quad \gamma^1(x_1,x_2,x_3) \;=\;& \gamma_i^0(x_1,x_2,x_3)\wedge \\
& (\pm)(\exists x_4)\gamma_1^0(x_1,\ldots,x_4) \wedge \ldots \wedge (\pm)(\exists x_4)\gamma_{128}^0(x_1,\ldots,x_4)
\end{aligned}
$$

$\vdots$

where the i in the $\gamma_i^0(x_1,\ldots,x_k)$'s varies over all possible values. This way of writing the formulas is shorter and sometimes more convenient, but to count how many of these formulas there are, we need to know how many formulas there are in the sets which generate them. This is easily seen when the attributive constituents are written explicitly as the basic conjunctions generated by some set of formulas. Then, if for some depth and level, the generating set contains n formulas, there are 2^n attributive constituents of that depth and level. This is how all the numbers of formulas of a particular form in this example are obtained.

The depth-2 attributive constituents are:

$$
\begin{aligned}
\text{Level 1}: \quad \gamma^2(x_1) \;=\;& \gamma_i^0(x_1)\wedge \\
& (\pm)(\exists x_2)\gamma_1^1(x_1,x_2) \wedge \ldots \wedge (\pm)(\exists x_2)\gamma_{2^{35}}^1(x_1,x_2)
\end{aligned}
$$

$$
\begin{aligned}
\text{Level 2}: \quad \gamma^2(x_1,x_2) \;=\;& \gamma_i^0(x_1,x_2)\wedge \\
& (\pm)(\exists x_3)\gamma_1^1(x_1,x_2,x_3) \wedge \ldots \wedge (\pm)(\exists x_3)\gamma_{2^{133}}^1(x_1,x_2,x_3)
\end{aligned}
$$

$$
\begin{aligned}
\text{Level 3}: \quad \gamma^2(x_1,x_2,x_3) \;=\;& \gamma_i^0(x_1,x_2,x_3)\wedge \\
& (\pm)(\exists x_4)\gamma_1^1(x_1,\ldots,x_4) \wedge \ldots \wedge (\pm)(\exists x_4)\gamma_{2^{519}}^1(x_1,\ldots,x_4)
\end{aligned}
$$

$\vdots$

To form *constituents* we now say for each possible kind of individual whether or not it exists. Since possible kinds of individuals are described by attributive constituents, constituents have the following forms:

Depth 1 : $(\pm)(\exists x_1)\gamma_1^0(x_1) \wedge (\pm)(\exists x_1)\gamma_2^0(x_1)$

Depth 2 : $(\pm)(\exists x_1)\gamma_1^1(x_1) \wedge \ldots \wedge (\pm)(\exists x_1)\gamma_{512}^1(x_1)$

Depth 3 : $(\pm)(\exists x_1)\gamma_1^2(x_1) \wedge \ldots \wedge (\pm)(\exists x_1)\gamma_{2^{(2^{35}+1)}}^2(x_1)$

$\vdots$

The above constituents are the particular ones which are sentences and in terms of which any sentence has a distributive normal form. They can be considered as "level 0" constituents. There are also constituents of greater levels and they contain free variables. They are formed by using the attributive constituents of levels greater than 1. However their structure is a little more complex since they say what kind of individuals exist relative to some set of individuals, so it is necessary to also describe this set.

Since an attributive constituent of depth d and level 1 describes an individual as fully as is possible with formulas of depths up to d (and not relative to any other individuals), each possible individual is of the kind described by exactly one such attributive constituent for each d. Since a depth-d constituent says precisely which kinds of individuals exist, where the individuals are described as fully as is possible with depth-$(d-1)$ formulas, a depth-d constituent can be said to describe a *depth-d world*. That this makes sense is confirmed by the fact that for any fixed depth, exactly one constituent is true for each possible world. This is the sense in which constituents describe possible kinds of worlds, and depth-d worlds are particular kinds of worlds.

It is because each depth-d world is decribed by some depth-d constituent (although not all depth-d constituents actually describe depth-d worlds) that we can expect sentences to be expressable as disjunctions of constituents, the particular constituents in the disjunction being those which describe the possible depth-d worlds (i.e. kinds of worlds) in which the sentence is true.

The above discussion of what different formulas describe really only applies to *consistent* constituents and attributive constituents. If a constituent or attributive constituent is inconsistent, then it doesn't describe anything. As we will later see, most constituents are inconsistent, and we can wonder if there are any other descriptions of depth-d worlds, or any other (finite) partial descriptions of worlds, in terms of which every formula can be expressed, that do not include such a large proportion of "unusable descriptions" or "descriptions of nothing". Here I mean descriptions that can be defined without a definition which includes a large proportion of inconsistent or unusable elements, so for example, to define constituents and then take the consistent ones to be the partial descriptions we want, is not the kind of definition I am referring to. Although it seems to me unlikely that such sets of partial descriptions can be defined, such questions are beyond the scope of this thesis.

Chapter 3

Definitions and existence of distributive normal form

For first-order languages with only unary predicates and no function or constant symbols, disjunctive normal forms are defined in chapter 1. This chapter gives the definition of distributive normal forms for languages with at least one predicate symbol of arity greater than 1.

In all the definitions of lists of formulas which are indexed by a subscript (often i), the method of indexing is the one described in chapter 1. In all of these definitions of formulas of different depths with one definition for depth 0 and another for depth d, the case for depth d only applies to $d \geq 1$.

3.1 The case for formulas of restricted languages

We first consider languages with only finitely many predicate symbols, any countable number of constant symbols and no function symbols. The following definitions are from Hintikka [1965a]. They are slightly more general than those described in the previous chapter. Since the free variables of a formula may be any finite subset of the variables, the general definitions can't be in terms of levels which only allow some particular sets. We also consider formulas which contain constants. The formulas and sets of formulas will be written with parameters indicating the possible free variables and constants they may have. Variables and constants are called *individual terms*. We use $z_1, z_2, \ldots$ and x (and sometimes other letters, possibly with subscripts) as metavariables to range over the variables $x_1, x_2, \ldots$ and the constants. If a formula has parameters $z_1, \ldots, z_k$, then the variables or constants to which these parameters refer are distinct.

If there are q atomic formulas over $\{z_1, \ldots, z_k\}$ which contain z_k, then we denote these formulas by $\alpha_1(z_1, \ldots, z_k), \ldots, \alpha_q(z_1, \ldots, z_k)$.

An *attributive constituent of depth d with free individual terms* $z_1, \ldots, z_k$ is denoted $\gamma^d(z_1, \ldots, z_k)$ and defined by

$$\gamma_i^0(z_1, \ldots, z_k) = (\pm)\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)\alpha_q(z_1, \ldots, z_k)$$

$$\begin{aligned}\gamma_i^d(z_1, \ldots, z_k) = {} & (\pm)\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)\alpha_q(z_1, \ldots, z_k) \wedge \\ & (\pm)(\exists x)\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_r^{d-1}(z_1, \ldots, z_k, x)\end{aligned}$$

where $\alpha_1(z_1, \ldots, z_k), \ldots, \alpha_q(z_1, \ldots, z_k)$ are all the atomic formulas over $\{z_1, \ldots, z_k\}$ which contain z_k.

A *constituent of depth d with free individual terms* $z_1, \ldots, z_k$ is denoted $\delta^d(z_1, \ldots, z_k)$ and defined by

$$\delta_i^0(z_1, \ldots, z_k) = (\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_p(z_1, \ldots, z_k)$$

$$\begin{aligned}\delta_i^d(z_1, \ldots, z_k) = {} & (\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_p(z_1, \ldots, z_k) \wedge \\ & (\pm)(\exists x)\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_r^{d-1}(z_1, \ldots, z_k, x)\end{aligned}$$

where $\alpha_1(z_1), \ldots, \alpha_p(z_1, \ldots, z_k)$ are all the atomic formulas over $\{z_1, \ldots, z_k\}$.

In both of the above definitions x is the first variable not in $z_1, \ldots, z_k$ and $\gamma_1^{d-1}(z_1, \ldots, z_k, x), \ldots, \gamma_r^{d-1}(z_1, \ldots, z_k, x)$ are all the attributive constituents of depth $d-1$ with free individual terms $z_1, \ldots, z_k, x$, and i is the index by which the formulas are listed. Although it is not really necessary, for convenience we assume that the atomic formulas with any set of free individual terms are given in some order. Then together with the standard order described in chapter 1, we have an order for the formulas $\delta_i^d(z_1, \ldots, z_k)$ to use for examples.

Constituents can be equivalently defined by

$$\delta_i^d(z_1, \ldots, z_k) = (\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_p(z_1, \ldots, z_{k-1}) \wedge \gamma_j^d(z_1, \ldots, z_k)$$

for all depths d, or

$$\delta_i^d(z_1, \ldots, z_k) = \delta_l^0(z_1, \ldots, z_{k-1}) \wedge \gamma_j^d(z_1, \ldots, z_k)$$

for $d \geq 1$ and as above for $d = 0$. This last formula is true for all depths, but is not a complete definition for depth 0.

The definitions of the above formulas can be written in the following more concise way.

Denote the set of atomic formulas over $\{z_1, \ldots, z_k\}$ which contain z_k by $\mathcal{A}(z_1, \ldots, z_k)$, and the set of constituents of depth d with free individual terms $z_1, \ldots, z_k$ by $\mathcal{C}^d(z_1, \ldots, z_k)$, and the set of attributive constituents of depth d with free individual terms $z_1, \ldots, z_k$ by $\mathcal{T}^d(z_1, \ldots, z_k)$. Then, using the notation defined in chapter 1:

$$\mathcal{T}^0(z_1, \ldots, z_k) = \Delta\mathcal{A}(z_1, \ldots, z_k)$$

$$\mathcal{T}^d(z_1, \ldots, z_k) = \Delta(\mathcal{A}(z_1, \ldots, z_k) \cup \exists x(\mathcal{T}^{d-1}(z_1, \ldots, z_k, x)))$$

$$\mathcal{C}^0(z_1, \ldots, z_k) = \Delta(\bigcup_{i=1}^{k} \mathcal{A}(z_1, \ldots, z_i))$$

$$\mathcal{C}^d(z_1, \ldots, z_k) = \Delta(\bigcup_{i=1}^{k} \mathcal{A}(z_1, \ldots, z_i) \cup \exists x(\mathcal{T}^{d-1}(z_1, \ldots, z_k, x))).$$

It is also possible to define constituents in a way that is not equivalent to this and still get the result that any formula can be expressed as a disjunction of constituents. For example, Rantala [1987] gives a definition which uses constituents for the attributive constituents in the recursive definition of constituent (and also explicitly adds inconsistent constituents). This makes it unnecessary to define attributive constituents and results in a definition which is slightly simpler to write down. However, it also makes the constituents unnecessarily complex in the sense that they are bigger, there are more of them, and there are more inconsistent constituents. All of these results are caused by the scopes of the quantifiers containing conjuncts which do not contain the variable bound by the quantifier and may thus be moved out of its scope; however these conjuncts may clash with each other in that some may be the negations of others, making the formula inconsistent. Stated another way, such a constituent is equivalent to one which is of the form of a basic conjunction generated by a multiset of formulas, with each element either negated or not independently of the others, and to be consistent either all repetitions of a formula in the generating multiset must be negated or all must be unnegated in the constituent.

Another alternative definition of constituents, but only for sentences, is given by Niiniluoto [1987] (p. 66). Although separate definitions for constituents and attributive constituents are given, this is not necessary as the constituents are just the attributive constituents with no free variables. This is somewhat similar to the definition I later give of *sentence constituents* as *existential conjunctions* with no free variables. The difference is that this particular definition of Niiniluoto's gives constituents that are unnecessarily big in the same type of way as Rantala's constituents (although they are not exactly the same as the sentence case of Rantala's definition). Since the use of constituents is limited by their size and by not being able to find which constituents are consistent, and I later consider some lower bounds for the number of constituents and the fraction which are inconsistent, I use the definition which gives the smallest and fewest constituents that I have seen, which is

Hintikka's original definition. (The definitions in Hintikka [1953] and Hintikka [1965a] are the same, though with slightly different terminology.)

Lemma 3.1 *For each model, for each set of individual terms $z_1, \ldots, z_k$, for each valuation, exactly one constituent of the form $\delta_i^d(z_1, \ldots, z_k)$ is true for each d.*

PROOF For each depth d and each set of individual terms $z_1, \ldots, z_k$, the constituents of depth d with free individual terms $z_1, \ldots, z_k$ are the basic conjunctions generated by some set of formulas, so by an observation in chapter 1, exactly one such constituent is true for any given model and valuation. □

We have the corresponding result for attributive constituents:

Lemma 3.2 *For each model, for each set of individual terms $z_1, \ldots, z_k$, for each valuation, exactly one attributive constituent of the form $\gamma_i^d(z_1, \ldots, z_k)$ is true for each d.*

PROOF The attributive constituents of the form $\gamma_i^d(z_1, \ldots, z_k)$ are the basic conjunctions generated by some set of formulas, so for any model and valuation exactly one of them is true. □

Corollary 3.3 *For every depth d, for every set of individual terms $z_1, \ldots, z_k$,*

$$\bigwedge_{i \in I} \neg\gamma_i^d(z_1, \ldots, z_k) \iff \bigvee_{i \notin I} \gamma_i^d(z_1, \ldots, z_k).$$

PROOF For some model and valuation, if $\bigwedge_{i \in I} \neg\gamma_i^d(z_1, \ldots, z_k)$ is true, then $\gamma_i^d(z_1, \ldots, z_k)$ is false for each $i \in I$, so $\gamma_i^d(z_1, \ldots, z_k)$ is true for some $i \notin I$ by lemma 3.2, so $\bigvee_{i \notin I} \gamma_i^d(z_1, \ldots, z_k)$ is true. Conversely, if for some model and valuation, $\bigvee_{i \notin I} \gamma_i^d(z_1, \ldots, z_k)$ is true, then $\gamma_i^d(z_1, \ldots, z_k)$ is true for some $i \notin I$, so $\gamma_i^d(z_1, \ldots, z_k)$ is false for each $i \in I$ by lemma 3.2, so $\bigwedge_{i \in I} \neg\gamma_i^d(z_1, \ldots, z_k)$ is true. □

Lemma 3.1 has the following corollary.

Corollary 3.4 *For every depth d, for every set of individual terms $z_1, \ldots, z_k$,*

1. *if $F \iff \bigvee_{i \in I} \delta_i^d(z_1, \ldots, z_k)$, then $\neg F \iff \bigvee_{i \notin I} \delta_i^d(z_1, \ldots, z_k)$.*
2. *if $F \iff \bigvee_{i \in I} \delta_i^d(z_1, \ldots, z_k)$ and $G \iff \bigvee_{i \in J} \delta_i^d(z_1, \ldots, z_k)$, then $F \vee G \iff \bigvee_{i \in I \cup J} \delta_i^d(z_1, \ldots, z_k)$.*
3. *if $F \iff \bigvee_{i \in I} \delta_i^d(z_1, \ldots, z_k)$ and $G \iff \bigvee_{i \in J} \delta_i^d(z_1, \ldots, z_k)$, then $F \wedge G \iff \bigvee_{i \in I \cap J} \delta_i^d(z_1, \ldots, z_k)$.*

PROOF

1. For any model and valuation, if $\delta_i^d(z_1,\ldots,z_k)$ is true for some $i \notin I$, then $\delta_i^d(z_1,\ldots,z_k)$ is not true for any $i \in I$ by lemma 3.1, so F is not true, so $\neg F$ is true. Thus $\bigvee_{i\notin I}\delta_i^d(z_1,\ldots,z_k) \implies \neg F$. Conversely, if F is not true then $\delta_i^d(z_1,\ldots,z_k)$ is not true for each $i \in I$, but by lemma 3.1 one $\delta_i^d(z_1,\ldots,z_k)$ is true, so it must be for some $i \notin I$, so $\neg F \implies \bigvee_{i\notin I}\delta_i^d(z_1,\ldots,z_k)$.
2. by the associativity and commutativity of disjunction (which is implicit in the notation used)
3. by 1. and 2. above and De Morgan's law □

Theorem 3.5 (Hintikka [1965a]) *Every formula of depth $\leq d$ whose free individual terms are contained in $z_1,\ldots,z_k$ can be expressed as a disjunction of constituents of depth d with free individual terms $z_1,\ldots,z_k$.*

PROOF First note that the set of conjuncts that occur negated or not in a constituent does not depend on the order of the parameters of the constituent. This is just by the definition of constituents and is used throughout this proof. We use induction on the depth d. For $d = 0$, any formula of depth 0 with free individual terms in $z_1,\ldots,z_k$ is a Boolean combination of atomic formulas over $z_1,\ldots,z_k$, so it has a disjunctive normal form which is a disjunction of formulas of the form $\delta_j^0(z_1,\ldots,z_k)$, which is the required form. Assume that for some $d \geq 1$, for every set $z_1,\ldots,z_k$ of individual terms, any formula of depth $< d$ with free individual terms in $z_1,\ldots,z_k$ is equivalent to a disjunction of some $\delta_j^{d-1}(z_1,\ldots,z_k)$'s. Let F be any formula of depth $\leq d$ with free individual terms in $z_1,\ldots,z_k$ written without any abbreviation. We use induction on the formation of the formula. If F is an atomic formula, then F is equivalent to a disjunction of formulas of the form $\delta_i^0(z_1,\ldots,z_k)$ as shown above, and each $\delta_i^0(z_1,\ldots,z_k)$ is equivalent to the disjunction of all the formulas of the form

$$\delta_i^0(z_1,\ldots,z_k) \wedge (\pm)(\exists x)\gamma_1^{d-1}(z_1,\ldots,z_k,x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_p^{d-1}(z_1,\ldots,z_k,x)$$

and each of the above formulas is some $\delta_j^d(z_1,\ldots,z_k)$, so F is equivalent to a disjunction of some $\delta_j^d(z_1,\ldots,z_k)$'s. If F is $(\exists x)G$ then by changing the bound variables if necessary, we can assume that x is the first variable not in $z_1,\ldots,z_k$. Then G is a formula of depth $\leq d-1$ with free individual terms in $z_1,\ldots,z_k,x$, so by hypothesis can be expressed as some disjunction

$$\bigvee_{i\in I}\delta_i^{d-1}(z_1,\ldots,z_k,x)$$

and each $\delta_i^{d-1}(z_1,\ldots,z_k,x)$ is of the form $\delta_{i_a}^0(z_1,\ldots,z_k) \wedge \gamma_{i_b}^{d-1}(z_1,\ldots,z_k,x)$. So

$$(\exists x)G \iff \bigvee_{i\in I}(\delta_{i_a}^0(z_1,\ldots,z_k) \wedge (\exists x)\gamma_{i_b}^{d-1}(z_1,\ldots,z_k,x))$$

by distributing the existential quantification over disjunction and since the $\delta^0_{i_a}(z_1,\ldots,z_k)$'s don't contain x. Now, each $\delta^0_{i_a}(z_1,\ldots,z_k) \wedge (\exists x)\gamma^{d-1}_{i_b}(z_1,\ldots,z_k,x)$ is equivalent to the disjunction of all the formulas of the form

$$\delta^0_{i_a}(z_1,\ldots,z_k)\wedge$$
$$(\pm)(\exists x)\gamma^{d-1}_1(z_1,\ldots,z_k,x) \wedge \ldots \wedge (\exists x)\gamma^{d-1}_{i_b}(z_1,\ldots,z_k,x) \wedge \ldots \wedge (\pm)(\exists x)\gamma^{d-1}_p(z_1,\ldots,z_k,x)$$

and each of the above formulas is some $\delta^d_j(z_1,\ldots,z_k)$, so F is equivalent to a disjunction of some $\delta^d_j(z_1,\ldots,z_k)$'s. If F is $G \vee H$ then G and H each have depth $\leq d$ and free individual terms in $z_1,\ldots,z_k$, so each is equivalent to a disjunction of some $\delta^d_j(z_1,\ldots,z_k)$'s, thus so is F. If F is $\neg G$ then G has depth $\leq d$ and free individual terms in $z_1,\ldots,z_k$, so is equivalent to a disjunction of some $\delta^d_j(z_1,\ldots,z_k)$'s, thus so is F by corollary 3.4. □

An obvious question to ask is whether the expression of a depth-d formula as a disjunction of depth-d constituents is unique. If there are any inconsistent constituents, then each can either be part of the disjunction or not without affecting the meaning of the disjunction, so the expression of a formula as a disjunction of constituents will not be unique. To show that there are inconsistent constituents, here is an example: Consider the language with just one binary predicate symbol P. $\delta^2() = \gamma^2()$ is inconsistent if it contains an inconsistent conjunct, so let $(\exists x_1)\gamma^1_i(x_1)$ be a conjunct where

$$\gamma^1_i(x_1) = P(x_1,x_1) \wedge \neg(\exists x_2)\gamma^0_1(x_1,x_2) \wedge \ldots \wedge \ldots$$

where

$$\gamma^0_1(x_1,x_2) = P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2)$$

Now, the inconsistency of $(\exists x_1)(P(x_1,x_1) \wedge \neg(\exists x_2)\gamma^0_1(x_1,x_2))$ is sufficient for the inconsistency of the whole formula, and this is just

$$(\exists x_1)(P(x_1,x_1) \wedge \neg(\exists x_2)(P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2)))$$

which is inconsistent.

The next obvious question is whether the set of consistent constituents that occur in the expression of a formula is unique.

Theorem 3.6 *The set of consistent constituents that occur in the expression of a depth-d formula as a disjunction of depth-d constituents (with the same free individual terms as the formula) is unique.*

PROOF For two disjunctions of depth-d constituents F_1 and F_2, if F_1 includes some consistent constituent $\delta^n_i(z_1,\ldots,z_k)$ that F_2 doesn't, then there is a model $\mathcal{M}$ and a valuation v for $\mathcal{M}$ such that $\mathcal{M}, v \models \delta^n_i(z_1,\ldots,z_k)$, so $\mathcal{M}, v \models F_1$ and by lemma 3.1, for each $j \neq i$, $\mathcal{M}, v \not\models \delta^n_j(z_1,\ldots,z_k)$, so $\mathcal{M}, v \not\models F_2$, so $F_1 \not\Rightarrow F_2$. □

Thus a consistent depth-d formula may be expressed uniquely as a (non-empty) disjunction of consistent depth-d constituents. Since we allow empty disjunctions, every formula of depth d may be expressed uniquely as a disjunction of consistent depth-d constituents. Although there is no effective procedure for finding which constituents are consistent, the following definitions are useful for purposes of discussion.

By theorem 3.5, any formula of depth d with free individual terms $z_1, \ldots, z_k$ is equivalent to a disjunction of constituents of depth e with free individual terms $z_1, \ldots, z_l$ for any $e \geq d$ and any $\{z_1, \ldots, z_l\} \supset \{z_1, \ldots, z_k\}$. Such a disjunction is called a *distributive normal form at depth e with free individual terms $z_1, \ldots, z_l$* of the formula. Usually we will consider distributive normal forms with the same free individual terms as the formula, in which case the individual terms need not be mentioned. Likewise, if the depth of a distributive normal form is not mentioned, it is taken to be the same as that of the formula. A distributive normal form at depth e of a formula of depth d, where $e > d$, is also called an *expansion* of the formula to depth e. Following Rantala [1987], we call the disjunction of consistent depth-d constituents to which a depth-d formula is equivalent the *distributive normal form$_0$* of the formula (at depth d). An expansion which contains only consistent constituents will be called the *expansion$_0$* of the formula. In particular, the constituents of depth d have expansions to depth e, where an expansion $\bigvee_{i \in I} \delta_i^e(z_1, \ldots, z_k)$ of $\delta_j^d(z_1, \ldots, z_k)$ may contain any of the inconsistent constituents of depth e with free individual terms $z_1, \ldots, z_k$ and must contain all of the consistent constituents of depth e with free individual terms $z_1, \ldots, z_k$ which imply $\delta_j^d(z_1, \ldots, z_k)$.

A formula may be represented by the set of constituents in any of its distributive normal forms. We will call the set of constituents in a distributive normal form a *set representation* (at the same depth as the distributive normal form) of the formula, and the set of constituents in the distributive normal form$_0$ the *set representation$_0$* of the formula. So the set representation$_0$ of any inconsistent formula is the empty set.

We can state conditions for the validity of a formula and for when one formula implies another in terms of their distributive normal forms:

Theorem 3.7 *A formula F of depth $\leq d$ is valid iff its distributive normal form at depth d contains all the consistent depth-d constituents.*

PROOF If there is a consistent depth-d constituent not in F then there is a model and valuation for which F is false by lemma 3.1. Conversely, if F contains all consistent depth-d constituents then F is valid by lemma 3.1 (existence part). □

We can tell when one formula implies another of the same depth (and with the same free individual terms) by comparing their set representations. The same condition applies to formulas of different depths if the less deep formula is first expanded to the greater depth. I use the same notation for a formula and its set representation. In a set-theoretic context the set representation is meant, otherwise the formula is meant.

Theorem 3.8 *For two formulas F_1 and F_2, one of depth d and the other of depth e with $e \geq d$ (and with the same free individual terms):*

1. *For any set representations at depth e of the formulas,*

$$\text{if } \ F_1 \subseteq F_2 \ \text{ then } \ F_1 \Longrightarrow F_2.$$

2. *For the set representation$_0$ of F_1 and any set representation of F_2, both at depth e,*

$$F_1 \subseteq F_2 \ \text{ iff } \ F_1 \Longrightarrow F_2.$$

PROOF

1. If $F_1 \subseteq F_2$ then since adding disjuncts to a disjunction gives a formula implied by the original one, $F_1 \Longrightarrow F_2$.
2. For the direction not shown above, if there is a consistent $\delta \in F_1 \setminus F_2$ then there is a model and valuation for which δ is true, thus also F_1 is true and by lemma 3.1 F_2 is not true, so $F_1 \not\Longrightarrow F_2$. □

In the next chapter I use theorem 3.8 to find a condition for when one formula implies another that doesn't require the formulas to be expanded to the greater depth. There are, however, a number of factors restricting the applicability of this result, the most important being that we can not determine algorithmically which constituents are consistent. If we could, it would give a decision procedure for first-order logic (see theorem 4.9). Theorem 3.8 gives another way of showing that the consistent constituents can not be found algorithmically, since if they could, it would give a method for deciding implication between formulas, which would in turn give a decision procedure for first-order logic. So we can either look for conditions for implication between formulas that are both necessary and sufficient, but not effectively checkable. Or we can look for conditions that are sufficient but not necessary for implication which are effectively checkable. The two cases in theorem 3.8 are examples of this. Although we do have an effectively checkable sufficient condition for implication, it is unclear whether this condition would actually occur much in practice. In addition, distributive normal forms, even if we ignore the size of the constituents and just look at the number of constituents, are extremely big. I give an indication of their size in the next chapter.

3.2 The special case for constant-free sentences

In constituents, free variables and constant symbols play the same role. I will refer to a sentence which does not contain any constant symbols as a *constant-free sentence*. As a special case of theorem 3.5, any constant-free sentence of depth $\leq d$ can be expressed as a disjunction of constituents of the form $\delta_i^d()$. I now consider what these particular constituents are.

It is possible to give a definition for these constituents that is slightly simpler than the general definition, and so may help to clarify what they are. To see how to get this definition, we start with some examples. For each $\gamma^0(z_1, \ldots, z_k)$, I will write $\beta(z_1, \ldots, z_k)$. These formulas form a basis for the inductive definitions given earlier, and later they will be used in a definition which doesn't include the general definition of attributive constituent.

$\delta_i^0()$ — there are none

$\delta_i^1() = (\pm)(\exists x_1)\beta_1(x_1) \wedge \ldots \wedge (\pm)(\exists x_1)\beta_q(x_1)$

$\delta_i^2() = (\pm)(\exists x_1)\gamma_1^1(x_1) \wedge \ldots \wedge (\pm)(\exists x_1)\gamma_{q_1}^1(x_1)$

where $\gamma_i^1(x_1) = \beta_j(x_1) \wedge (\pm)(\exists x_2)\beta_1(x_1, x_2) \wedge \ldots \wedge (\pm)(\exists x_2)\beta_{q_2}(x_1, x_2)$

$\delta_i^3() = (\pm)(\exists x_1)\gamma_1^2(x_1) \wedge \ldots \wedge (\pm)(\exists x_1)\gamma_{q_3}^2(x_1)$

where $\gamma_i^2(x_1) = \beta_j(x_1) \wedge (\pm)(\exists x_2)\gamma_1^1(x_1, x_2) \wedge \ldots \wedge (\pm)(\exists x_2)\gamma_{q_4}^1(x_1, x_2)$

where $\gamma_i^1(x_1, x_2) = \beta_j(x_1, x_2) \wedge (\pm)(\exists x_3)\beta_1(x_1, x_2, x_3) \wedge \ldots \wedge (\pm)(\exists x_3)\beta_{q_5}(x_1, x_2, x_3)$

Since in all the above formulas, the variables are used in the same order, we can replace the parameters with an additional superscript (on the left, since the right is already used for depth) which we call the *level* and which tells us what the free variables are and what order they are in.

By replacing each $(x_1, \ldots, x_k)$ by a left superscript k we get:

$^0\delta_i^1 = \boxed{(\pm)(\exists x_1)^1\beta_1 \wedge \ldots \wedge (\pm)(\exists x_1)^1\beta_q}$

$^0\delta_i^2 = \boxed{(\pm)(\exists x_1)^1\gamma_1^1 \wedge \ldots \wedge (\pm)(\exists x_1)^1\gamma_{q_1}^1}$

where $^1\gamma_i^1 = {}^1\beta_j \wedge \boxed{(\pm)(\exists x_2)^2\beta_1 \wedge \ldots \wedge (\pm)(\exists x_2)^2\beta_{q_2}}$

$^0\delta_i^3 = \boxed{(\pm)(\exists x_1)^1\gamma_1^2 \wedge \ldots \wedge (\pm)(\exists x_1)^1\gamma_{q_3}^2}$

where $^1\gamma_i^2 = {}^1\beta_j \wedge \boxed{(\pm)(\exists x_2)^2\gamma_1^1 \wedge \ldots \wedge (\pm)(\exists x_2)^2\gamma_{q_4}^1}$

where $^2\gamma_i^1 = {}^2\beta_j \wedge \boxed{(\pm)(\exists x_3)^3\beta_1 \wedge \ldots \wedge (\pm)(\exists x_3)^3\beta_{q_5}}$

The $^k\beta_i$'s will be called *primitive conjunctions* and are defined explicitly as follows:

The *atomic formulas of level k* are those over the variables $x_1, \ldots, x_k$ which contain x_k. If there are q atomic formulas of level k, they are denoted ${}^k\alpha_1, \ldots, {}^k\alpha_q$.

The *primitive conjunctions of level k* are the formulas of the form

$${}^k\beta_i = (\pm)^k\alpha_1 \wedge \ldots \wedge (\pm)^k\alpha_q$$

where $i = 1, \ldots, 2^q$ is the index by which they are listed. (These are the same primitive conjunctions as described in chapter 2.)

Now, the above formulas which are in boxes have the particular form

$$(\pm)(\exists x_i)^i\gamma_1^d \wedge \ldots \wedge (\pm)(\exists x_i)^i\gamma_q^d$$

which I have called an *existential conjunction.* (The term "existential conjunction" is used, for example, by Brink [1989] to refer just to the case of depth 1, that is "where the γ's are β's".) The constituents above are such existential conjunctions, and the attributive constituents above are conjunctions of a primitive conjunction and an existential conjunction. These existential conjunctions can be defined without the general definition for constituents. Those other than ${}^0\delta^d$ can be defined in a way that is a natural extension of the definition for ${}^0\delta^d$, and it would be convenient to use the notation ${}^k\delta^d$ for these formulas. However, since I have already used δ's for constituents, and I don't want to give the impression that existential conjunctions are constituents, I will use ϵ's for these formulas. So, writing ${}^0\epsilon^d$ for ${}^0\delta^d$ and ${}^k\epsilon^d$ for the existential conjunctions of other levels, we get

$${}^k\epsilon_i^1 = (\pm)(\exists x_{k+1})^{k+1}\beta_1 \wedge \ldots \wedge (\pm)(\exists x_{k+1})^{k+1}\beta_q$$

$${}^k\epsilon_i^d = (\pm)(\exists x_{k+1})^{k+1}\gamma_1^{d-1} \wedge \ldots \wedge (\pm)(\exists x_{k+1})^{k+1}\gamma_p^{d-1}.$$

And each ${}^k\gamma^d$ has the form ${}^k\beta \wedge {}^k\epsilon^d$, so we can substitute for the above ${}^{k+1}\gamma_i^{d-1}$'s to get the definition:

The *existential conjunctions of level k and depth d* are denoted ${}^k\epsilon_i^d$ and are defined by

$${}^k\epsilon_i^1 = (\pm)(\exists x_{k+1})^{k+1}\beta_1 \wedge \ldots \wedge (\pm)(\exists x_{k+1})^{k+1}\beta_q$$

$$\begin{aligned} {}^k\epsilon_i^d \ = \ & (\pm)(\exists x_{k+1})({}^{k+1}\beta_1 \wedge {}^{k+1}\epsilon_1^{d-1}) \wedge \ldots \wedge (\pm)(\exists x_{k+1})({}^{k+1}\beta_1 \wedge {}^{k+1}\epsilon_p^{d-1}) \\ & \wedge \ldots \wedge \\ & (\pm)(\exists x_{k+1})({}^{k+1}\beta_q \wedge {}^{k+1}\epsilon_1^{d-1}) \wedge \ldots \wedge (\pm)(\exists x_{k+1})({}^{k+1}\beta_q \wedge {}^{k+1}\epsilon_p^{d-1}) \end{aligned}$$

where ${}^{k+1}\beta_1, \ldots, {}^{k+1}\beta_q$ are all the primitive conjunctions of level $k+1$ and ${}^{k+1}\epsilon_1^{d-1}, \ldots, {}^{k+1}\epsilon_p^{d-1}$ are all the existential conjunctions of level $k+1$ and depth $d-1$.

Now, the depth-d constituents with no free variables or constants are exactly the existential conjunctions of depth d and level 0. I will call these constituents *sentence constituents* (since *constant-free-sentence-constituents* is clumsy).

As for the general definition, the above definition can be written in the following more concise way:

Denote the set of atomic formulas of level k by ${}^k\mathcal{A}$, the set of primitive conjunctions of level k by ${}^k\mathcal{B}$, and the set of existential conjunctions of level k and depth d by ${}^k\mathcal{E}^d$, then

$$
\begin{aligned}
{}^k\mathcal{B} &= \Delta\,{}^k\mathcal{A} \\
{}^k\mathcal{E}^1 &= \Delta(\exists x_{k+1}({}^{k+1}\mathcal{B})) \\
{}^k\mathcal{E}^d &= \Delta(\exists x_{k+1}(\bigwedge({}^{k+1}\mathcal{B}, {}^{k+1}\mathcal{E}^{d-1})))
\end{aligned}
$$

We now have, by theorem 3.5, that any constant-free sentence of depth $\leq d$ is equivalent to a disjunction of depth-d constituents of the form ${}^0\epsilon_i^d$. I have tried to prove this directly, without using the general definition of constituent, but have so far not been able to.

3.3 The remaining cases

Languages with infinitely many predicate symbols

The definitions which have been given for languages with only finitely many predicate symbols can easily be made to apply to languages with infinitely many. Each formula contains only a finite number of predicate symbols, thus is also a formula of the language which contains only these predicate symbols, and thus has a distributive normal form as defined for this restricted language (and for each language whose predicates are a finite superset of those of this language). Thus each formula has a distributive normal form relative to each finite set of predicates which contains those in the formula. We can consider all the formulas which have been defined in the process of defining constituents to be relative to some finite subset of the set of predicates. If there are only finitely many predicates, we can always use the definitions relative to the set of all predicates. Otherwise we get, for each finite subset of the set of predicates, a list of definitions like those given above. To make this more explicit we could let the formulas have an additional parameter indicating the predicates they contain. However this would make the formulas notationally more cumbersome and is not necessary for showing the existence of distributive normal form for languages with infinitely many predicate symbols. So I don't introduce more complicated notation.

Languages with identity

There are at least two approaches to defining distributive normal forms for languages with identity. The easiest way to see that these normal forms exist is to just consider the identity predicate $=$ as an ordinary binary predicate, and form constituents as before. Then for any formula F, there is a disjunction $\bigvee_{i \in I} \delta_i$ of constituents such that $F \Longleftrightarrow \bigvee_{i \in I} \delta_i$, thus $F \Longleftrightarrow_{=} \bigvee_{i \in I} \delta_i$. With this approach, there are many constituents that are inconsistent because the predicate $=$ doesn't satisfy the conditions of identity. (For example, half of the constituents which contain $x = y$ as a conjunct also contain $y \neq x$.)

With the other approach, the attributive constituents do not explicitly contain identity, but are considered as abbreviations of formulas which do contain identity. This, Hintikka calls an *exclusive interpretation of quantifiers*. As compared with the previous approach, it eliminates the vast number of constituents that are inconsistent because the identity predicate doesn't satisfy the conditions of identity. Niiniluoto [1987] (p. 68) gives a definition for the exclusive interpretation of quantifiers which is the special case of the definition given below applied to the specific constituents he is dealing with.

Hintikka [1965a] (p. 59) defines a formula of the form $(\exists x)F$ where *the quantifier has an exclusive interpretation* to be an abbreviation for the formula $(\exists x)(x \neq a_1 \wedge \ldots \wedge x \neq a_k \wedge F)$ where $a_1, \ldots, a_k$ are the free individual terms of F. But it seems to me that $a_1, \ldots, a_k$ should be the free individual terms of $(\exists x)F$ (I use this for the definition), since otherwise when x is free in F, the formula $(\exists x)F$ with the quantifier interpreted exclusively contains in the scope of the quantifier the conjunct $x \neq x$, and thus is inconsistent. Similarly, $(\forall x)F$ where *the quantifier has an exclusive interpretation* is an abbreviation for the formula $(\forall x)((x \neq a_1 \wedge \ldots \wedge x \neq a_k) \rightarrow F)$.

For any formula F, the formula which is the same as F except that all quantifiers are interpreted exclusively is denoted F_{ex}.

Attributive constituents for a language with identity are defined as for other languages except that all quantifiers occurring in them are interpreted exclusively. This has the effect of including more conjuncts in the scopes of the quantifiers. We use the same notation for attributive constituents as before since it will be obvious from the context which attributive constituents are meant.

Constituents for languages with identity are formed from the attributive constituents as before, except that they have some additional conjuncts. To give the definition it is necessary to introduce some additional notation. In this context $\delta_i^d(z_1, \ldots, z_k)$ will be used to denote a formula defined as a constituent was previously, but where all quantifiers occurring in them are interpreted exclusively. These formulas $\delta_i^d(z_1, \ldots, z_k)$ are however not generally constituents.

Constituents are denoted $\delta_i^d(z_1, \ldots, z_k; a_1, \ldots, a_n)$ where all of $z_1, \ldots, z_k, a_1, \ldots, a_n$ are individual terms which are all distinct from each other. A constituent $\delta_i^d(z_1, \ldots, z_k; a_1, \ldots, a_n)$ is a formula of the form

$$(a_1 = z_{i_1}) \wedge \ldots \wedge (a_n = z_{i_n}) \wedge \bigwedge_{\substack{i,j \in \{1,\ldots,k\} \\ i \neq j}} (z_i \neq z_j) \wedge \delta_i^d(z_1, \ldots, z_k)$$

where $z_{i_1}, \ldots, z_{i_n} \in \{z_1, \ldots, z_k\}$. If there are not any a's then the first part of the constituent is omitted, that is, a constituent $\delta_i^d(z_1, \ldots, z_k; \,)$ has the form

$$\bigwedge_{\substack{i,j \in \{1,\ldots,k\} \\ i \neq j}} (z_i \neq z_j) \wedge \delta_i^d(z_1, \ldots, z_k).$$

The *main part* of a constituent $\delta_i^d(z_1, \ldots, z_k; a_1, \ldots, a_n)$ is that part of the formula of the form

$$\bigwedge_{\substack{i,j \in \{1,\ldots,k\} \\ i \neq j}} (z_i \neq z_j) \wedge \delta_i^d(z_1, \ldots, z_k).$$

Thus the main part of a constituent of the form $\delta^d(z_1, \ldots, z_k; a_1, \ldots, a_n)$ is a constituent of the form $\delta^d(z_1, \ldots, z_k; \,)$.

Any constituent of the form $\delta_i^d(z_1, \ldots, z_k; a_1, \ldots, a_n)$ is said to be a constituent *over* the set $\{z_1, \ldots, z_k, a_1, \ldots, a_n\}$.

If F is a formula such that $F \Longleftrightarrow_= F_{ex}$, then we say that *the interpretation of the quantifiers in F can be changed.*

Lemma 3.9 *For every formula F, there is a formula G such that $F \Longleftrightarrow_= G$ and the interpretation of the quantifiers in G can be changed.*

PROOF We use induction on the formation of the formula. If F doesn't contain any quantifiers, then $F \Longleftrightarrow_= F_{ex}$. (In particular, this holds for all atomic formulas.) If F is of the form $(\exists x)H$ and H is equivalent to some formula G for which the interpretation of the quantifiers may be changed, then $F \Longleftrightarrow_= (\exists x)G$. If there are no free individual terms in $(\exists x)G$ then $(\exists x)G \Longleftrightarrow_= ((\exists x)G)_{ex}$. If the free individual terms of $(\exists x)G$ are $a_1, \ldots, a_k$ then G is equivalent to the disjunction of all the formulas of the form $G \wedge (\pm)(x =$

$a_1) \wedge \ldots \wedge (\pm)(x = a_k)$. So $(\exists x)G \Longleftrightarrow_= \bigvee(\exists x)(G \wedge (\pm)(x = a_1) \wedge \ldots \wedge (\pm)(x = a_k))$. Each disjunct $(\exists x)(G \wedge (\pm)(x = a_1) \wedge \ldots \wedge (\pm)(x = a_k))$ which contains some $x = a_i$ is equivalent to the formula obtained from this disjunct by substituting a_i for x and omitting $(\exists x)$ and redundant conjuncts (as in the following example). This gives a formula from which the variable x and the outermost layer of quantification (which had been over x) have been removed, and the interpretation of deeper quantifiers remains able to be changed, thus the interpretation of quantifiers in this formula can be changed. The only disjunct not of that form is $(\exists x)(G \wedge (x \neq a_1) \wedge \ldots \wedge (x \neq a_k))$ which is already in a form where the interpretation of quantifiers can be changed. Thus F is equivalent to a disjunction where each disjunct can have the interpretation of quantifiers changed, thus this disjunction is the required formula. If F is of the form $G \vee H$ and both G and H have equivalent formulas for which the interpretation of quantifiers can be changed, then the disjunction of these formulas is the required formula. If F is of the form $\neg G$ and G has an equivalent formula for which the interpretation of quantifiers can be changed, then the negation of that formula is the required formula. □

I now give an example of converting a formula to a form where the interpretation of the quantifiers can be changed: Let F be $(\exists x)(P(x,a) \wedge Q(x,b))$. Then

$$\begin{aligned}
F \Longleftrightarrow_= \; & (\exists x)(P(x,a) \wedge Q(x,b) \wedge (x = a) \wedge (x = b)) \vee \\
& (\exists x)(P(x,a) \wedge Q(x,b) \wedge (x = a) \wedge (x \neq b)) \vee \\
& (\exists x)(P(x,a) \wedge Q(x,b) \wedge (x \neq a) \wedge (x = b)) \vee \\
& (\exists x)(P(x,a) \wedge Q(x,b) \wedge (x \neq a) \wedge (x \neq b)) \\
\Longleftrightarrow_= \; & (P(a,a) \wedge Q(a,b) \wedge (a = b)) \vee \\
& (P(a,a) \wedge Q(a,b) \wedge (a \neq b)) \vee \\
& (P(b,a) \wedge Q(b,b) \wedge (a \neq b)) \vee \\
& (\exists x)(P(x,a) \wedge Q(x,b) \wedge (x \neq a) \wedge (x \neq b))
\end{aligned}$$

Lemma 3.10 *For each normal model, for each valuation for it, for each set of individual terms $z_1, \ldots, z_k$, exactly one formula of the form $\delta^d(z_1, \ldots, z_k)$ is true for each d.*

PROOF This result is like lemma 3.1 applied to normal models, except that the formulas $\delta^d(z_1, \ldots, z_k)$ now have their quantifiers interpreted exclusively. It is true because the formulas of the form $\delta^d(z_1, \ldots, z_k)$ are still the basic conjunctions generated by some set of formulas. □

Lemma 3.11 *For each normal model and valuation for the model, for each set of individual terms $z_1, \ldots, z_n$ and partition $\{z_a, \ldots, z_b\}, \{z_c, \ldots, z_d\}$ of these terms, for each depth d, if $\delta_i^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ is true for some i then $\delta_j^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ is false for all $j \neq i$, and there is some such partition for which one of the constituents is true.*

PROOF First note that the order of the parameters $z_a, \ldots, z_b$ and of $z_c, \ldots, z_d$ affects only the order of the conjuncts (including nested ones) of the constituents, and not what these conjuncts are. If $\delta_i^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ is true then the only ways in which $\delta_j^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ can differ from it are by containing a different $\delta^d(z_a, \ldots, z_b)$ in which case $\delta_j^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ is false by lemma 3.10, or by containing a different conjunct $z_e = z_f$ for some $z_e \in \{z_c, \ldots, z_d\}$ and $z_f \in \{z_a, \ldots, z_b\}$ in which case $\delta_j^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ is false since no distinct elements of $\{z_a, \ldots, z_b\}$ are equal. To show that there is some partition for which one of the constituents is true: For any valuation for a normal model, each formula $z_i = z_j$ (for $i, j \in \{1, \ldots, n\}$) is either true or false in such way that $=$ satisfies the conditions of an equivalence relation. From each equivalence class (there are finitely many) choose an element and let these elements be $z_a, \ldots, z_b$ (in any order). Let $z_c, \ldots, z_d$ be $\{z_1, \ldots, z_n\} \setminus \{z_a, \ldots, z_b\}$ (again in any order). Then for all $i, j \in \{a, \ldots, b\}$ such that $i \neq j$, the formula $z_i \neq z_j$ is true, and for all $z_j \in \{z_c, \ldots, z_d\}$ there is a $z_i \in \{z_a, \ldots, z_b\}$ such that $z_j = z_i$ is true (by the definition of an equivalence relation). Also, there is exactly one $\delta^d(z_a, \ldots, z_b)$ that is true, by lemma 3.10. The conjunction of these true formulas is a constituent of the form $\delta^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ and is true. □

In the above lemma, the partition need not be unique. To show that more than one of the constituents of some depth and over some set of individual terms can be true (for the same model and valuation), here is an example: For the language with just one binary predicate P, some of the constituents are:

$\delta_1^0(a; b)$ is $P(a, a) \wedge (b = a)$

$\delta_2^0(a; b)$ is $\neg P(a, a) \wedge (b = a)$

$\delta_1^0(b; a)$ is $P(b, b) \wedge (a = b)$

$\delta_2^0(b; a)$ is $\neg P(b, b) \wedge (a = b)$

So $\delta_1^0(a; b) \Longleftrightarrow_= \delta_1^0(b; a)$ and similarly $\delta_2^0(a; b) \Longleftrightarrow_= \delta_2^0(b; a)$ (and these formulas are consistent).

This lack of uniqueness seems to mean that the constituents can not be divided into sets such that there is always exactly one constituent in each set that is true. For constituents of the same form there need not be any that are true, and for constituents over the same individual terms there can be more than one that is true. Because there are distinct equivalent consistent constituents (of the same depth and over the same set of individual terms), if we aim to form normal forms which are disjunctions of constituents of some depth and over some set of individual terms, not even the consistent constituents in a normal form will be unique.

Hintikka [1965a] (p. 60) claims that: "Any formula (with identity) whose depth is no greater than d and whose free individual symbols are all among the a_i's and b's can be converted into a disjunction of the constituents just described". Thereby he seems to say that any formula of depth $\leq d$ whose free individual terms are in $a_1, \ldots, a_k, b_1, \ldots b_m$ can be expressed as a disjunction of constituents of the form $\delta^d(a_1, \ldots, a_k; b_1, \ldots, b_m)$. This is clearly false. For example, the formula $x \neq y$ can not be expressed as any disjunction of constituents of the form $\delta^0(x; y)$ since each such constituent includes the conjunct $y = x$. It is not even the case that for any formula of depth $\leq d$ whose free individual terms are in $z_1, \ldots, z_n$ there is *some* partition into $a_1, \ldots, a_k$ and $b_1, \ldots b_m$ of $z_1, \ldots, z_n$ such that the formula can be expressed as a disjunction of constituents of the form $\delta^d(a_1, \ldots, a_k; b_1, \ldots, b_m)$. For example, to express $P(x, y)$ as a disjunction of constituents of depth 0 over x, y, it is necessary to have both some disjuncts containing $x = y$ (or $y = x$) and some containing $x \neq y$ (or $y \neq x$), thus $P(x, y)$ can not be expressed as a disjunction of constituents of only one form, though it can be expressed as a disjunction of constituents over x, y.

It seems to me that the result really is that any formula of depth $\leq d$ whose free individual terms are in $z_1, \ldots, z_n$ is equivalent to a disjunction of constituents, where for each of these constituents there is some partition into $a_1, \ldots, a_k$ and $b_1, \ldots b_m$ of $z_1, \ldots, z_n$ such that the constituent is of the form $\delta^d(a_1, \ldots, a_k; b_1, \ldots, b_m)$. This can be expressed more briefly as: every formula of depth $\leq d$ whose free individual terms are in $z_1, \ldots, z_n$ is equivalent to a disjunction of constituents of depth d over $z_1, \ldots, z_n$. The longer form emphisizes the difference in form of this theorem from the case without identity by showing that this case involves sets of constituents of different forms.

Although there are different constituents of the same depth and over the same set of individual terms that can be true together, this only occurs for constituents that are equivalent. If a formula F implies an identity $a = b$, then a and b are said to be *equal according to F*.

Lemma 3.12 *If $\delta^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ and $\delta^d(z_e, \ldots, z_f; z_g, \ldots, z_h)$ are both constituents over $z_1, \ldots, z_n$ and are both true for some valuation for some normal model, then $\delta^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ and $\delta^d(z_e, \ldots, z_f; z_g, \ldots, z_h)$ are equivalent.*

PROOF Let δ_i be $\delta^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$ and δ_j be $\delta^d(z_e, \ldots, z_f; z_g, \ldots, z_h)$. Each constituent over $z_1, \ldots, z_n$ implies, for all distinct $z_i, z_j \in \{z_1, \ldots, z_n\}$, either $z_i = z_j$ or $z_i \neq z_j$ in such a way that it divides $z_1, \ldots, z_n$ into equivalence classes with respect to $=$. So if δ_i and δ_j are both true then they agree on the equivalence classes into which $z_1, \ldots, z_n$ is divided by $=$. So $z_a, \ldots, z_b$ and $z_e, \ldots, z_f$ are the same except for possible reordering and substitution of terms which are equal according to δ_i (and δ_j). Thus, by lemma 3.10, the formulas $\delta^d(z_a, \ldots, z_b)$ and $\delta^d(z_e, \ldots, z_f)$ which occur in δ_i and δ_j respectively are the same (i.e. are notational variations of each other) except for the substitution of individual terms which are equal according to δ_i (and δ_j). Thus δ_i and δ_j are equivalent. □

Now, considering equivalence classes of constituents of depth d over $z_1, \ldots, z_n$ (where the equivalence is that denoted $\Longleftrightarrow_=$), we get the following:

Corollary 3.13 *For each depth d and each set of individual terms $z_1, \ldots, z_n$, for any normal model and valuation for it, the constituents of depth d over $z_1, \ldots, z_n$ in exactly one equivalence class (with respect to $\Longleftrightarrow_=$) are true.*

PROOF Combine lemmas 3.11 and 3.12. □

Lemma 3.14 *For each depth d, every atomic formula whose free individual terms are in $z_1, \ldots, z_n$ is equivalent to a disjunction of constituents, where for each of these constituents there is some partition $\{z_a, \ldots, z_b\}, \{z_c, \ldots, z_d\}$ of $z_1, \ldots, z_n$ such that the constituent is of the form $\delta_i^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$.*

PROOF If F is an atomic formula that is not an identity then F is equivalent to the disjunction of all the formulas of the form

$$F \wedge \bigwedge_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} (\pm)(z_i = z_j).$$

Those of these disjuncts that are inconsistent because they don't satisfy the conditions of identity may be omitted. Now, each disjunct may be simplified as follows: if it contains a conjunct $z_i = z_j$ and both z_i and z_j occur in F then replace every occurrence of z_i in F by z_j and omit redundant conjuncts. (See the example below.) This gives a formula which contains, for every distinct individual terms x, y in the atomic part, the conjunct $x \neq y$, and possibly some conjuncts of the form $(\pm)(y = z)$ where at least one of y, z is not in the atomic part. Each of these disjuncts can now be expressed as a disjunction of constituents of the form $\delta^0(z_a, \ldots, z_b; z_c, \ldots, z_d)$ where $z_a, \ldots, z_b$ are the individual terms occurring in the atomic part or negations of equalities, and $z_c, \ldots, z_d$ are the other individual terms in $z_1, \ldots, z_n$. This is achieved by expressing the atomic part in the appropriate propositional disjunctive normal form and then including the relevant equalities and inequalities. Thus F is equivalent to a disjunction of constituents of depth 0 over $z_1, \ldots, z_n$. Each of these disjuncts C is equivalent to the disjunction of all the constituents D of depth d over $z_1, \ldots, z_n$ such that each conjunct of C is a conjunct of D. If F is $z_i = z_j$ (for some $i, j \in \{1, \ldots, n\}$), then by lemma 3.11, for each normal model and valuation for the model for which $z_i = z_j$ is true, there is some constituent of depth d over $z_1, \ldots, z_n$ which is true. Thus $z_i = z_j$ implies the disjunction of all these constituents. And each of these constituents implies $z_i = z_j$ (because each constituent over a set which includes z_i and z_j

either implies $z_i = z_j$ or implies $z_i \neq z_j$). Thus the disjunction of these constituents implies $z_i = z_j$. □

Theorem 3.15 *Every formula of depth $\leq d$ whose free individual terms are contained in $z_1, \ldots, z_n$ is equivalent to a disjunction of constituents, where for each of these constituents there is some partition $\{z_a, \ldots, z_b\}, \{z_c, \ldots, z_d\}$ of $z_1, \ldots, z_n$ such that the constituent is of the form $\delta_i^d(z_a, \ldots, z_b; z_c, \ldots, z_d)$.*

PROOF Let F be a formula of depth $\leq d$ with free individual terms in $z_1, \ldots, z_n$. By lemma 3.9 (and its proof) F can be written in a form in which the interpretation of quantifiers can be changed without changing its depth or free individual terms. Thus we can take F to be a formula with its quantifiers interpreted exclusively. We use induction on the formation of the formula. If F is an atomic formula then F is equivalent to a disjunction of constituents of depth d over $z_1, \ldots, z_n$ by lemma 3.14. Assume that for each proper subformula G of F, G can be expressed as a disjunction of constituents of any depth greater than or equal to that of G and over any set of individual terms containing those of G. If F is $G \vee H$ then G and H each have depth $\leq d$ and free individual terms in $z_1, \ldots, z_n$, so by hypothesis each can be expressed as a disjunction of constituents of depth d over $z_1, \ldots, z_n$, thus so can F. If F is $\neg G$ then G has depth $\leq d$ and free individual terms in $z_1, \ldots, z_n$, so by hypothesis can be expressed as a disjunction D of constituents of depth d over $z_1, \ldots, z_n$. F is not necessarily equivalent to the disjunction of the other constituents of depth d over $z_1, \ldots, z_n$ (since it is possible for more than one such constituent to be true for the same normal model and valuation). But, by corollary 3.13, F is equivalent to the disjunction of all the constituents of depth d over $z_1, \ldots, z_n$ which are not equivalent to any element of D.

If F is $(\exists x)G$ then by changing the bound variables if necessary, we can assume that x is the first variable not in $z_1, \ldots, z_n$. Then G has depth $\leq d - 1$ and free individual terms in $z_1, \ldots, z_n, x$. So by hypothesis, G can be expressed as the disjunction D of some constituents of depth $d - 1$ over $z_1, \ldots, z_n, x$. Each of these constituents C is of the form $\delta^{d-1}(z_a, \ldots, z_b; z_c, \ldots, z_d)$ for some partition $\{z_a, \ldots, z_b\}, \{z_c, \ldots, z_d\}$ of $\{z_1, \ldots, z_n, x\}$. If $x \in \{z_c, \ldots, z_d\}$ then C contains a conjunct $x = z_i$ for some $z_i \in \{z_a, \ldots, z_b\}$. Instead of C we may use the equivalent constituent which is just like C except that x and z_i have been substituted for each other. Thus each constituent C which is a disjunct of D may be taken to be of the form $\delta^{d-1}(z_a, \ldots, z_b; z_c, \ldots, z_d)$ with $x \in \{z_a, \ldots, z_b\}$. By distributing the existential quantification over disjunction, F is equivalent to a disjunction of formulas of the form $(\exists x)\delta^{d-1}(z_a, \ldots, z_b; z_c, \ldots, z_d)$. For each of these disjuncts, let $z_e, \ldots, z_f$ be $\{z_a, \ldots, z_b\} \setminus \{x\}$ (in any order). Then $\delta^{d-1}(z_a, \ldots, z_b; z_c, \ldots, z_d)$ may be written as $\delta^{d-1}(z_e, \ldots, z_f, x; z_c, \ldots, z_d)$ and is of the form

$$\bigwedge_{\substack{i,j \in \{a,\ldots,b\} \\ i \neq j}} (z_i \neq z_j) \ \wedge$$

$$
\begin{aligned}
& \delta^{d-1}(z_a, \ldots, z_b) \wedge \\
& (z_c = z_{i_1}) \wedge \ldots \wedge (z_d = z_{i_k}) \\
\Longleftrightarrow_{=} \quad & \\
& \bigwedge_{\substack{i,j \in \{a,\ldots,b\} \\ i \neq j}} (z_i \neq z_j) \wedge \\
& \delta^0(z_e, \ldots, z_f) \wedge \gamma_m^{d-1}(z_e, \ldots, z_f, x) \wedge \\
& (z_c = z_{i_1}) \wedge \ldots \wedge (z_d = z_{i_k}).
\end{aligned}
$$

So

$$
\begin{aligned}
& (\exists x)\delta^{d-1}(z_e, \ldots, z_f, x; z_c, \ldots, z_d) \\
\Longleftrightarrow_{=} \quad & \\
& \bigwedge_{\substack{i,j \in \{e,\ldots,f\} \\ i \neq j}} (z_i \neq z_j) \wedge \\
& \delta^0(z_e, \ldots, z_f) \wedge (\exists x)(\gamma_m^{d-1}(z_e, \ldots, z_f, x) \wedge (x \neq z_e) \wedge \ldots \wedge (x \neq z_f)) \wedge \\
& (z_c = z_{i_1}) \wedge \ldots \wedge (z_d = z_{i_k})
\end{aligned}
$$

by moving the conjuncts which don't contain x out of the scope of $(\exists x)$.

Each formula $\delta^0(z_e, \ldots, z_f) \wedge (\exists x)(\gamma_m^{d-1}(z_e, \ldots, z_f, x) \wedge (x \neq z_e) \wedge \ldots \wedge (x \neq z_f))$ is $\delta^0(z_e, \ldots, z_f) \wedge (\exists x)\gamma_m^{d-1}(z_e, \ldots, z_f, x)$ with the quantifier interpreted exclusively, and is equivalent to the disjunction of all the formulas of the form

$$
\begin{aligned}
& \delta^0(z_e, \ldots, z_f) \wedge \\
& (\pm)(\exists x)\gamma_1^{d-1}(z_e, \ldots, z_f, x) \wedge \ldots \wedge (\exists x)\gamma_m^{d-1}(z_e, \ldots, z_f, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_p^{d-1}(z_e, \ldots, z_f, x)
\end{aligned}
$$

which is a formula of the form $\delta^d(z_e, \ldots, z_f)$ (that is, a formula like a constituent for the case without identity, but with quantifiers interpreted exclusively). Let $\{\delta_i^d(z_e, \ldots, z_f)\}_{i \in I}$ be the formulas in this disjunction. Then

$$
\begin{aligned}
& (\exists x)\delta^{d-1}(z_e, \ldots, z_f, x; z_c, \ldots, z_d) \\
\Longleftrightarrow_{=} \quad & \\
& \bigwedge_{\substack{i,j \in \{e,\ldots,f\} \\ i \neq j}} (z_i \neq z_j) \wedge \\
& \bigvee_{i \in I} \delta_i^d(z_e, \ldots, z_f) \wedge \\
& (z_c = z_{i_1}) \wedge \ldots \wedge (z_d = z_{i_k}) \\
\Longleftrightarrow_{=} \quad & \\
& \bigvee_{i \in I} (\bigwedge_{\substack{i,j \in \{e,\ldots,f\} \\ i \neq j}} (z_i \neq z_j) \wedge
\end{aligned}
$$

$$\begin{aligned}&\delta_i^d(z_e, \ldots, z_f) \wedge \\ &(z_c = z_{i_1}) \wedge \ldots \wedge (z_d = z_{i_k})).\end{aligned}$$

Because the quantifiers in F are interpreted exclusively, none of the $z_{i_1}, \ldots, z_{i_k}$ can be x, thus $z_{i_1}, \ldots, z_{i_k} \in \{z_e, \ldots, z_f\}$. So each disjunct of the above formula is a constituent of the form $\delta^d(z_e, \ldots, z_f; z_c, \ldots, z_d)$. Thus F is a disjunction of constituents of depth d over $z_1, \ldots, z_n$. □

Any disjunction of depth-d constituents over $z_1, \ldots, z_n$ to which a formula F is equivalent is called a *distributive normal form at depth d over* $z_1, \ldots, z_n$ of F.

The following example illustrates some of the steps used in the above proof. It also shows that even depth-0 formulas may have very large distributive normal forms. To express $P(x) \vee Q(x, y, z)$ in distributive normal form:

$$\begin{aligned}
P(x) \iff_= \; & (P(x) \wedge (x = y) \wedge (y = z) \wedge (x = z)) \vee \\
& (P(x) \wedge (x = y) \wedge (y \neq z) \wedge (x \neq z)) \vee \\
& (P(x) \wedge (x \neq y) \wedge (y = z) \wedge (x \neq z)) \vee \\
& (P(x) \wedge (x \neq y) \wedge (y \neq z) \wedge (x = z)) \vee \\
& (P(x) \wedge (x \neq y) \wedge (y \neq z) \wedge (x \neq z)) \\
\iff_= \; & (P(x) \wedge (y = x) \wedge (z = x)) \vee \\
& (P(x) \wedge (x \neq z) \wedge (y = x)) \vee \\
& (P(x) \wedge (x \neq y) \wedge (z = y)) \vee \\
& (P(x) \wedge (x \neq y) \wedge (z = x)) \vee \\
& (P(x) \wedge (x \neq y) \wedge (y \neq z) \wedge (x \neq z))
\end{aligned}$$

$$\begin{aligned}
Q(x, y, z) \iff_= \; & (Q(x, y, z) \wedge (x = y) \wedge (y = z) \wedge (x = z)) \vee \\
& (Q(x, y, z) \wedge (x = y) \wedge (y \neq z) \wedge (x \neq z)) \vee \\
& (Q(x, y, z) \wedge (x \neq y) \wedge (y = z) \wedge (x \neq z)) \vee \\
& (Q(x, y, z) \wedge (x \neq y) \wedge (y \neq z) \wedge (x = z)) \vee \\
& (Q(x, y, z) \wedge (x \neq y) \wedge (y \neq z) \wedge (x \neq z)) \\
\iff_= \; & (Q(x, x, x) \wedge (y = x) \wedge (z = x)) \vee \\
& (Q(x, x, z) \wedge (x \neq z) \wedge (y = x)) \vee \\
& (Q(x, y, y) \wedge (x \neq y) \wedge (z = y)) \vee \\
& (Q(x, y, x) \wedge (x \neq y) \wedge (z = x)) \vee \\
& (Q(x, y, z) \wedge (x \neq y) \wedge (y \neq z) \wedge (x \neq z))
\end{aligned}$$

To find the formulas of the form $\delta^0(x,y,z)$:

$$\delta^0(x) = (\pm)P(x) \wedge (\pm)Q(x,x,x)$$

$$\begin{aligned}\delta^0(x,y) = {} & (\pm)P(x) \wedge (\pm)Q(x,x,x) \wedge \\ & (\pm)P(y) \wedge (\pm)Q(x,x,y) \wedge (\pm)Q(x,y,x) \wedge (\pm)Q(y,x,x) \wedge \\ & (\pm)Q(x,y,y) \wedge (\pm)Q(y,x,y) \wedge (\pm)Q(y,y,x) \wedge (\pm)Q(y,y,y)\end{aligned}$$

$$\begin{aligned}\delta^0(x,y,z) = {} & (\pm)P(x) \wedge (\pm)Q(x,x,x) \wedge \\ & (\pm)P(y) \wedge (\pm)Q(x,x,y) \wedge (\pm)Q(x,y,x) \wedge (\pm)Q(y,x,x) \wedge \\ & (\pm)Q(x,y,y) \wedge (\pm)Q(y,x,y) \wedge (\pm)Q(y,y,x) \wedge (\pm)Q(y,y,y) \wedge \\ & (\pm)P(z) \wedge (\pm)Q(x,x,z) \wedge (\pm)Q(x,z,x) \wedge (\pm)Q(z,x,x) \wedge \\ & (\pm)Q(x,z,z) \wedge (\pm)Q(z,x,z) \wedge (\pm)Q(z,z,x) \wedge \\ & (\pm)Q(y,y,z) \wedge (\pm)Q(y,z,y) \wedge (\pm)Q(z,y,y) \wedge \\ & (\pm)Q(y,z,z) \wedge (\pm)Q(z,y,z) \wedge (\pm)Q(z,z,y) \wedge \\ & (\pm)Q(x,y,z) \wedge (\pm)Q(x,z,y) \wedge (\pm)Q(y,x,z) \wedge (\pm)Q(y,z,x) \wedge \\ & (\pm)Q(z,x,y) \wedge (\pm)Q(z,y,x) \wedge (\pm)Q(z,z,z)\end{aligned}$$

Now use these formulas to form constituents of the following forms:

$$\begin{aligned}\delta^0(x,y,z;) &= (x \neq y) \wedge (y \neq z) \wedge (x \neq z) \wedge \delta^0(x,y,z) \\ \delta^0(x,y;z) &= (x \neq y) \wedge \delta^0(x,y) \wedge (z = z_i) \text{ where } z_i \in \{x,y\} \\ \delta^0(x,z;y) &= (x \neq z) \wedge \delta^0(x,z) \wedge (y = z_i) \text{ where } z_i \in \{x,z\} \\ \delta^0(y,z;x) &= (y \neq z) \wedge \delta^0(y,z) \wedge (x = z_i) \text{ where } z_i \in \{y,z\} \\ \delta^0(x;y,z) &= \delta^0(x) \wedge (y = x) \wedge (z = x) \\ \delta^0(y;x,z) &= \delta^0(y) \wedge (x = y) \wedge (z = y) \\ \delta^0(z;x,y) &= \delta^0(z) \wedge (x = z) \wedge (y = z)\end{aligned}$$

So there are $2^{30} + 3.2^{11} + 12$ constituents we have to consider, and this is only a 3-variable depth-0 formula.

Finally, we express the elements of the disjunctions to which $P(x)$ and $Q(x,y,z)$ are equivalent as disjunctions of the appropriate constituents. I will not write out these disjunctions in full since they are too big, but rather describe the disjuncts as all the formulas of a certain form:

We consider the disjuncts for $P(x)$ first:

$P(x) \wedge (y = x) \wedge (z = x)$ is equivalent to

$$(P(x) \wedge Q(x,x,x) \wedge (y = x) \wedge (z = x)) \vee$$
$$(P(x) \wedge \neg Q(x,x,x) \wedge (y = x) \wedge (z = x))$$

which is the disjunction of the constituents of the form

$$P(x) \wedge (\pm)Q(x,x,x) \wedge (y = x) \wedge (z = x).$$

$P(x) \wedge (x \neq z) \wedge (y = x)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq z) \wedge P(x) \wedge (\pm)Q(x,x,x) \wedge$$
$$(\pm)P(z) \wedge (\pm)Q(x,x,z) \wedge (\pm)Q(x,z,x) \wedge (\pm)Q(z,x,x) \wedge$$
$$(\pm)Q(x,z,z) \wedge (\pm)Q(z,x,z) \wedge (\pm)Q(z,z,x) \wedge (\pm)Q(z,z,z) \wedge (y = x).$$

$P(x) \wedge (x \neq y) \wedge (z = y)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq y) \wedge P(x) \wedge (\pm)Q(x,x,x) \wedge$$
$$(\pm)P(z) \wedge (\pm)Q(x,x,y) \wedge (\pm)Q(x,y,x) \wedge (\pm)Q(y,x,x) \wedge$$
$$(\pm)Q(x,y,y) \wedge (\pm)Q(y,x,y) \wedge (\pm)Q(y,y,x) \wedge (\pm)Q(y,y,y) \wedge (z = y).$$

$P(x) \wedge (x \neq y) \wedge (z = x)$ is equivalent to the disjunction of the constituents of the same form as above except that the final conjunct is $(z = x)$.

$P(x) \wedge (x \neq y) \wedge (y \neq z) \wedge (x \neq z)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq y) \wedge (y \neq z) \wedge (x \neq z) \wedge \delta^0(x,y,z)$$

where $\delta^0(x,y,z)$ has $P(x)$ as a conjunct.

Similarly, the disjuncts for $Q(x,y,z)$:

$Q(x,x,x) \wedge (y = x) \wedge (z = x)$ is equivalent to the disjunction of the constituents of the form

$$(\pm)P(x) \wedge Q(x,x,x) \wedge (y = x) \wedge (z = x)$$

$Q(x,x,z) \wedge (x \neq z) \wedge (y = x)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq z) \wedge (\pm)P(x) \wedge (\pm)Q(x,x,x) \wedge$$
$$(\pm)P(z) \wedge Q(x,x,z) \wedge (\pm)Q(x,z,x) \wedge (\pm)Q(z,x,x) \wedge$$
$$(\pm)Q(x,z,z) \wedge (\pm)Q(z,x,z) \wedge (\pm)Q(z,z,x) \wedge (\pm)Q(z,z,z) \wedge (y = x)$$

$Q(x,y,y) \wedge (x \neq y) \wedge (z = y)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq y) \wedge (\pm)P(x) \wedge (\pm)Q(x,x,x) \wedge$$
$$(\pm)P(y) \wedge (\pm)Q(x,x,y) \wedge (\pm)Q(x,y,x) \wedge (\pm)Q(y,x,x) \wedge$$
$$Q(x,y,y) \wedge (\pm)Q(y,x,y) \wedge (\pm)Q(y,y,x) \wedge (\pm)Q(y,y,y) \wedge (z = y)$$

$Q(x,y,x) \wedge (x \neq y) \wedge (z = x)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq y) \wedge (\pm)P(x) \wedge (\pm)Q(x,x,x) \wedge$$
$$(\pm)P(y) \wedge (\pm)Q(x,x,y) \wedge Q(x,y,x) \wedge (\pm)Q(y,x,x) \wedge$$
$$(\pm)Q(x,y,y) \wedge (\pm)Q(y,x,y) \wedge (\pm)Q(y,y,x) \wedge (\pm)Q(y,y,y) \wedge (z = x)$$

$Q(x,y,z) \wedge (x \neq y) \wedge (y \neq z) \wedge (x \neq z)$ is equivalent to the disjunction of the constituents of the form

$$(x \neq y) \wedge (y \neq z) \wedge (x \neq z) \wedge \delta^0(x,y,)$$

where $\delta^0(x,y,z)$ has $Q(x,y,z)$ as a conjunct.

So $P(x) \vee Q(x,y,z)$ is equivalent to the disjunction of all the constituents just described. The number of constituents in this disjunction is $2+2^9+2^9+2^9+2^{29}+1+2^8+2^8+2^8+2^{28} = 2307+2^{28}+2^{29}$.

Corollary 3.16 *Any formula of depth $\leq d$ whose free individual terms are $z_1,\ldots,z_k$ and which is of the form $F \wedge \bigwedge_{\substack{i,j\in\{1,\ldots,k\}\\ i\neq j}}(z_i \neq z_j)$ is equivalent to a disjunction of constituents of the form $\delta_i^d(z_1,\ldots,z_k;\ )$.*

PROOF By theorem 3.15, a formula G of depth $\leq d$ whose free individual terms are $z_1,\ldots,z_k$ is equivalent to a disjunction of constituents of depth d over $z_1,\ldots,z_k$. If G contains a conjunct $z_i \neq z_j$ for each $i,j \in \{1,\ldots,k\}$, $i \neq j$, then all the constituents in the disjunction must be of the form $\delta^d(z_1,\ldots,z_k;)$. □

Thus, slightly less formally, we can say that any formula for which all distinct free individual terms are explicitly not equal has a distributive normal form in terms of main parts of constituents. Since at most one constituent of the form $\delta^d(z_1,\ldots,z_k;)$ can be true for some normal model and valuation for it, the consistent constituents in a distributive normal form of depth d over $z_1,\ldots,z_k$ of such a formula are unique. Regarding the definition of constituents, Hintikka [1965a] (p. 60) says "It is obvious that formulae of the latter form" [equalities that are conjuncts of a constituent, but not of the main part] "cannot affect the consistency or inconsistency of the constituent in question. Hence they may be disregarded for many purposes". This seems to say that for many purposes, the only constituents that need be used are the main parts of constituents. I don't know what purposes are meant here. The above corollary has the following converse: If a formula G is equivalent to a disjunction of constituents of the form $\delta_i^d(z_1, \ldots, z_k;)$ then G implies $z_i \neq z_j$ for each $i, j \in \{1, \ldots, k\}$, $i \neq j$. Thus a formula of depth $\leq d$ whose free individual terms are $z_1, \ldots, z_k$ may be expressed as a disjunction of main parts of constituents of the form $\delta_i^d(z_1, \ldots, z_k;)$ iff it implies $z_i \neq z_j$ for each $i, j \in \{1, \ldots, k\}$, $i \neq j$. In particular, all constant-free sentences can be expressed as disjunctions of main parts of constituents. Thus, in a context where constant-free sentences are used to express "statements" (some examples of which are given in chapter 5), only main parts of constituents need be used. (This could be what Hintikka was referring to in the above quote.)

Languages with function symbols

Here also there are at least two different approaches. The first approach involves constructing a new language which does not contain function symbols and does contain identity. If the original language does not contain identity, we first extend it to contain identity in order to more easily prove the desired result.

All n-ary operations are particular $(n+1)$-ary relations. To be definite, an n-ary operation f is taken to be the relation such that $x = f(t_1, \ldots, t_n)$ iff $\langle t_1, \ldots, t_n, x \rangle \in f$. For a language L with identity, let L' be the language which contains the same symbols as L except that for each n, for each n-ary function symbol f of L, L' contains an $(n+1)$-ary relation symbol R_f instead of f. For any model $\mathcal{M} = \langle D, I \rangle$ for L, let $\mathcal{M}' = \langle D, I' \rangle$ be the model for L' where I' is the same as I for all symbols of L' that are the same as L and for each function symbol f of L, $I'(R_f) = I(f)$ (i.e. R_f is interpreted as the relation which is the function as which f is interpreted).

In the scope A of any quantifier in a formula, the *next unused variable* is the first variable z such that A is not in the scope of any quantifier over z, nor does z occur free in A.

For a formula which starts with some initial string of quantifiers, the *inside* of the formula is that part which is in the scope of all the initial quantifiers. For a formula which doesn't start with a quantifier, the *inside* is the whole formula.

For any formula F of a language L with identity, let F' be the formula of L' defined by:

For each atomic formula $A(t_1, \ldots, t_n)$ of F that contains function symbols

1. for each t_i, $i \in \{1, \ldots, n\}$, if t_i is a variable or constant symbol then leave it, otherwise start at the innermost level of nested function symbols and replace each occurrence of a function symbol and its arguments, which together I will call X, by the next unused

variable x and conjoin $x = X$ to the inside of the formula, and prefix the formula by $(\exists x)$.

2. replace each identity $x = f(a_1, \ldots, a_m)$ just introduced by $R_f(a_1, \ldots, a_m, x)$, let $z_1, \ldots, z_m, z_{m+1}, z_{m+2}$ be the next $m+2$ unused variables and replace $R_f(a_1, \ldots, a_m, x)$ by

$$((\forall z_1) \ldots (\forall z_{m+2})((R_f(z_1, \ldots, z_m, z_{m+1}) \wedge R_f(z_1, \ldots, z_m, z_{m+2})) \rightarrow (z_{m+1} = z_{m+2}))))$$
$$\rightarrow R_f(a_1, \ldots, a_m, x).$$

The resulting formula may be propositionally simplified so that the formula

$$(\forall z_1) \ldots (\forall z_{m+2})((R_f(z_1, \ldots, z_m, z_{m+1}) \wedge R_f(z_1, \ldots, z_m, z_{m+2})) \rightarrow (z_{m+1} = z_{m+2})))$$

only need be included once for each distinct R_f in the formula, as in the following example.

The formula obtained after the first step in converting F to F' will be denoted F^*.

We now consider an example of converting a formula F to F'. Let F be $R(f(f(a)), b)$ then F^* is

$$(\exists x_2)(\exists x_1)(R(x_2, b) \wedge (x_1 = f(a)) \wedge (x_2 = f(x_1))),$$

and F' is

$$(\exists x_2)(\exists x_1)(R(x_2, b) \wedge (((\forall x_3)(\forall x_4)(\forall x_5)((R_f(x_3, x_4) \wedge R_f(x_3, x_5)) \rightarrow (x_4 = x_5)))) \rightarrow$$
$$(R_f(a, x_1) \wedge R_f(x_1, x_2))).$$

The purpose for including the subformula $((\forall x_3)(\forall x_4)(\forall x_5)((R_f(x_3, x_4) \wedge R_f(x_3, x_5)) \rightarrow (x_4 = x_5)))$ in the above formula is that it is true exactly when R_f is interpreted as a function. We need such formulas for the result that F is valid iff F' is valid (see theorem 3.19).

The formulas F and F' are not equivalent since they are of different languages, but their relation to each other is something like equivalence. The following theorems state the sense in which a formula containing function symbols may be replaced by one which does not.

Lemma 3.17 *Each formula F of a language with identity is equivalent to F^*.*

PROOF After the first step in converting F to F', substituting terms that are equal according to the formula and omitting the quantification that becomes irrelevant results in the original formula F. □

Theorem 3.18 *For any formula F of a language L with identity, for any model $\mathcal{M}$ for L and any valuation v for $\mathcal{M}$, $\mathcal{M}, v \models F$ iff $\mathcal{M}', v \models F'$.*

PROOF First note that since L and L' have the same variables and $\mathcal{M}$ and $\mathcal{M}'$ have the same domains, any valuation for $\mathcal{M}$ may also be considered as a valuation for $\mathcal{M}'$ by considering it as a function from the variables to the domain, and for each model extending it to a function from the terms of the relevant language to the domain. If the result holds for all atomic formulas F then it holds for all formulas. Let F be an atomic formula. By lemma 3.17 we need only show that $\mathcal{M}, v \models F^*$ iff $\mathcal{M}', v \models F'$. F^* contains no nested function symbols, and each occurrence of a function symbol f is in a subformula of the form $x = f(t_1, \ldots, t_n)$. $\mathcal{M}, v \models (x = f(t_1, \ldots, t_n))$ iff $v(x) = v(f(t_1, \ldots, t_n))$ iff $v(x) = I(f)(v(t_1), \ldots, v(t_n))$ iff $\langle v(t_1), \ldots, v(t_n), v(x) \rangle \in I(f)$ iff $\langle v(t_1), \ldots, v(t_n), v(x) \rangle \in I'(R_f)$ iff $\mathcal{M}', v \models R_f(t_1, \ldots, t_n, x)$. So $\mathcal{M}', v \models F'$ whether or not $\mathcal{M}', v$ satisfies the formulas of the form $(\forall z_1) \ldots (\forall z_{m+2})((R_f(z_1, \ldots, z_m, z_{m+1}) \wedge R_f(z_1, \ldots, z_m, z_{m+2})) \rightarrow (z_{m+1} = z_{m+2})))$ ($\mathcal{M}', v$ actually does satisfy these formulas since in $\mathcal{M}'$, R_f is interpreted as a function). So, for this theorem, it is not necessary to include the formulas $(\forall z_1) \ldots (\forall z_{m+2})((R_f(z_1, \ldots, z_m, z_{m+1}) \wedge R_f(z_1, \ldots, z_m, z_{m+2})) \rightarrow (z_{m+1} = z_{m+2})))$ in the definition of F'. □

Theorem 3.19 *For any formula F of a language with identity, $\models F$ iff $\models F'$.*

PROOF By lemma 3.17 we need only show that $\models F^*$ iff $\models F'$. The formula F^* contains function symbols f only in subformulas of the form $x = f(t_1, \ldots, t_n)$. If F^* contains a function symbol f and is valid then it is true on all domains for all interpretations of f, so the formula in which each $x = f(t_1, \ldots, t_n)$ is replaced by $R_f(t_1, \ldots, t_n, x)$ is true on all domains for all interpretations of R_f as a function, so the formula in which each $R_f(t_1, \ldots, t_n, x)$ is now replaced by

$$((\forall z_1) \ldots (\forall z_{m+2})((R_f(z_1, \ldots, z_m, z_{m+1}) \wedge R_f(z_1, \ldots, z_m, z_{m+2})) \rightarrow (z_{m+1} = z_{m+2})))) \rightarrow R_f(t_1, \ldots, t_n, x),$$

which is the formula F', is true on all domains for all interpretations. Thus F' is valid. Each of these steps also goes the other direction, thus the converse also holds. □

Theorems 3.18 and 3.19 show that we can find whether or not a formula F with function symbols is true for a particular model and valuation, or is valid, if we can find a similar result for the formula F' without function symbols. So, for many purposes formulas without function symbols may be used instead of formulas containing function symbols, though the change increases the depth of the formula, makes it notably longer, and necessitates the inclusion of identity in the language. For such formulas without function symbols, distributive normal forms have been defined.

The other approach is to define constituents for languages with function symbols. We first consider the case where there are only finitely many function symbols in the language. It can then be extended to the case where there are infinitely many in a similar manner to the way the case for finitely many predicate symbols was extended to the case for infinitely many. Each formula contains only finitely many function symbols. All the definitions can be considered to be relative to some finite subset of the set of function symbols. Then each formula will have a distributive normal form in terms of constituents relative to any finite superset (of function symbols) of the function symbols in the formula. But we will not include these sets of function symbols in the notation for constituents since that would make it unnecessarily complex.

For formulas in which no terms containing function symbols contain any variables, constituents can be defined in a way which is considerably simpler than the definition I am about to give. A formula contains only finitely many terms. In forming constituents, use the terms of the formula which contain function symbols as the constant symbols were used in the formation of the atomic formulas. This doesn't work if there are terms which contain function symbols and variables. (This can be seen by considering the example $(\exists x)P(x, f(x))$.) So I don't consider this approach any further.

To form constituents and attributive constituents, we need to divide the atomic formulas into finite sets. We want to do this in such a way that given any formula, there will be a finite set of constituents for which we can show that the disjunction of some of its elements is equivalent to the formula. There might be more than one possible way in which this can be done. In this case, I don't think there is much point in asking what the best way is, since for the case without function symbols, distributive normal forms are already too big to literally represent formulas in (except in some special cases), though the *existence* of distributive normal form might be useful for proving general results. Thus my aim in this section is just to show the existence of a normal form which can be regarded as a generalization of the distributive normal form already defined.

The *functional depth* of a formula is the maximum depth of the terms it contains, where the *depth* of a term t, $d(t)$, is the number of levels of nested applications of function symbols, defined by: $d(x) = 0$ if x is a variable or a constant symbol; $d(f(t_1, \ldots, t_n)) = 1 + \max_{i \in \{1,\ldots,n\}} d(t_i)$ for any n-ary function symbol f and terms $t_1, \ldots, t_n$.

If $z_1, \ldots, z_k$ are individual terms then the atomic formulas of functional depth w over $\{z_1, \ldots, z_k\}$ which contain z_k are denoted ${}_w\alpha_i(z_1, \ldots, z_k)$.

For the case without function symbols, the atomic formulas were divided into different "levels". In this case they are divided by both "level" and functional depth. For example, for the language with one unary function symbol f and one binary predicate symbol P:

$$
\begin{array}{lll}
{}_0\alpha_1(x) = P(x,x) & {}_1\alpha_1(x) = P(x, f(x)) & \cdots \\
 & {}_1\alpha_2(x) = P(f(x), x) & \\
 & {}_1\alpha_3(x) = P(f(x), f(x)) &
\end{array}
$$

$$
\begin{array}{lll}
{}_0\alpha_1(x,y) = P(x,y) & {}_1\alpha_1(x,y) = P(x, f(y)) & \cdots \\
{}_0\alpha_2(x,y) = P(y,x) & {}_1\alpha_2(x,y) = P(y, f(x)) & \\
{}_0\alpha_3(x,y) = P(y,y) & {}_1\alpha_3(x,y) = P(y, f(y)) & \\
 & {}_1\alpha_4(x,y) = P(f(x), y) & \\
 & {}_1\alpha_5(x,y) = P(f(x), f(y)) & \\
 & {}_1\alpha_6(x,y) = P(f(y), x) & \\
 & {}_1\alpha_7(x,y) = P(f(y), y) & \\
 & {}_1\alpha_8(x,y) = P(f(y), f(x)) & \\
 & {}_1\alpha_9(x,y) = P(f(y), f(y)) & \\
 & & \\
\vdots & \vdots &
\end{array}
$$

An *attributive constituent of depth d and functional depth w with free individual terms* $z_1, \ldots, z_k$ is denoted ${}_w\gamma^d(z_1, \ldots, z_k)$ and defined by

$$
\begin{aligned}
{}_w\gamma_i^0(z_1, \ldots, z_k) &= (\pm)_0\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)_w\alpha_q(z_1, \ldots, z_k) \\
{}_w\gamma_i^d(z_1, \ldots, z_k) &= (\pm)_0\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)_w\alpha_q(z_1, \ldots, z_k) \wedge \\
&\quad (\pm)(\exists x)_w\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)_w\gamma_r^{d-1}(z_1, \ldots, z_k, x)
\end{aligned}
$$

where ${}_0\alpha_1(z_1, \ldots, z_k), \ldots, {}_w\alpha_q(z_1, \ldots, z_k)$ are all the atomic formulas of functional depths 0 to w over $\{z_1, \ldots, z_k\}$ which contain z_k.

A *constituent of depth d and functional depth w with free individual terms* $z_1, \ldots, z_k$ is denoted ${}_w\delta^d(z_1, \ldots, z_k)$ and defined by

$$
{}_w\delta_i^0(z_1, \ldots, z_k) = (\pm)_0\alpha_1(z_1) \wedge \ldots \wedge (\pm)_w\alpha_p(z_1, \ldots, z_k)
$$

$$
\begin{aligned}
{}_w\delta_i^d(z_1,\ldots,z_k) \;=\; & (\pm)_0\alpha_1(z_1) \wedge \ldots \wedge (\pm)_w\alpha_p(z_1,\ldots,z_k) \wedge \\
& (\pm)(\exists x)_w\gamma_1^{d-1}(z_1,\ldots,z_k,x) \wedge \ldots \wedge (\pm)(\exists x)_w\gamma_r^{d-1}(z_1,\ldots,z_k,x)
\end{aligned}
$$

where ${}_0\alpha_1(z_1), \ldots, {}_w\alpha_p(z_1,\ldots,z_k)$ are all the atomic formulas of functional depths 0 to w over $\{z_1,\ldots,z_k\}$.

In both of the above definitions x is the first variable not in $z_1,\ldots,z_k$ and ${}_w\gamma_1^{d-1}(z_1,\ldots, z_k,x), \ldots, {}_w\gamma_r^{d-1}(z_1,\ldots,z_k,x)$ are all the attributive constituents of depth $d-1$ and functional depth w with free individual terms $z_1,\ldots,z_k,x$, and i is the index by which the formulas are listed.

A constituent ${}_w\delta^d(z_1,\ldots,z_k)$ (as defined above) has the form

$$(\pm)_0\alpha_1(z_1) \wedge \ldots \wedge (\pm)_w\alpha_p(z_1,\ldots,z_{k-1}) \wedge {}_w\gamma^d(z_1,\ldots,z_k),$$

or equivalently,

$${}_w\delta^0(z_1,\ldots,z_{k-1}) \wedge {}_w\gamma^d(z_1,\ldots,z_k).$$

Lemma 3.20 *For each model and each valuation for it, for each set of individual terms $z_1,\ldots,z_k$, for each depth d, for each functional depth w, exactly one constituent of the form ${}_w\delta_i^d(z_1,\ldots,z_k)$ is true.*

PROOF The constituents of the form ${}_w\delta_i^d(z_1,\ldots,z_k)$ are the basic conjunctions generated by some set of formulas, so by an observation in chapter 1, exactly one such constituent is true for any given model and valuation. □

Theorem 3.21 *Every formula of depth $\leq d$ and functional depth $\leq w$ whose free individual terms are contained in $z_1,\ldots,z_k$ can be expressed as a disjunction of constituents of depth d and functional depth w with free individual terms $z_1,\ldots,z_k$.*

PROOF We use induction on the formation of the formula. Let F be a formula of depth $\leq d$ and functional depth $\leq w$ with free individual terms in $z_1,\ldots,z_k$. If F is an atomic formula, then F is an element of the set which generates the formulas ${}_w\delta_i^d(z_1,\ldots,z_k)$, thus F is equivalent to the disjunction of all the formulas ${}_w\delta_i^d(z_1,\ldots,z_k)$ in which F occurs unnegated. Assume that for each proper subformula G of F, G can be expressed as a disjunction of constituents of any depth greater than or equal to that of G, functional depth greater than or equal to that of G, and free individual terms containing those of G. If F is $(\exists x)G$ then by changing the bound variables if necessary, we can assume that x is the first variable not in $z_1,\ldots,z_k$. Then G is a formula of depth $\leq d-1$ and functional depth $\leq w$ with free individual terms in $z_1,\ldots,z_k,x$, so by hypothesis can be expressed as some disjunction

$$\bigvee_{i\in I} {}_w\delta_i^{d-1}(z_1,\ldots,z_k,x)$$

and each ${}_w\delta_i^{d-1}(z_1,\ldots,z_k,x)$ is of the form ${}_w\delta_{i_a}^0(z_1,\ldots,z_k) \wedge {}_w\gamma_{i_b}^{d-1}(z_1,\ldots,z_k,x)$. So

$$(\exists x)G \iff \bigvee_{i\in I}({}_w\delta_{i_a}^0(z_1,\ldots,z_k) \wedge (\exists x)_w\gamma_{i_b}^{d-1}(z_1,\ldots,z_k,x)).$$

Each ${}_w\delta_{i_a}^0(z_1,\ldots,z_k) \wedge (\exists x)_w\gamma_{i_b}^{d-1}(z_1,\ldots,z_k,x)$ is equivalent to the disjunction of all the formulas of the form

$$\begin{aligned}
&{}_w\delta_{i_a}^0(z_1,\ldots,z_k)\wedge \\
&(\pm)(\exists x)_w\gamma_1^{d-1}(z_1,\ldots,z_k,x) \wedge \ldots \wedge (\exists x)_w\gamma_{i_b}^{d-1}(z_1,\ldots,z_k,x) \wedge \ldots \wedge \\
&\qquad\qquad\qquad\qquad\qquad\qquad (\pm)(\exists x)_w\gamma_p^{d-1}(z_1,\ldots,z_k,x)
\end{aligned}$$

and each of the above formulas is some ${}_w\delta_j^d(z_1,\ldots,z_k)$, so F is equivalent to a disjunction of some ${}_w\delta_j^d(z_1,\ldots,z_k)$'s. If F is $G \vee H$ then G and H each have depth $\leq d$, functional depth $\leq w$ and free individual terms in $z_1,\ldots,z_k$, so each is equivalent to a disjunction of some ${}_w\delta_j^d(z_1,\ldots,z_k)$'s, thus so is F. If F is $\neg G$ then G has depth $\leq d$, functional depth $\leq w$ and free individual terms in $z_1,\ldots,z_k$, so is equivalent to a disjunction of some ${}_w\delta_j^d(z_1,\ldots,z_k)$'s, thus F is equivalent to the disjunction of all the other constituents of the form ${}_w\delta_j^d(z_1,\ldots,z_k)$ by lemma 3.20. □

As before, a disjunction of constituents of depth d and functional depth w with free individual terms $z_1,\ldots,z_k$ to which a formula F is equivalent is called a *distributive normal form at depth d and functional depth w with free individual terms* $z_1,\ldots,z_k$ of F. By lemma 3.20, the set of consistent constituents in such a distributive normal form is unique.

For a language with both function symbols and identity, each formula is equivalent to one which contains no nested function symbols (shown by lemma 3.17), so the only constituents we need are those of functional depth 1. Equality may be taken into account either by considering the identity predicate as an ordinary binary predicate or by the method using the exclusive interpretation of quantifiers as before, but now starting with the definition just given for constituents of functional depth 1. Rantala [1987] (p. 68) mentions this approach in the context of considering the identity predicate as an ordinary binary predicate.

This chapter has shown the existence of distributive normal form for all first-order languages. In the following chapter I consider mainly the simplest case, that of languages without function symbols or identity.

Chapter 4

Some properties of distributive normal forms

This chapter examines the size of distributive normal forms, finds a lower bound for the fraction of inconsistent constituents, and shows how the existence of an algorithm for determining whether or not an arbitrary constituent is consistent or even just determining how many consistent (or inconsistent) constituents of a given form there are would imply that first-order logic was decidable. Then some sufficient, though not necessary, decidable conditions for the inconsistency of a constituent are considered, and it is shown how they can be used to prove the completeness of first-order logic (by means of the *completeness theorem of the theory of distributive normal forms*). Only languages without function symbols or identity (but possibly containing constant symbols) are considered.

4.1 The size of distributive normal forms

I will only consider the simplest possible case, that of a language which has only one binary predicate symbol and no function or constant symbols. This gives a lower bound on the size of distributive normal forms. If we use the algorithm described in section 4.3, which is based on theorem 3.5, for converting formulas to distributive normal form, then the formulas of some given depth with some set of free individual terms will contain on average about half of the constituents of that depth with those free individual terms. Even if some of the inconsistent constituents found by this algorithm are omitted, we have no way of omitting them all (as is shown in section 4.3). constit.s And in the process of converting a formula to distributive normal form it is necessary to consider those inconsistent constituents which can later be found to be inconsistent and omitted. The total number of constituents of some depth with some set of free individual terms gives an indication of the size of the distributive normal forms of the formulas of that depth with those free individual terms. Here size is measured just as the number of constituents in a distributive normal form and ignores the size of these constituents, although with increasing depth the constituents do become bigger and not just more numerous.

I consider the case of constant-free sentences first. For each depth d, we ask how many depth-d constituents there are. To find this we need to find how many existential conjunctions of each depth and level there are.

The number of basic conjunctions generated by a set of n formulas is 2^n. This is all we use, together with the definitions of the formulas involved, to find the numbers of formulas of various different forms in this section. Let a_k be the number of atomic formulas of level k. Then the number of primitive conjunctions of level k is 2^{a_k}. If the number of existential conjunctions of depth d and level k is $e_{d,k}$, then

$$e_{1,k} = 2^{(2^{a_{k+1}})}$$

since there are $2^{a_{k+1}}$ primitive conjunctions of level $k+1$, and for greater depths

$$e_{d,k} = 2^{(2^{a_{k+1}}.e_{d-1,k+1})}$$

since there are $2^{a_{k+1}}.e_{d-1,k+1}$ formulas of the form $({}^{k+1}\beta_i \wedge {}^{k+1}\epsilon_j^{d-1})$.

So $e_{2,k} = 2^{2^{(a_{k+1}+2^{a_{k+2}})}}$ and $e_{3,k} = 2^{2^{(a_{k+1}+2^{(a_{k+2}+2^{a_{k+3}})})}}$.

For a language with only one binary predicate, $a_1 = 1$, $a_2 = 3$, $a_3 = 5$. And since constituents are existential conjunctions of level 0,

- the number of depth-1 constituents is 4
- the number of depth-2 constituents is 2^{512}
- the number of depth-3 constituents is $2^{(2^{(2^{35}+1)})}$.

For some indication of how big these numbers are, consider the following. An upper bound for the number of atoms in the universe may be estimated given an estimate for the total mass of the universe by dividing it by the mass of a hydrogen atom. Using the values in the table on p. 30 of Padmanabhan [1993] gives an estimate of 10^{80} for an upper bound on the number of atoms. Thus the number of depth-2 constituents is a lot bigger than the estimated number of atoms in the universe, and with increasing depth the numbers of constituents grow very rapidly.

There are two extremely compact ways of representing distributive normal forms. We can either number all the constituents (in a way that we can computably translate between the number and the constituent) and then indicate the presence or absence of each constituent by the value of a single bit in an appropriate location. Or we can number the distributive normal forms themselves and represent the whole formula by a single number. In both cases we need the same number of bits as there are constituents.

So already for the case of constant-free sentences most of the distributive normal forms of depth 2 or more are bigger than we can represent in practice. And the distributive normal forms with free individual terms are even bigger. In this case we ask, for each depth d and each set of individual terms $z_1, \ldots, z_k$, how many constituents of the form $\delta^d(z_1, \ldots, z_k)$ there are.

For any set of k individual terms $\{z_1, \ldots, z_k\}$, let a_k be the number of atomic formulas over $\{z_1, \ldots, z_k\}$ which contain z_k. a_k is well-defined since it depends only on the number of individual terms and not what the particular individual terms are. Then the number of attributive constituents of depth 0 over any fixed set of k individual terms is 2^{a_k}. If the number of attributive constituents of depth d over any fixed set of k individual terms is $t_{d,k}$, then for $d > 0$,

$$t_{d,k} = 2^{(a_k + t_{d-1,k+1})}.$$

So

$$t_{1,k} = 2^{(a_k + 2^{a_{k+1}})}$$

$$t_{2,k} = 2^{(a_k + 2^{(a_{k+1} + 2^{a_{k+2}})})}$$

$$t_{3,k} = 2^{(a_k + 2^{(a_{k+1} + 2^{(a_{k+2} + 2^{a_{k+3}})})})}.$$

If the number of constituents of depth d over any fixed set of k individual terms is $c_{d,k}$, then $c_{0,k} = 2^{a_1 + \cdots + a_k}$ and for greater depths

$$c_{d,k} = 2^{(a_1 + \ldots + a_k + t_{d-1,k+1})}.$$

So

$$c_{1,k} = 2^{(a_1 + \ldots + a_k + 2^{a_{k+1}})}$$

$$c_{2,k} = 2^{(a_1 + \ldots + a_k + 2^{(a_{k+1} + 2^{a_{k+2}})})}$$

$$c_{3,k} = 2^{(a_1 + \ldots + a_k + 2^{(a_{k+1} + 2^{(a_{k+2} + 2^{a_{k+3}})})})}.$$

In general

$$c_{d,k} = 2^{(a_1 + \ldots + a_k + 2^{(a_{k+1} + 2^{(\cdots^{(a_{k+d-1} + 2^{a_{k+d}})} \cdots)})})}.$$

For a language with only one binary predicate, $a_i = 2i - 1$ for each i. So the number of depth-d constituents over any fixed set of k individual terms, for a few of the smallest k's and d's, is given by:

	$k=0$	$k=1$	$k=2$	$k=3$
$d=0$	—	2	8	32
$d=1$	4	512	2^{36}	2^{137}
$d=2$	2^{512}	$2^{(1+2^{35})}$	$2^{(4+2^{133})}$	$2^{(9+2^{519})}$
$d=3$	$2^{2^{(1+2^{35})}}$	$2^{(1+2^{(3+2^{133})})}$	$2^{(4+2^{(5+2^{519})})}$	$2^{(9+2^{(7+2^{2057})})}$

Thus for depths of 2 or more, most formulas can not be represented literally in distributive normal form. Given the vast number of constituents, it would be interesting to find some indication of how many of them are inconsistent. We will see (in theorem 4.19) that there is no algorithm which can find, for each depth and set of individual terms how many constituents are consistent / inconsistent. But we can show that most constituents are inconsistent (see section 4.4). We might also be interested in certain subsets of formulas which *can* be literally represented in distributive normal form. For example, we could consider the distributive normal forms of depth 2 with no free individual terms that contain only one constituent. And the constituents of depth 2 could be divided into sets according to how many attributive constituents occur positively in them. Those with only one attributive constituent would be the easiest to represent. It might be interesting to study some of these smaller sets of constituents. (For example, on p. 86–88 a particular set of constituents is classified according to whether they are consistent or inconsistent.) But, it seems that for most sentences that express some natural mathematical property, their distributive normal forms are not the small ones.

4.2 Analysis of implication

We already have, by theorem 3.8, a condition in terms of distributive normal forms for implication between formulas. We could apply this theorem directly by expanding the less deep formula to the greater depth. However, distributive normal forms are very big, and expanding a formula to a greater depth makes it much bigger still. In this section I consider a way of comparing distributive normal forms to determine when one formula implies another, which does not require the formulas to be expanded to the greater depth. It is based on a natural tree-structure of certain formulas which are part of the distributive normal forms. And while it might not be such a direct application of theorem 3.8, it does still make use of this theorem. Also, the representation of constituents as trees or sets of trees that we use here provides a convenient way of stating certain necessary conditions for consistency of a constituent. (*Trivial inconsistency* is formulated in terms of this tree-structure in section 4.5.)

automated constituent seems that theoretical sets of certain definition of

I first give some definitions. As for logical terminology, many of the concepts associated with trees are referred to in the literature by various different names. The trees I use are defined, for example, in Barwise [1977] and Smullyan [1968]. They are a special case of trees as defined in graph theory (as connected acyclic graphs) that is commonly used in computer science (and described informally in Harel [1992], Cohen [1991], and Aho and Ullman [1992]). I simply define the concepts I will use without mentioning the alternative terminology.

A *tree* is a triple $\mathcal{T} = \langle S, l, R \rangle$ where

- S is a set, the elements of which are called *nodes*
- l is a function, $l : S \to \mathbb{N}$ where $\mathbb{N}$ is the set of natural numbers, and $l(x)$ is called the *level* of x
- R is a relation, $R \subseteq S \times S$, where if $\langle x, y \rangle \in R$ then x is called a *predecessor* of y and y is called a *successor* of x

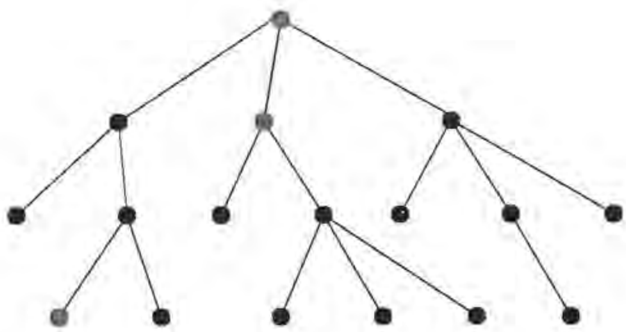
such that

- there is a unique node of level 1, and it is called the *root* of $\mathcal{T}$
- every node other than the root has a unique predecessor
- if $\langle x, y \rangle \in R$ then $l(y) = l(x) + 1$.

A *path* in a tree is a sequence of nodes where the first term is the root and for each term x, the next term in the sequence, if there is one, is a successor of x. A node with no successors is called a *leaf*. A *maximal path* is a path whose last term is a leaf, or which is infinite. If a node x occurs before a node y in any path then x is said to be *above* y in the tree and y is *below* x. If for each node $y \in Y$, y is below x, then Y is said to be *below* x. For any node x of a tree $\mathcal{T}$, the *subtree generated by* x is the tree whose root is x and where the successors of every node are the same as in $\mathcal{T}$. The subtrees generated by the nodes of $\mathcal{T}$ are called the *principal subtrees* of $\mathcal{T}$. A tree may be specified by giving its root and the subtrees generated by each of its successors. For a node x, each subtree generated by a successor of x is called a *branch* of x. A tree is *finitely branching* if each node has only finitely many branches. A tree is *finite* if it has a finite number of nodes. Two nodes with the same predecessor are called *siblings*. For each node x that has successors, the set of all successors of x is called a *sibling-set*. A sibling-set is said to be of the same *level* as its elements.

A set of trees is called a *forest*.

Trees are often represented as diagrams with the nodes as dots, the root at the top, and the successors of each node below it and joined to it with a line. A simple example is:



A *labelled tree* is a pair $\langle \mathcal{T}, f \rangle$ where $\mathcal{T}$ is some tree $\langle S, l, R \rangle$ and f is a function $f : S \to Q$ where Q is a set whose elements are called *labels*, in particular $f(x)$ is called the *label* of x.

An *isomorphism* between two trees $\langle S_1, l_1, R_1 \rangle$ and $\langle S_2, l_2, R_2 \rangle$ is a bijection $f : S_1 \to S_2$ such that for all $x, y \in S_1$, $\langle x, y \rangle \in R_1$ iff $\langle f(x), f(y) \rangle \in R_2$. If there is an isomorphism between two trees they are said to be *isomorphic*.

If $\langle \langle S_1, l_1, R_1 \rangle, g_1 \rangle$ and $\langle \langle S_2, l_2, R_2 \rangle, g_2 \rangle$ are labelled trees where both have the same set of labels Q, then a function $f : S_1 \to S_2$ is said to *preserve labels* if for all $x \in S_1$, $g_2(f(x)) = g_1(x)$.

A node of a labelled tree is said to have *duplicate branches* if it has two or more branches that have a label-preserving isomorphism between them. For any finite labelled tree $\langle \mathcal{T}, f \rangle$, we can find a labelled tree for which no node has duplicate branches by the following method: Start at the greatest level n at which there are some nodes whose only branches are leaves. For each node at level n, for each set of duplicate branches that it has, delete all but one of those branches. Decrease the level by 1 and repeat the process. Continue until the root is reached. The tree obtained in this manner is called $\langle \mathcal{T}, f \rangle$ *with duplicates removed*.

So far, we have only used existential conjunctions for the case of constant-free sentences since then the constituents are existential conjunctions. But we can also define existential conjunctions with different sets of free individual terms, and then both attributive constituents and constituents will have the form of a conjunction of some depth-0 formula and an existential conjunction, the difference between attributive constituents and constituents being a difference in what the depth-0 formula is:

For $d \geq 1$, an *existential conjunction of depth d with free individual terms* $z_1, \ldots, z_k$ is denoted $\epsilon_i^d(z_1, \ldots, z_k)$ and defined by

$$\epsilon_i^d(z_1, \ldots, z_k) = (\pm)(\exists x)\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_p^{d-1}(z_1, \ldots, z_k, x)$$

where x is the first variable not in $z_1, \ldots, z_k$ and $\gamma_1^{d-1}(z_1, \ldots, z_k, x), \ldots, \gamma_p^{d-1}(z_1, \ldots, z_k, x)$ are all the attributive constituents of depth $d-1$ with free individual terms $z_1, \ldots, z_k, x$.

Then constituents and attributive constituents of depths greater than 0 have the following forms:

$$
\begin{aligned}
\gamma_i^d(z_1,\ldots,z_k) &= \gamma_p^0(z_1,\ldots,z_k) \wedge \epsilon_r^d(z_1,\ldots,z_k)\\
\delta_j^d(z_1,\ldots,z_k) &= \delta_q^0(z_1,\ldots,z_k) \wedge \epsilon_s^d(z_1,\ldots,z_k).
\end{aligned}
$$

The above formulas may be naturally represented as labelled trees. In describing how this is done, I will refer to the nodes by their labels rather than their names. This is somewhat imprecise, but I think it facilitates communication. It must be understood however that different nodes in a tree can have the same label. All the trees which I use for representing constituents and attributive constituents are finite and have formulas for their labels. We first represent an existential conjunction $\epsilon^d(z_1,\ldots,z_k)$ by the set

$$\{X \mid (\exists x)X \text{ occurs unnegated in } \epsilon^d(z_1,\ldots,z_k)\}.$$

Then the above attributive constituent is represented by the tree with $\gamma_p^0(z_1,\ldots,z_k)$ as the root and the elements of $\epsilon_r^d(z_1,\ldots,z_k)$ as its branches, and the above constituent is represented by the tree with $\delta_q^0(z_1,\ldots,z_k)$ as the root and the elements of $\epsilon_s^d(z_1,\ldots,z_k)$ as its branches. Each branch is again a tree of this form. There are two limit cases to be considered, the first being that of formulas of depth 0. They have as their trees just a root, and so form the leaves of the trees and provide a basis for this recursive definition. The other case is when the formulas have no free individual terms, in which case there is no root, so the formula is represented by a set of trees instead of a tree. This is when the constituent or attributive constituent is an existential conjunction. This can only occur at the outermost level of the recursion, that is when we start with an existential conjunction, and never with an attributive constituent which is in another constituent (or attributive constituent). So each existential conjunction is represented by a forest and each constituent or attributive constituent with free individual terms by a tree.

To give an example of this tree representation of a constituent, we consider a language with only one binary predicate symbol. In writing the constituents and attributive constituents for this example, all conjuncts which are not negations are shown explicitly. The formula

$$
\begin{aligned}
\delta^2(x_1) \;=\;& \delta_1^0(x_1) \wedge \\
& \neg(\exists x_2)\gamma_1^1(x_1,x_2) \wedge \ldots \wedge \neg(\exists x_2)\gamma_{2^{31}-2}^1(x_1,x_2) \wedge (\exists x_2)\gamma_{2^{31}-1}^1(x_1,x_2)\\
& (\exists x_2)\gamma_{2^{31}}^1(x_1,x_2) \wedge \neg(\exists x_2)\gamma_{2^{31}+1}^1(x_1,x_2) \wedge \ldots \wedge \neg(\exists x_2)\gamma_{2^{35}}^1(x_1,x_2)
\end{aligned}
$$

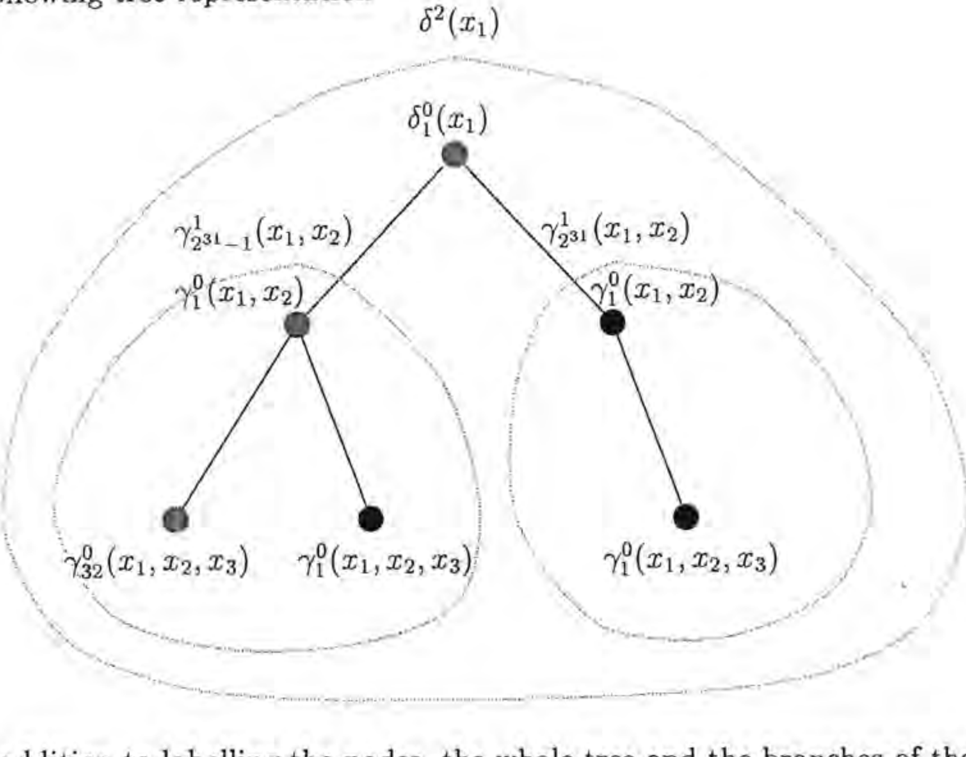
where

$$
\begin{aligned}
\gamma_{2^{31}-1}^1(x_1,x_2) \;=\;& \gamma_1^0(x_1,x_2)\\
& (\exists x_3)\gamma_1^0(x_1,x_2,x_3) \wedge \neg(\exists x_3)\gamma_2^0(x_1,x_2,x_3) \wedge \ldots \wedge \neg(\exists x_3)\gamma_{31}^0(x_1,x_2,x_3) \wedge\\
& (\exists x_3)\gamma_{32}^0(x_1,x_2,x_3)
\end{aligned}
$$

and

$$
\begin{aligned}
\gamma_{2^{31}}^1(x_1,x_2) \;=\;& \gamma_1^0(x_1,x_2)\\
& (\exists x_3)\gamma_1^0(x_1,x_2,x_3) \wedge \neg(\exists x_3)\gamma_2^0(x_1,x_2,x_3) \wedge \ldots \wedge \neg(\exists x_3)\gamma_{32}^0(x_1,x_2,x_3)
\end{aligned}
$$

has the following tree representation



where in addition to labelling the nodes, the whole tree and the branches of the root have been circled and labelled.

Together with the previously introduced representation of a formula as a set of constituents, this now gives us a representation for each constant-free sentence as a set of forests, and for each formula with free individual terms as a set of trees. (I use "set of trees" instead of "forest" to describe the representation of a formula with free individual terms because it more closely follows the difference in structure between the cases for formulas with and without free individual terms.) I will call this representation for formulas, constituents, and attributive constituents the *sets-and-trees representation*, and the *sets-and-trees representation*$_0$ is the one using the set representation$_0$ (defined on p. 42). I use the terminology of the sets-and-trees representation of constituents and attributive constituents also to refer to the syntactic parts of these formulas. For example, in a constituent

$$\begin{aligned}\delta^d(z_1, \ldots, z_k) &= \delta^0(z_1, \ldots, z_k) \wedge \\ &\quad (\pm)(\exists x)\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_r^{d-1}(z_1, \ldots, z_k, x),\end{aligned}$$

$\delta^0(z_1, \ldots, z_k)$ is called the root, and each $\gamma_i^{d-1}(z_1, \ldots, z_k, x)$ for which $(\exists x)\gamma_i^{d-1}(z_1, \ldots, z_k, x)$ is a conjunct of $\delta^d(z_1, \ldots, z_k)$ is called a branch.

For $e > d$, if we can describe those depth-e constituents which occur in the expansion of a depth-d constituent to depth e in terms of the tree-structures of the formulas involved, then we get a condition for implication between constituents, which when put together with theorem 3.8 gives a condition for implication between formulas. And there is a condition for when a depth-e constituent is in the expansion of a depth-d constituent to depth e for which it is only necessary to consider the top d levels of the trees involved.

Since the representations for formulas with and without free individual terms are different, we consider these two cases separately. For both we use the following lemma.

Lemma 4.1 *For each model, for each set of individual terms $z_1, \ldots, z_k$, for each valuation, exactly one existential conjunction of the form $\epsilon_i^d(z_1, \ldots, z_k)$ is true for each d.*

PROOF For each depth d and each set of individual terms $z_1, \ldots, z_k$, the existential conjunctions of depth d with free individual terms $z_1, \ldots, z_k$ are the basic conjunctions generated by some set of formulas, so exactly one such existential conjunction is true for any given model and valuation. □

Case for constant-free sentences

As a special case of theorem 3.8, we have (again using the same symbol for a formula and its set representation): For any constant-free sentences F_1 and F_2, one of depth d and the other of depth e with $e \geq d$,

1. for any set representations at depth e of the sentences, if $F_1 \subseteq F_2$ then $F_1 \Longrightarrow F_2$
2. for the set representation$_0$ of F_1 and any set representation of F_2, both at depth e, $F_1 \subseteq F_2$ iff $F_1 \Longrightarrow F_2$.

Now, the specific question we are asking in this case is: If $e > d$, given some $\epsilon_i^d()$, which constituents $\epsilon_j^e()$ are in the expansion of $\epsilon_i^d()$ to depth e?

Using the sets-and-trees representation, each sentence constituent is represented by a forest. With this representation we can answer the question for the case where $d = 1$ and e is any greater depth by (a special case of) the following lemma.

Lemma 4.2 *If $e > 1$ then an existential conjunction $\epsilon_i^1(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the $\epsilon_j^e(z_1, \ldots, z_k)$'s such that the elements of $\epsilon_i^1(z_1, \ldots, z_k)$ are precisely the roots of the elements of $\epsilon_j^e(z_1, \ldots, z_k)$.*

PROOF The existential conjunction $\epsilon_i^1(z_1, \ldots, z_k)$ is of the form

$$(\pm)(\exists x)\beta_1(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\beta_q(z_1, \ldots, z_k, x)$$

and the existential conjunctions of depth e over $z_1, \ldots, z_k$ are of the form

$$(\pm)(\exists x)(\beta_1(z_1, \ldots, z_k, x) \wedge \epsilon_1^{e-1}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_1(z_1, \ldots, z_k, x) \wedge \epsilon_p^{e-1}(z_1, \ldots, z_k, x))$$

$$\wedge \ldots \wedge$$

$$(\pm)(\exists x)(\beta_q(z_1, \ldots, z_k, x) \wedge \epsilon_1^{e-1}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_q(z_1, \ldots, z_k, x) \wedge \epsilon_p^{e-1}(z_1, \ldots, z_k, x)).$$

For each $a \in \{1, \ldots, q\}$, $\beta_a(z_1, \ldots, z_k, x) \iff (\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_1^{e-1}(z_1, \ldots, z_k, x)) \vee \ldots \vee (\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_p^{e-1}(z_1, \ldots, z_k, x))$, by lemma 4.1. So, by distributing the existential quantifier over disjunction,

$$\begin{aligned}(\exists x)\beta_a(z_1, \ldots, z_k, x) \iff & (\exists x)(\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_1^{e-1}(z_1, \ldots, z_k, x)) \vee \ldots \vee \\ & (\exists x)(\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_p^{e-1}(z_1, \ldots, z_k, x))\end{aligned}$$

and

$$\begin{aligned}\neg(\exists x)\beta_b(z_1, \ldots, z_k, x) \iff & \neg(\exists x)(\beta_b(z_1, \ldots, z_k, x) \wedge \epsilon_1^{e-1}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge \\ & \neg(\exists x)(\beta_b(z_1, \ldots, z_k, x) \wedge \epsilon_p^{e-1}(z_1, \ldots, z_k, x)).\end{aligned} \tag{4.1}$$

So $(\exists x)\beta_a(z_1, \ldots, z_k, x)$ is equivalent to the disjunction of all the formulas of the form

$$(\pm)(\exists x)(\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_1^{e-1}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_p^{e-1}(z_1, \ldots, z_k, x)) \tag{4.2}$$

for which at least one conjunct is not a negation. Let

$$\begin{aligned} A &= \{a \in \{1, \ldots, q\} \mid (\exists x)\beta_a(z_1, \ldots, z_k, x) \text{ is a conjunct of } \epsilon_i^1(z_1, \ldots, z_k)\}, \\ B &= \{b \in \{1, \ldots, q\} \mid \neg(\exists x)\beta_b(z_1, \ldots, z_k, x) \text{ is a conjunct of } \epsilon_i^1(z_1, \ldots, z_k)\}.\end{aligned}$$

Then for each $a \in A$, substitute (the disjunction of all the formulas (4.2) of which at least one conjunct is not a negation) for $(\exists x)\beta_a(z_1, \ldots, z_k, x)$ in $\epsilon_i^1(z_1, \ldots, z_k)$. And for each $b \in B$, substitute the formula on the right-hand side of (4.1) for $\neg(\exists x)\beta_b(z_1, \ldots, z_k, x)$ in $\epsilon_i^1(z_1, \ldots, z_k)$. In the resulting formula, distributing conjunction over disjunction results in the formula which is the disjunction of all the formulas of the form $\epsilon_j^e(z_1, \ldots, z_k)$ such that for each $a \in A$ there is at least one $n \in \{1, \ldots, p\}$ such that $(\exists x)(\beta_a(z_1, \ldots, z_k, x) \wedge \epsilon_n^{e-1}(z_1, \ldots, z_k, x))$ is a conjunct of $\epsilon_j^e(z_1, \ldots, z_k)$ and for each $b \in B$, for all $n \in \{1, \ldots, p\}$, $\neg(\exists x)(\beta_b(z_1, \ldots, z_k, x) \wedge \epsilon_n^{e-1}(z_1, \ldots, z_k, x))$ is a conjunct of $\epsilon_j^e(z_1, \ldots, z_k)$. Restating this in terms of the sets-and-trees representation of the existential conjunctions, $\epsilon_i^1(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the formulas $\epsilon_j^e(z_1, \ldots, z_k)$ such that each element of $\epsilon_i^1(z_1, \ldots, z_k)$ is the root of some element of $\epsilon_j^e(z_1, \ldots, z_k)$ and all roots of elements of $\epsilon_j^e(z_1, \ldots, z_k)$ are elements of $\epsilon_i^1(z_1, \ldots, z_k)$. □

As a special case of the above lemma, we get:

If $e > 1$, then the consistent constituents that are in the expansion of $\epsilon_i^1()$ to depth e are the consistent constituents $\epsilon_j^e()$ such that the elements of $\epsilon_i^1()$ are precisely the roots of the elements of $\epsilon_j^e()$.

Let F_1 and F_2 be constant-free sentences, F_1 of depth 1 and F_2 of any greater depth e, where F_1 is represented by $\{\epsilon_i^1()\}_{i\in I}$ and F_2 by $\{\epsilon_j^e()\}_{j\in J}$. If $\{\epsilon_i^1()\}_{i\in I}$ is the sets-and-trees representation$_0$ of F_1, then $F_1 \Longrightarrow F_2$ iff the expansions of all of the $\epsilon_i^1() \in F_1$ to depth e are contained in $\{\epsilon_j^e()\}_{j\in J}$ iff for each $\epsilon_i^1() \in F_1$, all constituents of depth e with elements the set of whose roots is $\epsilon_i^1()$ are in $\{\epsilon_j^e()\}_{j\in J}$. So there is just as much to check as if all the expanding is actually done. However, if $\{\epsilon_j^e()\}_{j\in J}$ is the sets-and-trees representation$_0$ of F_2, then $F_2 \Longrightarrow F_1$ iff each $\epsilon_j^e() \in F_2$ is in the expansion to depth e of some $\epsilon_i^1() \in F_1$, and $\epsilon_j^e()$ is in the expansion of $\epsilon_i^1()$ iff the roots of the elements of $\epsilon_j^e()$ are the elements of $\epsilon_i^1()$. So this can be checked with considerably less work than expanding F_1 to depth e.

To check if a formula implies one of a greater depth we may as well just expand it to the greater depth. So from now on I will not consider this case in looking for a more efficient way of checking implication between formulas. We have just seen that for constant-free sentences if we have the set representation$_0$ of the deeper sentence available there is a much simpler way of checking if it implies a sentence of depth 1 than by expanding to the greater depth. But since we generally can't expect to know which constituents are consistent, I now state the corresponding partial result for any sets-and-trees representations of the sentences:

For constant-free sentences F_1 of depth 1 and F_2 of any greater depth e, where F_1 is represented by $\{\epsilon_i^1()\}_{i\in I}$ and F_2 by $\{\epsilon_j^e()\}_{j\in J}$, if for each $\epsilon_j^e() \in F_2$ there is some $\epsilon_i^1() \in F_1$ such that the set of roots of the elements of $\epsilon_j^e()$ is $\epsilon_i^1()$, then $F_2 \Longrightarrow F_1$.

We now consider depths $d > 1$.

For a labelled tree $\mathcal{T} = \langle\langle N, l, R\rangle, f\rangle$, for $d \geq 1$, let N' be $\{x \in N \mid l(x) \leq d\}$, l' be l restricted to N', R' be R restricted to N', and f' be f restricted to N'. Then the labelled tree $\langle\langle N', l', R'\rangle, f'\rangle$ with duplicates removed is called the *top-d tree* of $\mathcal{T}$. $\mathcal{T}$ Thus, we can say less formally that the top-d tree of $\mathcal{T}$ is the labelled tree which consists of the top d levels of $\mathcal{T}$ (if they exist), with duplicates removed. The *top-d forest* of a forest $\mathcal{F}$ is the set whose elements are the top-d trees of the elements of $\mathcal{F}$.

Lemma 4.3 *For any depths d and e with $d < e$, an existential conjunction $\epsilon_i^d(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the $\epsilon_j^e(z_1, \ldots, z_k)$'s such that $\epsilon_i^d(z_1, \ldots, z_k)$ is the top-d forest of $\epsilon_j^e(z_1, \ldots, z_k)$.*

PROOF We first show this for $e = d+1$. We use induction on the depth d. The case for $d = 1$ is shown by lemma 4.2. Assume that for some depth $d-1$, for every set of individual terms $z_1, \ldots, z_k$, each $\epsilon_i^{d-1}(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the $\epsilon_j^d(z_1, \ldots, z_k)$'s such that the set of top-$(d-1)$ trees of the elements of $\epsilon_j^d(z_1, \ldots, z_k)$ is $\epsilon_i^{d-1}(z_1, \ldots, z_k)$. Then an existential conjunction $\epsilon_i^d(z_1, \ldots, z_k)$ has the form

$$(\pm)(\exists x)(\beta_1(z_1, \ldots, z_k, x) \wedge \epsilon_1^{d-1}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_1(z_1, \ldots, z_k, x) \wedge \epsilon_p^{d-1}(z_1, \ldots, z_k, x))$$
$$\wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_q(z_1, \ldots, z_k, x) \wedge \epsilon_1^{d-1}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_q(z_1, \ldots, z_k, x) \wedge \epsilon_p^{d-1}(z_1, \ldots, z_k, x))$$

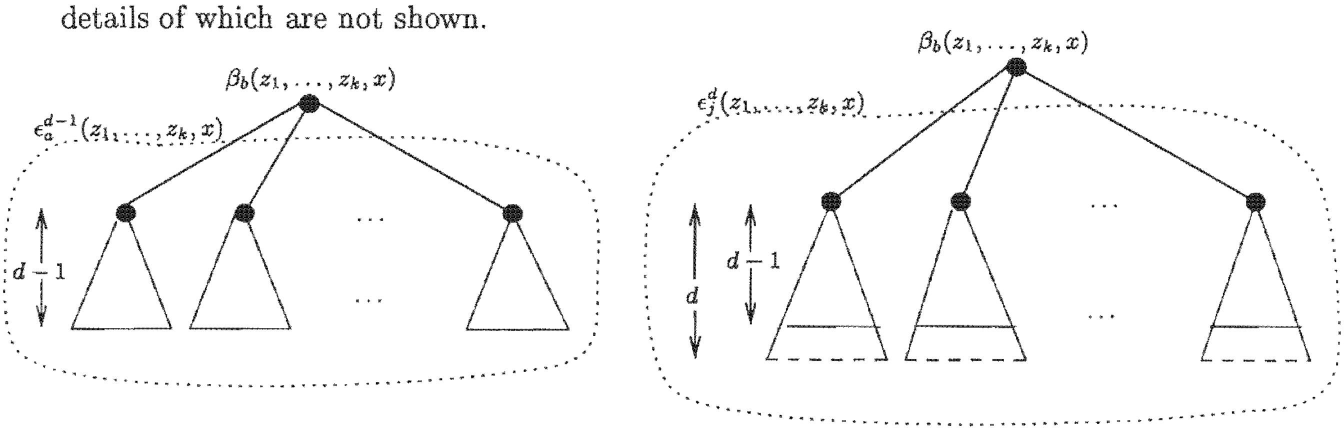
and an existential conjunction $\epsilon_j^{d+1}(z_1, \ldots, z_k)$ has the form

$$(\pm)(\exists x)(\beta_1(z_1, \ldots, z_k, x) \wedge \epsilon_1^{d}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_1(z_1, \ldots, z_k, x) \wedge \epsilon_r^{d}(z_1, \ldots, z_k, x))$$
$$\wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_q(z_1, \ldots, z_k, x) \wedge \epsilon_1^{d}(z_1, \ldots, z_k, x)) \wedge \ldots \wedge$$
$$(\pm)(\exists x)(\beta_q(z_1, \ldots, z_k, x) \wedge \epsilon_r^{d}(z_1, \ldots, z_k, x)).$$

By the induction hypothesis, each $\epsilon_a^{d-1}(z_1, \ldots, z_k, x)$ is equivalent to the disjunction $\bigvee_{j \in J_a} \epsilon_j^d(z_1, \ldots, z_k, x)$ of all the $\epsilon_j^d(z_1, \ldots, z_k, x)$'s such that the top-$(d-1)$ forest of $\epsilon_j^d(z_1, \ldots, z_k, x)$ is $\epsilon_a^{d-1}(z_1, \ldots, z_k, x)$. So

$$\beta_b(z_1, \ldots, z_k, x) \wedge \epsilon_a^{d-1}(z_1, \ldots, z_k, x) \iff \bigvee_{j \in J_a} (\beta_b(z_1, \ldots, z_k, x) \wedge \epsilon_j^d(z_1, \ldots, z_k, x)).$$

And these formulas have the following tree structures, where triangles represent trees the details of which are not shown.



If $\epsilon_a^{d-1}(z_1,\ldots,z_k,x)$ is the top-$(d-1)$ forest of $\epsilon_j^d(z_1,\ldots,z_k,x)$ then $\beta_b(z_1,\ldots,z_k,x) \wedge \epsilon_a^{d-1}(z_1,\ldots,z_k,x)$ is the top-d tree of $\beta_b(z_1,\ldots,z_k,x)\wedge\epsilon_j^d(z_1,\ldots,z_k,x)$, and if $\beta_b(z_1,\ldots,z_k,x)\wedge \epsilon_a^{d-1}(z_1,\ldots,z_k,x)$ is the top-d tree of $\beta_b(z_1,\ldots,z_k,x)\wedge\epsilon_j^d(z_1,\ldots,z_k,x)$ then $\epsilon_a^{d-1}(z_1,\ldots,z_k,x)$ is the top-$(d-1)$ forest of $\epsilon_j^d(z_1,\ldots,z_k,x)$. Now,

$$(\exists x)(\beta_b(z_1,\ldots,z_k,x) \wedge \epsilon_a^{d-1}(z_1,\ldots,z_k,x)) \iff \bigvee_{j\in J_a} (\exists x)(\beta_b(z_1,\ldots,z_k,x) \wedge \epsilon_j^d(z_1,\ldots,z_k,x))$$

and

$$\neg(\exists x)(\beta_b(z_1,\ldots,z_k,x)\wedge\epsilon_a^{d-1}(z_1,\ldots,z_k,x)) \iff \bigwedge_{j\in J_a} \neg(\exists x)(\beta_b(z_1,\ldots,z_k,x)\wedge\epsilon_j^d(z_1,\ldots,z_k,x)).$$

For each conjunct $(\exists x)(\beta_b(z_1,\ldots,z_k,x)\wedge\epsilon_a^{d-1}(z_1,\ldots,z_k,x))$ of $\epsilon_i^d(z_1,\ldots,z_k)$, substitute the above equivalent formula, and for each conjunct $\neg(\exists x)(\beta_b(z_1,\ldots,z_k,x)\wedge\epsilon_a^{d-1}(z_1,\ldots,z_k,x))$ of $\epsilon_i^d(z_1,\ldots,z_k)$, substitute the above equivalent formula. In the resulting formula, distributing conjunction over disjunction results in the formula which is the disjunction of all the formulas of the form $\epsilon_j^{d+1}(z_1,\ldots,z_k)$ such that each element of $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d tree of some element of $\epsilon_j^{d+1}(z_1,\ldots,z_k)$ and each top-d tree of an element of $\epsilon_j^{d+1}(z_1,\ldots,z_k)$ is an element of $\epsilon_i^d(z_1,\ldots,z_k)$. In other words, $\epsilon_i^d(z_1,\ldots,z_k)$ is equivalent to the disjunction of all the formulas $\epsilon_j^{d+1}(z_1,\ldots,z_k)$ such that $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of $\epsilon_j^{d+1}(z_1,\ldots,z_k)$.

We now show the result for any $e > d$. We use induction on $e-d$. If $e-d=1$ the result is shown above. If the result holds for $e-d=n$ then each existential conjunction $\epsilon_i^d(z_1,\ldots,z_k)$ is equivalent to the disjunction of all the $\epsilon_j^{d+n}(z_1,\ldots,z_k)$'s such that $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of $\epsilon_j^{d+n}(z_1,\ldots,z_k)$. By the case shown above, each $\epsilon_j^{d+n}(z_1,\ldots,z_k)$ is equivalent to the disjunction of all the $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$'s such that $\epsilon_j^{d+n}(z_1,\ldots,z_k)$ is the top-$(d+n)$ forest of $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$. If $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of $\epsilon_j^{d+n}(z_1,\ldots,z_k)$ and $\epsilon_j^{d+n}(z_1,\ldots,z_k)$ is the top-$(d+n)$ forest of $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$ then $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$. And if $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$ then $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of the top-$(d+n)$ forest of $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$. Thus $\epsilon_i^d(z_1,\ldots,z_k)$ is equivalent to the disjunction of all the $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$'s such that $\epsilon_i^d(z_1,\ldots,z_k)$ is the top-d forest of $\epsilon_l^{d+n+1}(z_1,\ldots,z_k)$. So the result also holds for $e-d=n+1$. □

By this lemma, any existential conjunction is equivalent to some disjunction of existential conjunctions of any greater depth. Any such disjunction will be called an *expansion* of the existential conjunction to the greater depth. By lemma 4.1, the set of consistent existential conjunctions occurring in an expansion of some existential conjunction to a greater depth is unique. This terminology causes some ambiguity since an expansion of any formula has already been defined (on p. 42) as a disjunction of constituents to which it is equivalent. However, I haven't come across any need to expand existential conjunctions in terms of constituents and by an *expansion* of an existential conjunction, I will mean an expansion in terms of existential conjunctions.

Restating the above lemma and taking into account that inconsistent existential conjunctions may be or not be in an expansion, we get:

If $d < e$, a consistent existential conjunction $\epsilon_j^e(z_1 \ldots, z_k)$ is in the expansion to depth e of $\epsilon_i^d(z_1, \ldots, z_k)$ iff $\epsilon_i^d(z_1, \ldots, z_k)$ is the top-d forest of $\epsilon_j^e(z_1, \ldots, z_k)$.

As an immediate consequence of lemma 4.3, we get:

For any existential conjunction $\epsilon^d(z_1, \ldots, z_k)$ and any $c < d$, the top-c forest of $\epsilon^d(z_1, \ldots, z_k)$ is some existential conjunction $\epsilon^c(z_1, \ldots, z_k)$ and $\epsilon^d(z_1, \ldots, z_k) \Longrightarrow \epsilon^c(z_1, \ldots, z_k)$.

We now combine theorem 3.8 with the special case of lemma 4.3 for which the existential conjunctions are sentence constituents:

Theorem 4.4 *Given constant-free sentences F_1 of depth d and F_2 of any greater depth e,*

1. *For the sets-and-trees representation$_0$ of F_2 and any sets-and-trees representation of F_1, $F_2 \Longrightarrow F_1$ iff for each $\epsilon_j^e() \in F_2$, there is some $\epsilon_i^d() \in F_1$, such that $\epsilon_i^d()$ is the top-d forest of $\epsilon_j^e()$.*
2. *For any sets-and-trees representations of the sentences, If for each $\epsilon_j^e() \in F_2$, there is some $\epsilon_i^d() \in F_1$, such that $\epsilon_i^d()$ is the top-d forest of $\epsilon_j^e()$, then $F_2 \Longrightarrow F_1$.*

PROOF By theorem 3.8 and lemma 4.3. □

Case for formulas with free individual terms

In the general case, a formula can be expanded both to greater depths and to larger sets of free individual terms. I consider only the case for formulas with the same free individual terms.

With the sets-and-trees representation, a constituent with free individual terms is represented by a tree, thus a formula with free individual terms by a set of trees.

We ask the same question as before: If $e > d$, what is the expansion of a constituent of depth d to depth e? In this case, specifically:

Given some $\delta_i^d(z_1, \ldots, z_k)$, which constituents $\delta_j^e(z_1, \ldots, z_k)$ are in the expansion of $\delta_i^d(z_1, \ldots, z_k)$ to depth e?

The structure of a constituent is like that of an existential conjunction, but joined to the set of trees is a root, making the constituent into just one tree.

Lemma 4.5 *For any depths d and e with $d < e$, a constituent $\delta_i^d(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the constituents $\delta_j^e(z_1, \ldots, z_k)$ such that $\delta_i^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\delta_j^e(z_1, \ldots, z_k)$.*

PROOF For $d = 0$, each $\delta_i^0(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the formulas of the form $\delta_i^0(z_1, \ldots, z_k) \wedge \epsilon_c^e(z_1, \ldots, z_k)$ which are exactly the formulas of the form $\delta_j^e(z_1, \ldots, z_k)$ whose root is $\delta_i^0(z_1, \ldots, z_k)$. For $d > 0$, a constituent $\delta_i^d(z_1, \ldots, z_k)$ has the form $\delta_a^0(z_1, \ldots, z_k) \wedge \epsilon_b^d(z_1, \ldots, z_k)$. By lemma 4.3, $\epsilon_b^d(z_1, \ldots, z_k)$ is equivalent to the disjunction $\bigvee_{c \in C} \epsilon_c^e(z_1, \ldots, z_k)$ of all the $\epsilon_c^e(z_1, \ldots, z_k)$'s such that $\epsilon_b^d(z_1, \ldots, z_k)$ is the top-d forest of $\epsilon_c^e(z_1, \ldots, z_k)$. So

$$\begin{aligned} \delta_i^d(z_1, \ldots, z_k) &\iff \delta_a^0(z_1, \ldots, z_k) \wedge \bigvee_{c \in C} \epsilon_c^e(z_1, \ldots, z_k) \\ &\iff \bigvee_{c \in C} (\delta_a^0(z_1, \ldots, z_k) \wedge \epsilon_c^e(z_1, \ldots, z_k)) \end{aligned}$$

Now, if $\epsilon_b^d(z_1, \ldots, z_k)$ is the top-d forest of $\epsilon_c^e(z_1, \ldots, z_k)$ then $\delta_a^0(z_1, \ldots, z_k) \wedge \epsilon_b^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\delta_a^0(z_1, \ldots, z_k) \wedge \epsilon_c^e(z_1, \ldots, z_k)$. And if $\delta_a^0(z_1, \ldots, z_k) \wedge \epsilon_b^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\delta_a^0(z_1, \ldots, z_k) \wedge \epsilon_c^e(z_1, \ldots, z_k)$ then $\epsilon_b^d(z_1, \ldots, z_k)$ is the top-d forest of $\epsilon_c^e(z_1, \ldots, z_k)$. Thus $\delta_i^d(z_1, \ldots, z_k)$ is equivalent to the disjunction of all $\delta_j^e(z_1, \ldots, z_k)$'s such that $\delta_i^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\delta_j^e(z_1, \ldots, z_k)$. □

Lemma 4.6 *For any depths d and e with $d < e$, an attributive constituent $\gamma_i^d(z_1, \ldots, z_k)$ is equivalent to the disjunction of all the attributive constituents $\gamma_j^e(z_1, \ldots, z_k)$ such that $\gamma_i^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\gamma_j^e(z_1, \ldots, z_k)$.*

PROOF The proof of this lemma is just like the proof of lemma 4.5, but with all δ's replaced by γ's. □

Corollary 4.7 *For any attributive constituent $\gamma^d(z_1, \ldots, z_k)$ and any $c \leq d$, the top-c tree of $\gamma^d(z_1, \ldots, z_k)$ is some attributive constituent $\gamma^{c-1}(z_1, \ldots, z_k)$ and $\gamma^d(z_1, \ldots, z_k) \Longrightarrow \gamma^{c-1}(z_1, \ldots, z_k)$.*

PROOF This is immediate from lemma 4.6. □

Putting lemma 4.5 together with theorem 3.8, we get:

Theorem 4.8 *Given formulas F_1 of depth d and F_2 of any greater depth e, both with free variables $z_1, \ldots, z_k$,*

1. *For the sets-and-trees representation$_0$ of F_2 and any sets-and-trees representation of F_1, $F_2 \Longrightarrow F_1$ iff for each $\delta^e_j(z_1, \ldots, z_k) \in F_2$, there is some $\delta^d_i(z_1, \ldots, z_k) \in F_1$ such that $\delta^d_i(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\delta^e_j(z_1, \ldots, z_k)$.*

2. *For any sets-and-trees representations of F_1 and F_2,* *If for each $\delta^e_j(z_1, \ldots, z_k) \in F_2$, there is some $\delta^d_i(z_1, \ldots, z_k) \in F_1$ such that $\delta^d_i(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree of $\delta^e_j(z_1, \ldots, z_k)$, then $F_2 \Longrightarrow F_1$.*

PROOF By theorem 3.8 and lemma 4.5. □

4.3 Undecidability of consistency of constituents

The proof of theorem 3.5 gives us a method of converting any formula to distributive normal form. If F is a formula of depth $\leq d$ whose free individual terms are in $z_1, \ldots, z_k$, then we can find a disjunction of constituents of the form $\delta^d(z_1, \ldots, z_k)$ which is equivalent to F by the following algorithm:

First write F using just the logical connectives $\neg, \vee$ and the quantifier $\exists$, and change the bound variables so that for each $(\exists x)$ in F, x is the first unused variable in the alphabetical order $(x_1, x_2, \ldots)$ which is not in $z_1, \ldots, z_k$. If $d = 0$ then F has depth 0, so can be transformed to a disjunctive normal form with disjuncts of the form $\delta^0(z_1, \ldots, z_k)$, otherwise

- if F is atomic : write F in disjunctive normal form as $\bigvee_{i \in I} \delta^0_i(z_1, \ldots, z_k)$, then for each $i \in I$, replace $\delta^0_i(z_1, \ldots, z_k)$ by the disjunction of all the formulas of the form

$$\delta^0_i(z_1, \ldots, z_k) \wedge (\pm)(\exists x)\gamma^{d-1}_1(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma^{d-1}_p(z_1, \ldots, z_k, x)$$

  where x is the first variable not in $z_1, \ldots, z_k$ and $\gamma^{d-1}_1(z_1, \ldots, z_k, x), \ldots, \gamma^{d-1}_p(z_1, \ldots, z_k, x)$ are all the attributive constituents of that form.

- if F is $(\exists x)G$: G is a formula of depth $\leq d-1$ with free individual terms in $z_1, \ldots, z_k, x$, (and x is the first variable not in $z_1, \ldots, z_k$), so use this algorithm to convert G to a disjunction of the form $\bigvee_{i \in I} \delta^{d-1}_i(z_1, \ldots, z_k, x)$, here each $\delta^{d-1}_i(z_1, \ldots, z_k, x)$ has the form $\delta^0_{i_a}(z_1, \ldots, z_k) \wedge \gamma^{d-1}_{i_b}(z_1, \ldots, z_k, x)$, so in $(\exists x)G$, move the quantifier to get $\bigvee_{i \in I}(\delta^0_{i_a}(z_1, \ldots, z_k) \wedge (\exists x)\gamma^{d-1}_{i_b}(z_1, \ldots, z_k, x))$, and then replace each $\delta^0_{i_a}(z_1, \ldots, z_k) \wedge (\exists x)\gamma^{d-1}_{i_b}(z_1, \ldots, z_k, x)$ by the disjunction of all the formulas of the form

$$\delta^0_{i_a}(z_1, \ldots, z_k) \wedge$$
$$(\pm)(\exists x)\gamma^{d-1}_1(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\exists x)\gamma^{d-1}_{i_b}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma^{d-1}_p(z_1, \ldots, z_k, x)$$

where $\gamma_1^{d-1}(z_1, \ldots, z_k, x), \ldots, \gamma_p^{d-1}(z_1, \ldots, z_k, x)$ are all the attributive constituents of that form.

- if F is $G \vee H$: use this algorithm to convert each of G and H to a disjunction of formulas of the form $\delta^d(z_1, \ldots, z_k)$.
- if F is $\neg G$: use this algorithm to convert G to a disjunction of formulas of the form $\delta^d(z_1, \ldots, z_k)$, then F is the disjunction of all the formulas of the form $\delta^d(z_1, \ldots, z_k)$ which are not disjuncts of G.

may give depth by

Theorem 4.9 *For a language with at least one n-ary predicate symbol for some $n \geq 2$, there is no algorithm such that, for each constituent (of each depth), the algorithm can decide whether or not the constituent is consistent.*

PROOF The above algorithm converts any formula to distributive normal form. If there is an algorithm for finding the inconsistent constituents then we would have a decision procedure for first-order logic since a formula of depth d is valid iff its distributive normal form contains all the consistent depth-d constituents (by corollary 3.7). Since for languages with a predicate symbol of arity ≥ 2 the logic is not decidable (Hodel [1995], p. 208–209), there can't be such an algorithm. □

Once we have proved some further results we will be able to show that we can't generally even find an algorithm which tells us how many constituents of each form are consistent (see theorem 4.19). But we can obtain some lower bounds for the number, or fraction, of inconsistent constituents of each form, as I do in the next section.

4.4 A lower bound for the fraction of inconsistent constituents

This section gives an intuitive explanation of some conditions that are sufficient for a constituent to be inconsistent, as an introduction to the formal definition of trivial inconsistency (found in the next section). It then uses these conditions to find a lower bound for the fraction of inconsistent constituents.

Constituents of depth 0 are the formulas $\delta_i^0(z_1, \ldots, z_k)$ of the form $(\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_p(z_1, \ldots, z_k)$. That is, they are the basic conjunctions generated by the set $\{\alpha_1(z_1), \ldots, \alpha_p(z_1, \ldots, z_k)\}$. Each constituent $\delta_i^0(z_1, \ldots, z_k)$ is satisfied by some model and valuation, defined as follows: Let $\mathcal{M}$ be a model whose domain contains k elements $e_1, \ldots, e_k$

and whose interpretation I is defined by, for each formula $P(z_a, \ldots, z_b)$ in the generating set $\{\alpha_1(z_1), \ldots, \alpha_p(z_1, \ldots, z_k)\}$, if $P(z_a, \ldots, z_b)$ occurs unnegated in $\delta_i^0(z_1, \ldots, z_k)$ then let $\langle e_a, \ldots, e_b\rangle \in I(P)$ and if $P(z_a, \ldots, z_b)$ occurs negated in $\delta_i^0(z_1, \ldots, z_k)$ then let $\langle e_a, \ldots, e_b\rangle \notin I(P)$. (It may be assumed, for the sake of a complete definition, that for each predicate symbol P, $I(P)$ contains only those tuples mentioned above.) For those z_i which are constants, let $I(z_i) = e_i$. Let v be the valuation defined by $v(z_i) = e_i$ for those $i \in \{1, \ldots, k\}$ for which z_i is a variable. Then $\mathcal{M}, v \models \delta_i^0(z_1, \ldots, z_k)$.

Constituents of depth 1 with no free individual terms are of the form $(\pm)(\exists x)\gamma_1^0(x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_q^0(x)$. That is, they are the basic conjunctions generated by the set $\{(\exists x)\gamma_1^0(x), \ldots, (\exists x)\gamma_q^0(x)\}$. Each is satisfied by some model and valuation: For any model $\mathcal{M}$, for each element a of the domain, $\mathcal{M}, v(a/x) \models \gamma_i^0(x)$ for exactly one $i \in \{1, \ldots, q\}$ because $\gamma_1^0(x), \ldots, \gamma_q^0(x)$ are the basic conjunctions generated by some set of formulas, so exactly one of them is true for a given model and valuation ("each a satisfies exactly one $\gamma_i^0(x)$"). And each $\gamma_i^0(x)$ is consistent (shown in the same way as the fact that depth-0 constituents are consistent is shown above). So the formulas $(\exists x)\gamma_i^0(x)$ can each be made true or false independently of the others by the existence or lack of it of a suitable individual a. So each constituent $\delta^1()$ is consistent.

With even one free individual term, there are inconsistent depth-1 constituents. For example, any constituent of the form $\delta^1(x) = \delta^0(x) \wedge \neg(\exists y)\gamma_1^0(x,y) \wedge \ldots \wedge \neg(\exists y)\gamma_q^0(x,y)$ is inconsistent because it implies both the existence of some individual and that there are no individuals. For depths of 2 or more there are inconsistent constituents whether there are free individual terms or not. An example of an inconsistent depth-2 constituent with no free individual terms is given on page 41.

One possible way of obtaining a lower bound for the number (or fraction) of inconsistent constituents would be to try to count the *trivially inconsistent* constituents, using any definition of trivial inconsistency (see the next section). I don't know how difficult this would be, but a much simpler consideration gives us a lower bound for the fraction of inconsistent constituents which is extremely close to 1 (shown below). In other words, for depths of 2 or more, almost all the constituents are inconsistent. With such a high fraction of constituents inconsistent by this consideration alone, there doesn't seem to be much point in looking for higher lower bounds. But it would be interesting to try to find a lower bound for the fraction of not-(trivially-inconsistent) constituents that are actually inconsistent. (This would also be considerably more difficult and I have not attempted it.)

Constituents of depths $d > 0$ are of the form

$$\begin{aligned}\delta_i^d(z_1, \ldots, z_k) \;&=\; (\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_r(z_1, \ldots, z_k) \wedge \\ &\quad (\pm)(\exists x)\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_q^{d-1}(z_1, \ldots, z_k, x)\end{aligned}$$

Let p be the number of formulas $\alpha_1(z_1), \ldots, \alpha_r(z_1, \ldots, z_k)$. For $d = 1$ all the attributive constituents of the form $\gamma_j^{d-1}(z_1, \ldots, z_k, x)$ are consistent, for $d > 1$ some are inconsistent.

(How many of these formulas are inconsistent will be considered later.) We will first get a lower bound on the number of inconsistent constituents of the form $\delta^d(z_1, \ldots, z_k)$ in terms of a lower bound on the number of inconsistent attributive constituents of the form $\gamma^{d-1}(z_1, \ldots, z_k, x)$. For each inconsistent attributive constituent $\gamma_j^{d-1}(z_1, \ldots, z_k, x)$, to be consistent $\delta_i^d(z_1, \ldots, z_k)$ must contain $\neg(\exists x)\gamma_j^{d-1}(z_1, \ldots, z_k, x)$. So if there are at least n inconsistent attributive constituents of the form $\gamma_j^{d-1}(z_1, \ldots, z_k, x)$, then for $\delta_i^d(z_1, \ldots, z_k)$ to be consistent, at least n out of $p+q$ conjuncts are fixed, the others may either contain negation or not. So there are $\leq 2^{p+q-n}$ consistent constituents of the form $\delta_i^d(z_1, \ldots, z_k)$, so the fraction of consistent constituents is at most

$$\frac{2^{p+q-n}}{2^{p+q}} = \frac{1}{2^n}$$

and the fraction of inconsistent constituents is at least

$$1 - \frac{1}{2^n}.$$

Thus, for example, if there are at least 4 inconsistent attributive constituents of depth $d-1$ then over 90% of constituents of depth d (and every greater depth) are inconsistent. Actually, as soon as there are any inconsistent constituents (or attributive constituents), there are a lot more than 4 (shown below).

The above discussion also applies to attributive constituents, so we get the same lower bound for the number of inconsistent attributive constituents of depth d in terms of the number at depth $d-1$. To get actual lower bounds from this "relative" lower bound, we need some lower bound on the number of inconsistent attributive constituents of depth 1 with at least one free individual term. These attributive constituents are of the form (for $k \geq 1$)

$$\begin{aligned} \gamma_i^1(z_1, \ldots, z_k) &= \gamma_j^0(z_1, \ldots, z_k) \wedge \\ &\quad (\pm)(\exists x)\gamma_1^0(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_s^0(z_1, \ldots, z_k, x). \end{aligned} \tag{4.3}$$

I first consider the case where k is 1. Then we are dealing with attributive constituents of the form

$$\begin{aligned} \gamma^1(x) &= \gamma^0(x) \wedge \\ &\quad (\pm)(\exists y)\gamma_1^0(x, y) \wedge \ldots \wedge (\pm)(\exists y)\gamma_p^0(x, y) \end{aligned}$$

where $\gamma^0(x)$ is of the form

$$(\pm)P_1(x, \ldots, x) \wedge \ldots \wedge (\pm)P_m(x, \ldots, x)$$

where the arities of the predicates $P_1, \ldots, P_m$ are $n_1, \ldots, n_m$ respectively, and the formulas $\gamma_i^0(x, y)$ are of the form

$$\bigwedge_{\substack{z_1, \ldots, z_{n_1} \in \{x, y\} \\ \text{at least 1 } z \text{ is } y}} (\pm)P_1(z_1, \ldots, z_{n_1}) \wedge \ldots \wedge \bigwedge_{\substack{z_1, \ldots, z_{n_m} \in \{x, y\} \\ \text{at least 1 } z \text{ is } y}} (\pm)P_m(z_1, \ldots, z_{n_m}).$$

For each $\gamma^0(x)$, there is a unique $\gamma^0(x,y)$ such that if x is substituted for y in $\gamma^0(x,y)$ we get a formula which is equivalent to $\gamma^0(x)$, in fact it is the same as $\gamma^0(x)$ except that it contains repetitions of some conjuncts. It is the formula such that for each $i \in \{1,\ldots,m\}$, if $P_i(x,\ldots,x)$ is a conjunct of $\gamma^0(x)$, then each $P_i(z_1,\ldots,z_{n_i})$ is a conjunct of $\gamma^0(x,y)$, and if $\neg P_i(x,\ldots,x)$ is a conjunct of $\gamma^0(x)$, then each $\neg P_i(z_1,\ldots,z_{n_i})$ is a conjunct of $\gamma^0(x,y)$. This particular $\gamma^0(x,y)$ describes the y which is x, and thus must occur in $\gamma^1(x)$ for $\gamma^1(x)$ to be consistent. Half of the constituents / attributive constituents of the form $\gamma^1(x)$ do not satisfy this condition and thus are inconsistent. (We later see that for a language with only one binary predicate, these are the only inconsistent constituents of this form.)

The general case for $k \geq 1$ is similar. In (4.3), $\gamma_j^0(z_1,\ldots,z_k)$ expresses how z_k is related to $z_1,\ldots,z_{k-1}$ (if they exist) and how it is related to itself. For $\gamma_i^1(z_1,\ldots,z_k)$ to be consistent, z_k must exist as one of the x's. This condition can be expressed syntactically: There is a unique $\gamma_m^0(z_1,\ldots,z_k,x)$ such that if z_k is substituted for x in it, then ignoring repetitions of conjuncts, we get $\gamma_j^0(z_1,\ldots,z_k)$. To show this: Let $\mathbf{A}$ be the set of atomic formulas over $z_1,\ldots,z_k$ which contain z_k, then $\gamma_j^0(z_1,\ldots,z_k)$ is a basic conjunction generated by $\mathbf{A}$. And the formulas of the form $\gamma^0(z_1,\ldots,z_k,x)$ are the basic conjunctions generated by $\mathbf{B}$ where $\mathbf{B}$ is the set of atomic formulas over $z_1,\ldots,z_k,x$ which contain x. $\mathbf{B}$ can be obtained from $\mathbf{A}$ as follows: Each element of $\mathbf{A}$ contains some occurrences of z_k. For each $A \in \mathbf{A}$, let $\mathbf{B}_A$ be the set of all formulas which are A with x substituted for *at least one* occurrence of z_k. (So for those elements A of $\mathbf{A}$ in which z_k only occurs once, $\mathbf{B}_A$ is $\{A(x/z_k)\}$.) Then $\bigcup_{A\in\mathbf{A}} \mathbf{B}_A = \mathbf{B}$. Now, define $\gamma_m^0(z_1,\ldots,z_k,x)$ by: if A is a conjunct of $\gamma_j^0(z_1,\ldots,z_k)$ then let B be a conjunct of $\gamma_m^0(z_1,\ldots,z_k,x)$ for each $B \in \mathbf{B}_A$, and if $\neg A$ is a conjunct of $\gamma_j^0(z_1,\ldots,z_k)$ then let $\neg B$ be a conjunct of $\gamma_m^0(z_1,\ldots,z_k,x)$ for each $B \in \mathbf{B}_A$. Then $\gamma_m^0(z_1,\ldots,z_k,x)$ is the required formula. It will be called the *expanded copy* of $\gamma_j^0(z_1,\ldots,z_k)$. For $\gamma_i^1(z_1,\ldots,z_k)$ to be consistent, it must contain as a conjunct $(\exists x)\gamma_m^0(z_1,\ldots,z_k,x)$ (that is, it must contain positively $\gamma_m^0(z_1,\ldots,z_k,x)$) where $\gamma_m^0(z_1,\ldots,z_k,x)$ is the expanded copy of $\gamma_j^0(z_1,\ldots,z_k)$. (If $\mathcal{M},v \models \gamma_i^1(z_1,\ldots,z_k)$ then $\mathcal{M},v \models \gamma_j^0(z_1,\ldots,z_k)$, let $e = v(z_k)$, then $\mathcal{M},v(e/x) \models \gamma_m^0(z_1,\ldots,z_k,x)$, so $\mathcal{M},v \models (\exists x)\gamma_m^0(z_1,\ldots,z_k,x)$.) Thus at least half of the attributive constituents of the form $\gamma_i^1(z_1,\ldots,z_k)$ are inconsistent.

Now, the smallest possible number of attributive constituents is for a language with only one binary predicate, in which case there are $2^{(1+2^3)} = 512$ attributive constituents of depth 1 with any 1 (fixed) free individual term. With more free individual terms there are more attributive constituents.

So, for $k \geq 1$, at least 256 attributive constituents of the form $\gamma_i^1(z_1,\ldots,z_k)$ are inconsistent. Using this in the lower bound obtained earlier, we get that the fraction of depth 2 (or greater) constituents, irrespective of whether or not they have free individual terms, which are inconsistent is greater than or equal to

$$1 - \frac{1}{2^{256}}.$$

The case that remains to be considered is that of depth 1. All the constituents of depth 1 with no free individual terms are consistent. If there are any free individual terms then the same consideration that showed that at least half of the attributive constituents of depth 1 are inconsistent shows that at least half of the constituents of depth 1 are inconsistent.

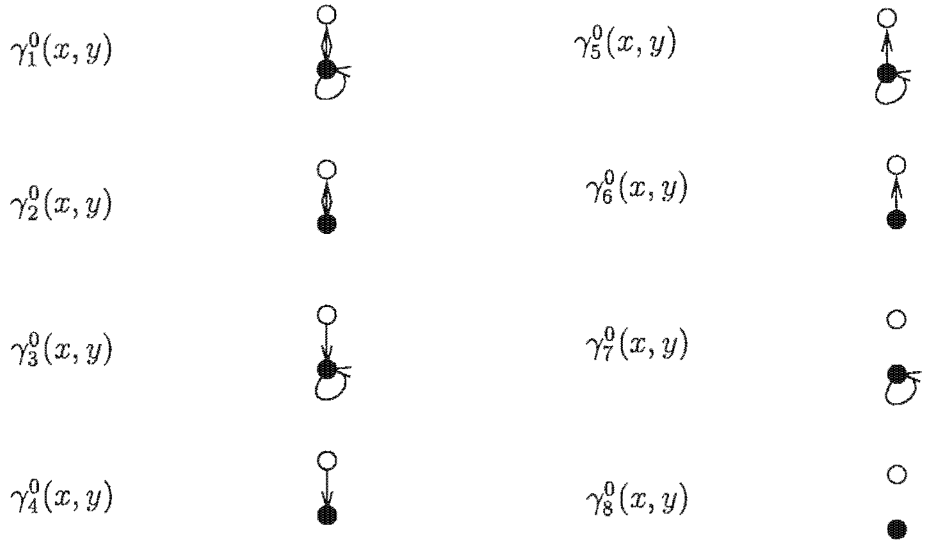
In summary, at depth 0 there are no inconsistent constituents. At depth 1, with no free individual terms there are no inconsistent constituents and with any free individual terms at least half of the constituents are inconsistent. And at depths 2 or greater, nearly all the constituents are inconsistent.

A special case

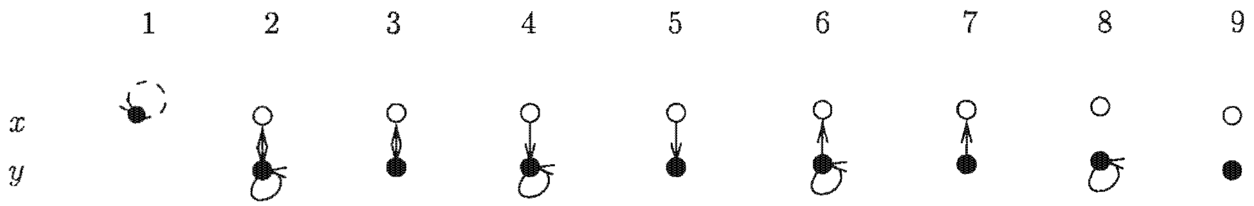
I now show that for the special case of a language which contains just one binary predicate, the only inconsistent constituents of depth 1 with 1 free individual term are those found to be inconsistent above, thus that exactly half of these constituents are consistent and half inconsistent. For the cases of no free individual terms and of one free individual term, the constituents are the attributive constituents. (This is immediate from the definition.) So the formulas we are now concerned with are those of the form

$$\delta^1(x) = \gamma^1(x) = \gamma^0(x) \wedge (\pm)(\exists y)\gamma^0_1(x, y) \wedge \ldots \wedge (\pm)(\exists y)\gamma^0_p(x, y).$$

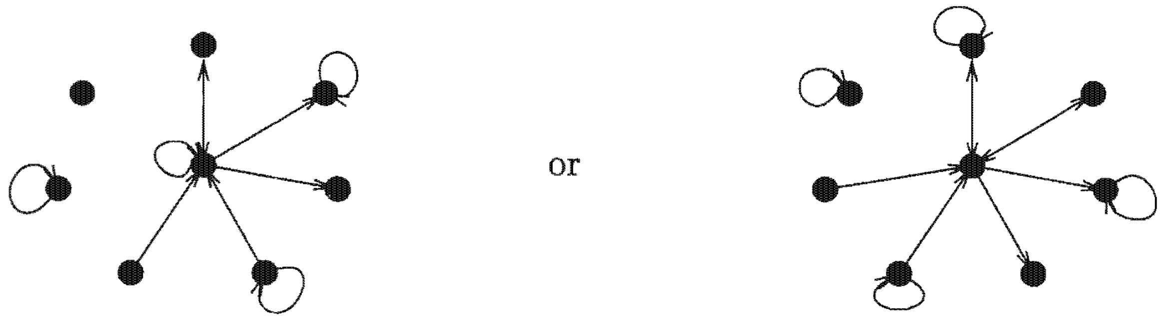
Consider the language which contains just the binary predicate P. A binary relation R on a small finite set S can be conveniently represented by a diagram, where the elements of S are represented as dots, possibly colored in or labelled to identify which elements they represent, and if $\langle x, y\rangle \in R$ then an arrow is drawn from the dot representing x to the dot representing y. The attributive constituents $\gamma^0(x, y)$ describe some individual y (actually, for any model and valuation they describe the individual to which y gets mapped, but I won't say this every time) relative to some individual x. That is, they describe the relation between x and y, given by the values of $P(x, y)$ and $P(y, x)$, and the relation of y to itself given by the value of $P(y, y)$, but they do not describe how x is related to itself. To represent these formulas I use uncolored dots for x, the reference individual, and colored-in dots for y, the individual being described. It is of course possible for x and y to be the same individual. This is not included in the following diagrams, but is taken into account later. The attributive constituents $\gamma^0(x, y)$ are represented as follows:



So $\gamma^1(x)$ can be specified by giving values for:



The parts of the above diagram labelled 1 to 9 correspond to the conjuncts of $\gamma^1(x)$, where each conjunct can either contain negation or not. In 1 the presence or absence of negation says whether or not the individual x is related to itself, represented in the diagram by the dashed arrow. In 2 to 9 the presence or absence of negation says whether or not an individual of a particular kind relative to x exists. We now consider what dependencies there are between (the formulas represented by) 1 to 9. By considering the case in which y is x, we see that $1 \Longrightarrow 2$ and $\neg 1 \Longrightarrow 9$. So, to be consistent, if $\gamma^1(x)$ contains 1 then it must contain 2, and if $\gamma^1(x)$ contains $\neg 1$ then it must contain 9. Half of the constituents / attributive constituents of the form $\gamma^1(x)$ do not satisfy this condition and are thus inconsistent. Since there are no dependencies between 2 to 9, the other half are all consistent. In particular, each is satisfied in some model of the form



where the center element in each diagram is x and is part of the model, and each of the surrounding elements is either part of the model or not.

4.5 Trivially inconsistent constituents

Those depth-1 constituents that were found to be inconsistent in the previous section are examples of *trivially inconsistent* constituents. There are no decidable conditions which are necessary and sufficient for the inconsistency of constituents. But there are decidable conditions which are sufficient, though not necessary, for inconsistency. (Thus it is necessary, but not sufficient, that a constituent not satisfy these conditions to be consistent.) Any condition which can be checked mechanically and is sufficient for the inconsistency of a constituent can be included in the definition of trivial inconsistency. The main significance of trivial inconsistency is that there are decidable sufficient conditions for inconsistency which are *adequate* for the *completeness theorem of the theory of distributive normal forms* (see section 4.7) to hold. Hintikka [1965a] gives two definitions of trivial inconsistency. It was thought that both were adequate for the completeness theorem, but due to an error in the attempted proof of this result (which I discuss on p. 114), it is unclear whether or not these definitions are adequate. There is, however, a definition of trivial inconsistency (given in the next section) closely related to one of Hintikka's which is adequate for the completeness theorem. This section gives Hintikka's definition of trivial inconsistency and an equivalent formulation of it in terms of the tree-structure of constituents and attributive constituents.

We consider some necessary conditions for consistency.

An attributive constituent $\gamma_i^d(z_1,\ldots,z_k)$ is of the form

$$\begin{aligned}&\gamma_j^0(z_1,\ldots,z_k)\wedge\\&(\pm)(\exists x)\gamma_1^{d-1}(z_1,\ldots,z_k,x)\wedge\ldots\wedge(\pm)(\exists x)\gamma_r^{d-1}(z_1,\ldots,z_k,x)\end{aligned}$$

where for each $a\in\{1,\ldots,r\}$, $\gamma_a^{d-1}(z_1,\ldots,z_k,x)$ is of the form

$$\begin{aligned}&\gamma_b^0(z_1,\ldots,z_k,x)\wedge\\&(\pm)(\exists y)\gamma_1^{d-2}(z_1,\ldots,z_k,x,y)\wedge\ldots\wedge(\pm)(\exists y)\gamma_s^{d-2}(z_1,\ldots,z_k,x,y).\end{aligned}$$

The case we saw in the previous section was when $d=1$, in which case each attributive constituent $\gamma_a^{d-1}(z_1,\ldots,z_k,x)$ is just a root $\gamma_b^0(z_1,\ldots,z_k,x)$. For $\gamma_i^1(z_1,\ldots,z_k)$ to be consistent, z_k must exist as one of the x's. This can be expressed formally by the condition that the expanded copy $\gamma_c^0(z_1,\ldots,z_k,x)$ of $\gamma_j^0(z_1,\ldots,z_k)$ must be a branch of $\gamma_j^0(z_1,\ldots,z_k)$. In the general case, instead of the formulas $\gamma_a^{d-1}(z_1,\ldots,z_k,x)$ being depth-0 descriptions of x, they include such descriptions $\gamma_b^0(z_1,\ldots,z_k,x)$ as their roots. And it is necessary that in $\gamma_i^d(z_1,\ldots,z_k)$, $\gamma_j^0(z_1,\ldots,z_k)$ have some branch $\gamma_a^{d-1}(z_1,\ldots,z_k,x)$ of which the expanded copy of $\gamma_j^0(z_1, \ldots, z_k)$ is the root, to be consistent (because for $d > 1$, an attributive constituent $\gamma^d(z_1, \ldots, z_k)$ implies the attributive constituent $\gamma^1(z_1, \ldots, z_k)$ which is its top-2 tree, by corollary 4.7). In other words, if $\gamma_i^d(z_1, \ldots, z_k)$ is consistent then $\gamma_j^0(z_1, \ldots, z_k)$ has some successor which is the expanded copy of itself. This necessary condition for consistency can be strengthened. To do this, we use the following terminology:

The constituents and attributive constituents previously defined are called constituents and attributive constituents *of the first kind* by Hintikka [1965a]. They can be rewritten in an equivalent form as follows: An attributive constituent of the form

$$\begin{aligned}\gamma_i^d(z_1, \ldots, z_k) &= (\pm)\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)\alpha_q(z_1, \ldots, z_k) \wedge \\ &\quad (\pm)(\exists x)\gamma_1^{d-1}(z_1, \ldots, z_k, x) \wedge \ldots \wedge (\pm)(\exists x)\gamma_r^{d-1}(z_1, \ldots, z_k, x)\end{aligned}$$

can be written as

$$\begin{aligned}\gamma_i^d(z_1, \ldots, z_k) &= (\pm)\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)\alpha_q(z_1, \ldots, z_k) \wedge \\ &\quad \bigwedge_{a \in A}(\exists x)\gamma_a^{d-1}(z_1, \ldots, z_k, x) \wedge \bigwedge_{b \in B} \neg(\exists x)\gamma_b^{d-1}(z_1, \ldots, z_k, x)\end{aligned}$$

where $\{A, B\}$ is a partition of $\{1, \ldots, r\}$. The above attributive constituent is equivalent to

$$\begin{aligned}&(\pm)\alpha_1(z_1, \ldots, z_k) \wedge \ldots \wedge (\pm)\alpha_q(z_1, \ldots, z_k) \wedge \\ &\bigwedge_{a \in A}(\exists x)\gamma_a^{d-1}(z_1, \ldots, z_k, x) \wedge (\forall x)(\bigvee_{a \in A} \gamma_a^{d-1}(z_1, \ldots, z_k, x)).\end{aligned}$$

And a constituent

$$\begin{aligned}\delta_i^d(z_1, \ldots, z_k) &= (\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_p(z_1, \ldots, z_k) \wedge \\ &\quad \bigwedge_{a \in A}(\exists x)\gamma_a^{d-1}(z_1, \ldots, z_k, x) \wedge \bigwedge_{b \in B} \neg(\exists x)\gamma_b^{d-1}(z_1, \ldots, z_k, x)\end{aligned}$$

is equivalent to

$$\begin{aligned}&(\pm)\alpha_1(z_1) \wedge \ldots \wedge (\pm)\alpha_p(z_1, \ldots, z_k) \wedge \\ &\bigwedge_{a \in A}(\exists x)\gamma_a^{d-1}(z_1, \ldots, z_k, x) \wedge (\forall x)(\bigvee_{a \in A} \gamma_a^{d-1}(z_1, \ldots, z_k, x)).\end{aligned}$$

Both of the above equivalences are shown as follows: For $\{A, B\}$ a partition of $\{1, \ldots, r\}$ as in the above formulas,

$$\begin{aligned}&\bigwedge_{a \in A}(\exists x)\gamma_a^{d-1}(z_1, \ldots, z_k, x) \wedge \bigwedge_{b \in B} \neg(\exists x)\gamma_b^{d-1}(z_1, \ldots, z_k, x) \\ \iff &\bigwedge_{a \in A}(\exists x)\gamma_a^{d-1}(z_1, \ldots, z_k, x) \wedge \neg(\bigvee_{b \in B}(\exists x)\gamma_b^{d-1}(z_1, \ldots, z_k, x))\end{aligned}$$

$$\iff \bigwedge_{a\in A}(\exists x)\gamma_a^{d-1}(z_1,\ldots,z_k,x) \wedge \neg((\exists x)\bigvee_{b\in B}\gamma_b^{d-1}(z_1,\ldots,z_k,x))$$
$$\iff \bigwedge_{a\in A}(\exists x)\gamma_a^{d-1}(z_1,\ldots,z_k,x) \wedge (\forall x)(\neg\bigvee_{b\in B}\gamma_b^{d-1}(z_1,\ldots,z_k,x))$$
$$\iff \bigwedge_{a\in A}(\exists x)\gamma_a^{d-1}(z_1,\ldots,z_k,x) \wedge (\forall x)(\bigwedge_{b\in B}\neg\gamma_b^{d-1}(z_1,\ldots,z_k,x))$$
$$\iff \bigwedge_{a\in A}(\exists x)\gamma_a^{d-1}(z_1,\ldots,z_k,x) \wedge (\forall x)(\bigvee_{a\in A}\gamma_a^{d-1}(z_1,\ldots,z_k,x)).$$

The final equivalence above holds because $\bigwedge_{b\in B}\neg\gamma_b^{d-1}(z_1,\ldots,z_k,x) \iff \bigvee_{a\in A}\gamma_a^{d-1}(z_1,\ldots,z_k,x)$ by corollary 3.3. Constituents and attributive constituents defined in this equivalent form are said to be *of the second kind* (Hintikka [1965a]). I use the same notation for constituents (and attributive constituents) of the first and second kinds. Where it is necessary to distinguish between them, the context will make clear which kind is meant. The definitions are as follows:

Attributive constituents of the second kind are the same as the first kind for depth 0 and for greater depths d are defined by

$$\gamma_i^d(z_1,\ldots,z_k) \;=\; (\pm)\alpha_1(z_1,\ldots,z_k)\wedge\ldots\wedge(\pm)\alpha_q(z_1,\ldots,z_k)\wedge \bigwedge_{a\in A}(\exists x)\gamma_a^{d-1}(z_1,\ldots,z_k,x)\wedge(\forall x)(\bigvee_{a\in A}\gamma_a^{d-1}(z_1,\ldots,z_k,x))$$

where $\alpha_1(z_1,\ldots,z_k),\ldots,\alpha_q(z_1,\ldots,z_k)$ are all the atomic formulas over $\{z_1,\ldots,z_k\}$ which contain z_k.

Constituents of the second kind are the same as the first kind for depth 0 and for greater depths d are defined by

$$\delta_i^d(z_1,\ldots,z_k) \;=\; (\pm)\alpha_1(z_1)\wedge\ldots\wedge(\pm)\alpha_p(z_1,\ldots,z_k)\wedge \bigwedge_{a\in A}(\exists x)\gamma_a^{d-1}(z_1,\ldots,z_k,x)\wedge(\forall x)(\bigvee_{a\in A}\gamma_a^{d-1}(z_1,\ldots,z_k,x))$$

where $\alpha_1(z_1),\ldots,\alpha_p(z_1,\ldots,z_k)$ are all the atomic formulas over $\{z_1,\ldots,z_k\}$.

In both of the above definitions, A is some subset of $\{a \mid \gamma_a^{d-1}(z_1,\ldots,z_k,x)$ is an attributive constituent of the second kind of depth $d-1$ with free individual terms $z_1,\ldots,z_k,x\}$.

As for constituents of the first kind, I use the terminology of roots and branches obtained from the sets-and-trees representation for constituents and attributive constituents of the second kind. In fact, the syntactic structure of constituents and attributive constituents of the second kind corresponds closely with their sets-and-trees representation in that the attributive constituents which occur in them are exactly the branches of the tree. Or equivalently, if all the attributive constituents are written out in terms of depth-0 attributive constituents, all these depth-0 attributive constituents (and for constituents, also the depth-0 constituent which is the root) are exactly the nodes of the tree. Hintikka [1953] (p. 39) says: "For the purpose of formulating metatheorems concerning constituents and normal forms, the constituents and normal forms of the second kind possess some considerable advantage over those of the first kind." I'm not sure what advantage is meant here, but one use of constituents of the second kind is that the definition of trivial inconsistency can be stated in terms of them, though no more easily than in terms of the sets-and-trees representation.

Trivial inconsistency is defined in such a way that if any of certain formulas which are implied by a constituent or attributive constituent is *directly contradictory* (as defined below), then the constituent or attributive constituent is trivially inconsistent.

A formula is *directly contradictory* if

- it is a conjunction of a number of formulas one of which is the negation of another, or
- it is of the form $(\exists x)A$ and A is directly contradictory, or
- it is a conjunction of which some conjunct is directly contradictory.

We can now express the strengthened necessary condition for consistency:

An attributive constituent

$$\begin{aligned}\gamma_i^d(z_1,\ldots,z_k) \;=\; & \gamma_a^0(z_1,\ldots,z_k)\wedge \\ & \bigwedge_{b\in B}(\exists x)\gamma_b^{d-1}(z_1,\ldots,z_k,x)\wedge(\forall x)(\bigvee_{b\in B}\gamma_b^{d-1}(z_1,\ldots,z_k,x))\end{aligned}$$

implies

$$\gamma_a^0(z_1,\ldots,z_k)\wedge\bigvee_{b\in B}(\gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x))$$

for each $i\in\{1,\ldots,k\}$. And

$$\gamma_a^0(z_1,\ldots,z_k)\wedge\bigvee_{b\in B}(\gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x))$$

can only be consistent if there is some $b\in B$ for which

$$\gamma_a^0(z_1,\ldots,z_k)\wedge(\gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x))$$

is not directly contradictory. Any $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ for which $\gamma_a^0(z_1,\ldots,z_k)\wedge(\gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x))$ is not directly contradictory will be called a *partial expansion* to depth $d-1$ of $\gamma_a^0(z_1,\ldots,z_k)$ with respect to z_i. This terminology is chosen because there is a unique partial expansion to depth 0 of $\gamma_a^0(z_1,\ldots,z_k)$ with respect to z_k, and it is the expanded copy of $\gamma_a^0(z_1,\ldots,z_k)$.

I now consider some examples. For the language with one binary predicate symbol P the above condition is used to show the inconsistency of certain attributive constituents of the form $\gamma^1(z_1, x)$. Let

$$\begin{aligned}
\gamma_a^0(z_1, x) &= P(z_1, x) \wedge \neg P(x, z_1) \wedge P(x, x) \\
\gamma_b^0(z_1, x, y) &= P(z_1, y) \wedge \neg P(y, z_1) \wedge P(x, y) \wedge P(y, x) \wedge P(y, y) \\
\gamma_c^1(z_1, x) &= \gamma_a^0(z_1, x) \wedge (\exists y)\gamma_b^0(z_1, x, y) \wedge (\forall y)\gamma_b^0(z_1, x, y).
\end{aligned}$$

Then, applying the condition in which we substitute x for y,

$$\begin{aligned}
\gamma_c^1(z_1, x) \implies\ & P(z_1, x) \wedge \neg P(x, z_1) \wedge P(x, x) \wedge \\
& P(z_1, x) \wedge \neg P(x, z_1) \wedge P(x, x) \wedge P(x, x) \wedge P(x, x)
\end{aligned}$$

which is not directly contradictory. And, applying the condition in which we substitute z_1 for y,

$$\begin{aligned}
\gamma_c^1(z_1, x) \implies\ & P(z_1, x) \wedge \neg P(x, z_1) \wedge P(x, x) \wedge \\
& P(z_1, z_1) \wedge \neg P(z_1, z_1) \wedge P(x, z_1) \wedge P(z_1, x) \wedge P(z_1, z_1)
\end{aligned}$$

which is directly contradictory. Thus $\gamma_c^1(z_1, x)$ is inconsistent. Now, let

$$\begin{aligned}
\gamma_d^0(z_1, x, y) &= P(z_1, y) \wedge P(y, z_1) \wedge \neg P(x, y) \wedge P(y, x) \wedge P(y, y) \\
\gamma_e^0(z_1, x, y) &= \neg P(z_1, y) \wedge \neg P(y, z_1) \wedge \neg P(x, y) \wedge P(y, x) \wedge \neg P(y, y) \\
\gamma_f^1(z_1, x) &= \gamma_a^0(z_1, x) \wedge \bigwedge_{i \in \{b,d\}} (\exists y)\gamma_i^0(z_1, x, y) \wedge (\forall y) \bigvee_{i \in \{b,d\}} \gamma_i^0(z_1, x, y) \\
\gamma_g^1(z_1, x) &= \gamma_a^0(z_1, x) \wedge \bigwedge_{i \in \{b,e\}} (\exists y)\gamma_i^0(z_1, x, y) \wedge (\forall y) \bigvee_{i \in \{b,e\}} \gamma_i^0(z_1, x, y).
\end{aligned}$$

Then, in $\gamma_f^1(z_1, x)$, applying the condition in which we substitute x for y does not show any inconsistency because $\gamma_f^1(z_1, x)$ contains the branch $\gamma_b^0(z_1, x, y)$ which we have seen above gives a result which is not directly contradictory. And, applying the condition in which we substitute z_1 for y,

$$\begin{aligned}
\gamma_f^1(z_1, x) \implies\ & \gamma_a^0(z_1, x) \wedge \\
& (\gamma_b^0(z_1, x, y)(z_1/y) \vee (P(z_1, z_1) \wedge P(z_1, z_1) \wedge \neg P(x, z_1) \wedge P(z_1, x) \wedge P(z_1, z_1))).
\end{aligned}$$

Now, $\gamma_b^0(z_1, x, y)(z_1/y)$ is directly contradictory, so also $\gamma_a^0(z_1, x) \wedge \gamma_b^0(z_1, x, y)(z_1/y)$ is directly contradictory; but $\gamma_a^0(z_1, x) \wedge P(z_1, z_1) \wedge P(z_1, z_1) \wedge \neg P(x, z_1) \wedge P(z_1, x) \wedge P(z_1, z_1)$ is not directly contradictory. So this consideration does not show whether or not $\gamma_f^1(z_1, x)$ is consistent. Note that for an attributive constituent of the form $\gamma^1(z_1, x)$, for x as described by the depth-0 part of $\gamma^1(z_1, x)$ there is a unique branch which $\gamma^1(z_1, x)$ must have in order not to be found inconsistent by this consideration (in the above examples, this branch is $\gamma_b^0(z_1, x, y)$). But for z_1 there is not a unique such branch, as is shown by the attributive constituents $\gamma_f^1(z_1, x)$ and $\gamma_g^1(z_1, x)$, where one contains $\gamma_d^0(z_1, x, y)$ and the other contains $\gamma_e^0(z_1, x, y)$, and both $\gamma_d^0(z_1, x, y)$ and $\gamma_e^0(z_1, x, y)$ are partial expansions of $\gamma_a^0(z_1, x)$ to depth 0 with respect to z_1. For depth-0 constituents, partial expansions can be defined in the same way as for attributive constituents, and then there are unique partial expansions to depth 0 with respect to *any* of the free individual terms. This is because depth-0 constituents fully describe the depth-0 relations between all the free individual terms. The reason we don't have unique partial expansions for attributive constituents is that they can differ in conjuncts which, after substituting, describe the way free individual terms other than the last are related to each other (since this is not specified by the attributive constituent for which we are obtaining a partial expansion). In the above example, $\gamma_d^0(z_1, x, y)$ and $\gamma_e^0(z_1, x, y)$ differ in this way when z_1 is substituted for y.

Since a constituent $\delta^0(z_1, \ldots, z_k)$ has a unique partial expansion to depth 0 with respect to any z in $\{z_1, \ldots, z_k\}$, this partial expansion can be called the *expanded copy of* $\delta^0(z_1, \ldots, z_k)$ *with respect to* z. Expanded copies of depth-0 constituents can also be defined more directly by: For each $i \in \{1, \ldots, k\}$, the *expanded copy* of $\delta^0(z_1, \ldots, z_k)$ with respect to z_i is the constituent $\delta^0(z_1, \ldots, z_k, x)$ which contains every conjunct of $\delta^0(z_1, \ldots, z_k)$ and for each atomic formula A over $z_1, \ldots, z_k, x$ which contains x, if $A(z_i/x)$ is a conjunct of $\delta^0(z_1, \ldots, z_k)$ then A is a conjunct of $\delta^0(z_1, \ldots, z_k, x)$ and if $\neg A(z_i/x)$ is a conjunct of $\delta^0(z_1, \ldots, z_k)$ then $\neg A$ is a conjunct of $\delta^0(z_1, \ldots, z_k, x)$.

We have seen that for an attributive constituent $\gamma^d(z_1, \ldots, z_k)$ to be consistent, its root $\gamma^0(z_1, \ldots, z_k)$ must have, for each $i \in \{1, \ldots, k\}$, a branch which is a partial expansion of $\gamma^0(z_1, \ldots, z_k)$ with respect to z_i. If $\gamma^d(z_1, \ldots, z_k)$ is consistent then each of its principal subtrees is consistent, so a similar condition to the one above applies to all nodes which have successors: If $\gamma^d(z_1, \ldots, z_k)$ is consistent, then each node $\gamma^0(z_1, \ldots, z_l)$ which has successors has, for each $i \in \{1, \ldots, l\}$, a branch which is a partial expansion of $\gamma^0(z_1, \ldots, z_l)$ with respect to z_l.

We now consider another necessary condition for consistency of constituents and attributive constituents.

In an attributive constituent $\gamma^d(z_1, \ldots, z_k)$ of the second kind, to *omit a layer of quantifiers* means, for some x_i over which quantification occurs in $\gamma^d(z_1, \ldots, z_k)$, to remove all occurrences of $(\exists x_i)$ and $(\forall x_i)$ together with all atomic formulas containing x_i and omit unnecessary repetitions of formulas. In addition, we must either consider formulas which differ only in their bound variables as notational variations of each other, or change the bound variables to be used in alphabetic order. Either way, it is then possible to get the same result by omitting different layers of quantifiers. I will assume that the bound variables are changed to be used in alphabetic order, even though this is not part of the original definition, since it allows for greater precision. Omitting layers of quantifiers is described by Hintikka [1965a] (p. 68–71) and the definition is stated clearly by Niiniluoto [1987] (p. 71). We can see what kind of formulas are obtained by omitting layers of quantifiers from attributive constituents by considering some examples. Omitting any layer of quantifiers from an attributive constituent $\gamma^d(z_1,\ldots,z_k)$ results in a formula which is somewhat like an attributive constituent of the form $\gamma^{d-1}(z_1,\ldots,z_k)$, but may within the scope of an existential quantifier, or at the outermost level, contain formulas which are conjunctions where either one conjunct is the negation of another, or conflicting lists of kinds of individuals are asserted to exist. In other words, the formula obtained after the particular bound variable has been removed is like an attributive constituent, but is "over-defined" at some level. Only if this "over-definition" is actually redundant (i.e. repetition) is the formula obtained after omitting repetitions (and changing bound variables) an attributive constituent of the form $\gamma^{d-1}(z_1,\ldots,z_k)$. If the result of omitting some layer of quantifiers from $\gamma^d(z_1,\ldots,z_k)$ is not an attributive constituent $\gamma^{d-1}(z_1,\ldots,z_k)$ then this result is inconsistent. Since $\gamma^d(z_1,\ldots,z_k)$ implies this inconsistent formula, it is also inconsistent. And, since only one of the formulas of the form $\gamma^{d-1}(z_1,\ldots,z_k)$ is true for any particular model and valuation, the results of omitting each of the layers of quantifiers from $\gamma^d(z_1,\ldots,z_k)$ must all be the same (up to notational variation) for $\gamma^d(z_1,\ldots,z_k)$ to be consistent since $\gamma^d(z_1,\ldots,z_k)$ implies each of these formulas. Omitting a layer of quantifiers in constituents of the second kind is defined in the same way as in attributive constituents.

The necessary conditions for the consistency of constituents and attributive constituents which we have just seen are used in the following definition of trivial inconsistency. Since I later consider a different definition of trivial inconsistency, I call this kind *one-trivial inconsistency.*

An attributive constituent

$$\gamma^d(z_1,\ldots,z_k) = \gamma^0(z_1,\ldots,z_k) \wedge \bigwedge_{b\in B}(\exists x)\gamma_b^{d-1}(z_1,\ldots,z_k,x) \wedge (\forall x)(\bigvee_{b\in B}\gamma_b^{d-1}(z_1,\ldots,z_k,x))$$

is called *one-trivially inconsistent* if it does not satisfy the following necessary conditions for consistency:

- for each $i \in \{1,\ldots,k\}$, $\gamma^0(z_1,\ldots,z_k)$ has some branch which is a partial expansion to depth $d-1$ of $\gamma^0(z_1,\ldots,z_k)$ with respect to z_i;
- omitting the different layers of quantifiers all give the same result, which is an attributive constituent of depth $d-1$ with free individual terms $z_1,\ldots,z_k$;
- it does not contain any attributive constituent which is one-trivially inconsistent.

A constituent

$$\delta^d(z_1,\ldots,z_k) = \delta^0(z_1,\ldots,z_k) \wedge \bigwedge_{b\in B}(\exists x)\gamma_b^{d-1}(z_1,\ldots,z_k,x) \wedge (\forall x)(\bigvee_{b\in B}\gamma_b^{d-1}(z_1,\ldots,z_k,x))$$

is called *one-trivially inconsistent* if it does not satisfy the following necessary conditions for consistency:

- for each $i \in \{1, \ldots, k\}$, $\delta^0(z_1, \ldots, z_k)$ has some branch which is a partial expansion to depth $d-1$ of $\delta^0(z_1, \ldots, z_k)$ with respect to z_i;
- omitting the different layers of quantifiers all give the same result, which is a constituent of depth $d-1$ with free individual terms $z_1, \ldots, z_k$;
- it does not contain any attributive constituent which is one-trivially inconsistent.

We can rewrite the above definition taking into account the condition that the attributive constituents contained in the constituent or attributive constituent are not one-trivially inconsistent. If omitting each different layer of quantifiers results in a constituent or attributive constituent, then this condition also applies to the branches of each node in the tree. But, even if omitting the different layers of quantifiers in $\gamma^d(z_1, \ldots, z_k)$ result in *the same* attributive constituent, this does not necessarily apply to the principal subtrees of $\gamma^d(z_1, \ldots, z_k)$. And, the existence of the required partial expansions of the root does not imply the existence of partial expansions of other nodes. Also, if omitting the different layers of quantifiers all give the same result, then this result is a constituent or attributive constituent as required above. So we can express the above definition of one-trivial inconsistency as follows:

A constituent or attributive constituent is *one-trivially inconsistent* if it does not satisfy the following conditions:

- each node N which has successors has, for each free individual term z in N, some branch which is a partial expansion of N to depth n with respect to z, where n is the depth of the branches of N
- in each principal subtree, omitting the different layers of quantifiers all give the same result.

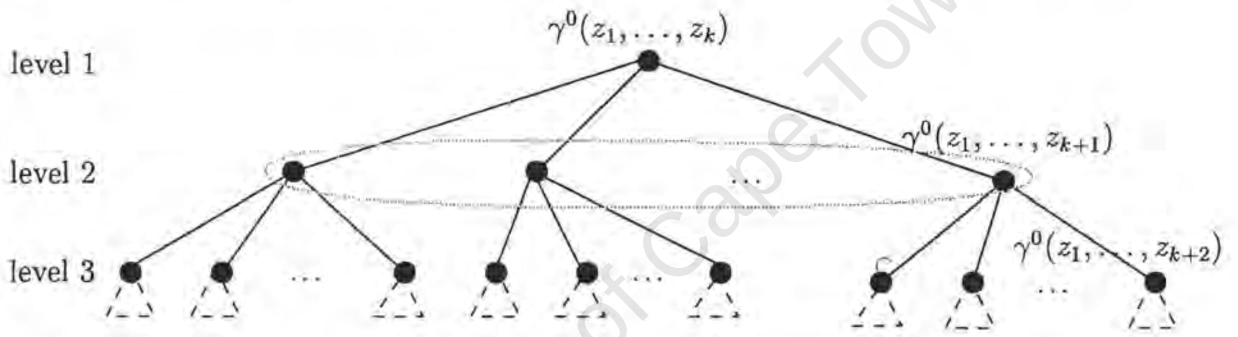
The above definition of one-trivial inconsistency is equivalent to the one in Hintikka [1965a] (p. 67–71). There are a number of other definitions in the literature. Hintikka [1965a] (p. 66–67) gives a definition which, although not exactly equivalent to this one, is according to *ibid.* (p. 71) "equally powerful" in the sense that for any constituent or attributive constituent which is trivially inconsistent by one definition but not by the other, its expansion to depth 1-greater than its own is also trivially inconsistent by the other definition. Hintikka [1970] (p. 270–272) gives a definition of trivial inconsistency which is equivalent to the one in Hintikka [1965a] (p. 67–71), though it uses slightly different terminology. Niiniluoto [1987] (p. 71) gives a definition in terms of a "truncation requirement" which is the same condition regarding the omission of layers of quantifiers as in the definition above, and a "repetition requirement" which seems to be the special case of the above requirement regarding the existence of partial expansions applied to sentence constituents.

In the above definition of one-trivial inconsistency, the condition regarding the existence of partial expansions is already expressed in terms of the tree-structure of constituents and attributive constituents. individual The other condition in this definition, involving the omission of layers of quantifiers, can also be expressed in terms of the sets-and-trees representation. For this, we will use sibling-sets (defined on p. 70) and the following terminology: The *predecessor* of a sibling-set is the predecessor of each of its elements. The *predecessor-set* of a sibling-set A is the sibling-set of which the predecessor of A is an element. The *successor-sets* of a sibling-set A are the sibling-sets whose elements are the successors of the elements of A. These definitions can also be stated as follows: Given a tree $\mathcal{T}$, let the *quotient tree* of $\mathcal{T}$ be the tree whose elements are the sibling-sets of $\mathcal{T}$ and whose successor/predecessor relation R is defined by: xRy iff the predecessor of the elements of y is an element of x. (This terminology is chosen because on the set of nodes of the tree omitting its root, the relation of being siblings is an equivalence relation and the sibling-sets are its equivalence classes, and the quotient tree is the tree whose nodes are these equivalence classes which preserves the relation of the original tree.) Then predecessor-sets in $\mathcal{T}$ are predecessors in the quotient tree of $\mathcal{T}$ and successor-sets in $\mathcal{T}$ are successors in the quotient tree of $\mathcal{T}$. The *nodes of level* i in a forest are those of level i in a tree in the forest, and for $i > 1$ a *sibling-set of level* i in a forest is a sibling-set of level i in a tree in the forest, and the set whose elements are all the nodes of level 1 is the *sibling-set of level 1*. In a tree (or forest) $\mathcal{T}$, the *subforest generated by a sibling-set* S is the forest whose nodes are the elements of S and those nodes which are below any element of S and whose predecessor/successor relation is the restriction of that of $\mathcal{T}$ to those nodes.

The expanded copy $\gamma^0(z_1, \ldots, z_k, z_{k+1})$ of a depth-0 attributive constituent $\gamma^0(z_1, \ldots, z_k)$ describes a kind of individual relative to one more individual than does $\gamma^0(z_1, \ldots, z_k)$. For an attributive constituent $\gamma^0(z_1, \ldots, z_k, z_{k+1})$, we define a particular attributive constituent $\gamma^0(z_1, \ldots, z_k)$ which describes a kind of individual relative to one fewer individual than does $\gamma^0(z_1, \ldots, z_k, z_{k+1})$ as follows: The *contracted copy* of $\gamma^0(z_1, \ldots, z_k, z_{k+1})$ is the formula obtained from $\gamma^0(z_1, \ldots, z_k, z_{k+1})$ by deleting all the conjuncts which contain z_k and then, in the resulting formula (which is of the form $\gamma^0(z_1, \ldots, z_{k-1}, z_{k+1})$), substituting z_k for z_{k+1}. Note that if $\gamma^0(z_1, \ldots, z_k)$ is the contracted copy of $\gamma^0(z_1, \ldots, z_k, z_{k+1})$ then $(\exists z_{k+1})\gamma^0(z_1, \ldots, z_k, z_{k+1}) \Longrightarrow (\exists z_k)\gamma^0(z_1, \ldots, z_k)$ (since any conjunction implies the conjunction of some of its conjuncts, and a change of bound variable results in an equivalent formula). The definition of contracted copy follows the structure of an attributive constituent in the sense that in an attributive constituent formulas of the form of the contracted copy of some depth-0 attributive constituent occur at one level of quantification less, until the outermost level is reached. In terms of the tree-structure, looking at (the labels of) the nodes of the different levels, going one level further away from the root, they are of the form of the expanded copies of those of the given level; and going one level closer to the root, they are of the form of the contracted copies of those of the given level.

The definition of the contracted copy of a depth-0 attributive constituent is generalized to attributive constituents of any depth as follows: The *contracted copy* of an attributive constituent $\gamma^d(z_1, \ldots, z_k, z_{k+1})$ is the formula obtained from it by deleting all atomic formulas which contain z_k and then substituting z_k for z_{k+1}, and then changing the bound variables from $z_{k+2}, \ldots, z_{k+d+1}$ to $z_{k+1}, \ldots, z_{k+d}$ (simultaneously) and in the resulting tree omitting duplicates. Thus the contracted copy of $\gamma^d(z_1, \ldots, z_k, z_{k+1})$ is of the form $\gamma^d(z_1, \ldots, z_k)$. In constituents and attributive constituents, existential conjunctions occur as subforests generated by sibling-sets. In those trees and forests used to represent constituents and attributive constituents, for a sibling-set S, the *contracted copy* of the subforest F generated by S is the forest whose trees are the contracted copies of the trees of F, with duplicates removed.

We now consider the effect of omitting a layer of quantifiers from an attributive constituent in terms of its tree-structure. In an attributive constituent of the form



omitting the first layer of quantifiers corresponds to "omitting" the circled level of the tree. The condition that omitting this layer of quantifiers results in an attributive constituent can be expressed as: for each sibling-set T of level 3, the contracted copy of the subforest generated by T must be the same. More precisely, there must be some forest F such that for each sibling-set T of level 3, the contracted copy of the subforest generated by T is F. For consistency, F must in fact be the top-$(d-1)$ forest of the subforest generated by the sibling-set of level 2, which it is if omitting the first and last layers of quantifiers give the same result. Omitting the second layer of quantifiers corresponds to "omitting" the third level of the tree, which is "omitting" the second level in each subtree generated by a node of level 2 in the original tree. Each of these subtrees must be consistent for the whole tree to be consistent, and if for each of them, omitting the first layer of quantifiers gives the required top-tree, then for the whole tree, omitting the second layer of quantifiers gives the required top-tree.

The following lemma will be used in section 4.7. To state it, I use the notation $\gamma^d(\{z_1, \ldots, z_k\}, z_i)$ for an attributive constituent of depth d whose free individual terms are $z_1, \ldots, z_k$ (but not necessarily in this order) where the last free individual term is z_i. For any constituent $\delta^d(z_1, \ldots, z_k)$, for each $i \in \{1, \ldots, k\}$, the attributive constituent

$\gamma^d(\{z_1,\ldots,z_k\},z_i)$ formed from $\delta^d(z_1,\ldots,z_k)$ by deleting some of the conjuncts from its root is called a *part of* $\delta^d(z_1,\ldots,z_k)$.

Lemma 4.10 *If $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ is a partial expansion to depth $d-1$ of $\delta_a^0(z_1,\ldots,z_k)$ with respect to z_i for some $i \in \{1,\ldots,k\}$, then substituting z_i for x in the root of $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ gives an attributive constituent $\gamma^0(\{z_1,\ldots,z_k\},z_i)$ which is a part of $\delta_a^0(z_1,\ldots,z_k)$, and substituting z_i for x in the subforest generated by the sibling-set of level 2 in $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ has the same effect as deleting all conjuncts which contain x from this subforest.*

PROOF If $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ is a partial expansion to depth $d-1$ of $\delta_a^0(z_1,\ldots,z_k)$ with respect to z_i, then $\delta_a^0(z_1,\ldots,z_k) \wedge \gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x)$ is not directly contradictory. Let the root of $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ be called $\gamma_c^0(z_1,\ldots,z_k,x)$. Then $\gamma_c^0(z_1,\ldots,z_k,x)(z_i/x)$ contains all the atomic formulas over $z_1,\ldots,z_k$ which contain z_i (some of them occur more than once), and $\delta_a^0(z_1,\ldots,z_k)$ contains all the atomic formulas over $z_1,\ldots,z_k$. So, since $\delta_a^0(z_1,\ldots,z_k) \wedge \gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x)$ is not directly contradictory, each conjunct of $\gamma_c^0(z_1,\ldots,z_k,x)(z_i/x)$ is a conjunct of $\delta_a^0(z_1,\ldots,z_k)$. The nodes in the subtrees $\gamma^{d-2}(z_1,\ldots,z_k,x,y)$ generated by successors of $\gamma_c^0(z_1,\ldots,z_k,x)$ are all of the form $\gamma^0(z_1,\ldots,z_k,x,y,\ldots,z)$ for some z. Since $\delta_a^0(z_1,\ldots,z_k) \wedge \gamma_b^{d-1}(z_1,\ldots,z_k,x)(z_i/x)$ is not directly contradictory, for each such node $\gamma^0(z_1,\ldots,z_k,x,y,\ldots,z)$, $\gamma^0(z_1,\ldots,z_k,x,y,\ldots,z)$ (z_i/x) is not directly contradictory. Thus z_i and x are related in the same way to z. So deleting all conjuncts which contain x from $\gamma^0(z_1,\ldots,z_k,x,y,\ldots,z)$ gives the same result as substituting z_i for x in $\gamma^0(z_1,\ldots,z_k,x,y,\ldots,z)$ and then omitting repetitions of conjuncts. So substituting z_i for x in the subforest generated by the sibling-set of level 2 in $\gamma_b^{d-1}(z_1,\ldots,z_k,x)$ has the same effect as the first step in obtaining the contracted copy of this subforest. So, if after substituting z_i for x in this subforest, we substitute x for y and then change the bound variables to be used in their order (starting with y), we get the contracted copy of this subforest. □

The following lemma expresses the effect of omitting a layer of quantifiers in terms of the tree-structure. And the following theorem expresses the condition regarding the omission of layers of quantifiers that is part of the definition of one-trivial inconsistency in terms of the tree-structure.

Lemma 4.11 *In a constituent or attributive constituent X of depth $d \geq 1$ with free individual terms $z_1,\ldots,z_k$, the formula that results from omitting the nth layer of quantifiers from X (for $1 \leq n \leq d$) is the formula obtained from X by, for each sibling-set S of level $n+1$, replacing the subforest generated by S by the conjunction of the contracted copies of all the subforests generated by the successor-sets of S.*

PROOF Since omitting the nth layer of quantifiers is omitting the first layer of quantifiers in each subtree generated by a node of level n (which is immediate from the definition of omitting a layer of quantifiers), we need only show that the omission of the first layer of quantifiers from X is the formula obtained from X by, for the sibling-set S of level 2, replacing the subforest generated by S by the conjunction of the contracted copies of the subforests generated by the sibling-sets of level 3.

The constituent or attributive constituent X has the following form (if X is a constituent, the initial $\gamma^0(z_1, \ldots, z_k)$ is replaced by a $\delta^0(z_1, \ldots, z_k)$):

$$
\begin{array}{l}
\gamma^0(z_1, \ldots, z_k)\wedge \\
\bigwedge_{a\in A}(\exists z_{k+1})[\gamma^0(z_1, \ldots, z_{k+1})\wedge \\
\qquad \bigwedge_{b\in B}(\exists z_{k+2})[\gamma^0(z_1, \ldots, z_{k+2})\wedge \\
\qquad\qquad \bigwedge_{c\in C}(\exists z_{k+3})[\gamma^0(z_1, \ldots, z_{k+3})\wedge \\
\qquad\qquad\qquad \ldots \\
\qquad\qquad\qquad \bigwedge_{e\in E}(\exists z_{k+d-1})[\gamma^0(z_1, \ldots, z_{k+d-1})\wedge \\
\qquad\qquad\qquad\qquad \bigwedge_{f\in F}(\exists z_{k+d})\gamma^0_f(z_1, \ldots, z_{k+d})\wedge \\
\qquad\qquad\qquad\qquad (\forall z_{k+d}) \bigvee_{f\in F} \gamma^0_f(z_1, \ldots, z_{k+d})]_e\wedge \\
\qquad\qquad\qquad (\forall z_{k+d-1}) \bigvee_{e\in E}[\ ]_e\wedge \\
\qquad\qquad\qquad \ldots]_c\wedge \\
\qquad\qquad (\forall z_{k+3}) \bigvee_{c\in C}[\ ]_c\,]_b\wedge \\
\qquad (\forall z_{k+2}) \bigvee_{b\in B}[\ ]_b\,]_a\wedge \\
(\forall z_{k+1}) \bigvee_{a\in A}[\ ]_a
\end{array}
$$

where each empty pair of brackets [ ] with some subscript is an abbreviation for the contents of the first occurrence of brackets with the same subscript. Omitting the first layer of quantifiers, we get

$$
\begin{array}{l}
\gamma^0(z_1,\ldots,z_k)\wedge \\
\bigwedge_{a\in A}[\bigwedge_{b\in B}(\exists z_{k+1})[\gamma^0(z_1,\ldots,z_{k+1})\wedge \\
\qquad\qquad \bigwedge_{c\in C}(\exists z_{k+2})[\gamma^0(z_1,\ldots,z_{k+2})\wedge \\
\qquad\qquad\qquad\qquad \ldots \\
\qquad\qquad\qquad\qquad \bigwedge_{e\in E}(\exists z_{k+d-2})[\gamma^0(z_1,\ldots,z_{k+d-2})\wedge \\
\qquad\qquad\qquad\qquad\qquad\qquad \bigwedge_{f\in F}(\exists z_{k+d-1})\gamma^0_f(z_1,\ldots,z_{k+d-1})\wedge \\
\qquad\qquad\qquad\qquad\qquad\qquad (\forall z_{k+d-1})\bigvee_{f\in F}\gamma^0_f(z_1,\ldots,z_{k+d-1})]_e\wedge \\
\qquad\qquad\qquad\qquad (\forall z_{k+d-1})\bigvee_{e\in E}[\;]_e\wedge \\
\qquad\qquad\qquad\qquad \ldots]_c\wedge \\
\qquad\qquad (\forall z_{k+2})\bigvee_{c\in C}[\;]_c\,]_b\wedge \\
(\forall z_{k+1})\bigvee_{b\in B}[\;]_b\,]_a
\end{array}
$$

where $\gamma^0(z_1,\ldots,z_{k+1}),\gamma^0(z_1,\ldots,z_{k+2}),\ldots$ in the above formula are the results of omitting the atomic formulas containing z_{k+1} from $\gamma^0(z_1,\ldots,z_{k+2}),\gamma^0(z_1,\ldots,z_{k+3}),\ldots$ of the previous formula, and then changing the bound variables from $z_{k+2},\ldots,z_{k+d}$ to $z_{k+1},\ldots,z_{k+d-1}$ throughout the whole formula. Thus each $[\;]_c$ in the above formula is the contracted copy of $[\;]_c$ of the previous formula as required. □

For those trees that represent constituents and attributive constituents, by the *depth* of a principal subtree, I mean the depth of the formula which is represented by that subtree. Thus, such subtrees which have n levels are of depth $n-1$.

Theorem 4.12 *For a constituent or attributive constituent X of depth $d \geq 1$, the following are equivalent*

(1.) for each principal subtree Y of X with depth ≥ 1, omitting each different layer of quantifiers in Y gives the same result

(2.) for each sibling-set S in X, for each successor-set T of S, the contracted copy of the subforest generated by T is the top-$(d-i+1)$ forest of the subforest generated by S (where i is the level of S).

Proof If (1.) then omitting each different layer of quantifiers in X gives the same result, so omitting the first layer of quantifiers gives the same result as omitting the last layer which is the top-d tree of X. So by lemma 4.11, the formula obtained from X by replacing the subforest generated by the sibling-set S of level 2 by the conjunction of the contracted copies of the subforests generated by the successor-sets of S is the top-d tree of X. So each such contracted copy is the top-$(d-1)$ forest of the subforest generated by S. Similarly, since for each principal subtree Y of X with depth ≥ 1, omitting the first and last layers of quantifiers give the same result, for each successor-set T of the sibling-set S of level 2 in Y, the contracted copy of the subforest generated by T is the top-n forest of the subforest generated by S where n is the number of levels in the subforest generated by T. Each sibling-set is the sibling-set of level 2 in some principal subtree. Thus (2.).

If (2.) then for the sibling-set S of level 2 in X, for each successor-set T of S, the contracted copy of the subforest generated by T is the top-$(d-1)$ forest of the subforest generated by S. So by lemma 4.11, omitting the first layer of quantifiers in X gives the top-d tree of X which is the result of omitting the last layer of quantifiers from X. Similarly, in each principal subtree Y of X, omitting the first and last layers of quantifiers give the same result. To show that omitting the other layers of quantifiers from X also give the top-d tree of X: To omit the second layer of quantifiers in X is to omit the first layer in each subtree Y generated by a node of level 2 in X. Since, for each such Y, this gives the top-$(d-1)$ tree of Y, the result of omitting the result of omitting the second layer of quantifiers in X is the top-d tree of X. Generally, omitting the nth layer of quantifiers, for $n \geq 2$, in X is to omit the first layer in each subtree generated by a node of level n, and results in the top-d tree of X. Similarly, in each principal subtree Y of X, omitting each different layer of quantifiers in Y gives the same result. Thus (1.).

The structure of a constituent or attributive constituent without any free individual terms is like that of X above, except without a root. The above proof does not involve the root of X, so it also holds for constituents of the form $\delta^d()$. □

Corollary 4.13 *A constituent or attributive constituent of depth d is one-trivially inconsistent if it does not satisfy the following conditions*

- *each node N which has successors has, for each free individual term z in N, some branch which is a partial expansion of N to depth n with respect to z, where n is the depth of the branches of N*
- *for each sibling-set S of level i, for each successor-set T of S, the contracted copy of the subforest generated by T is the top-$(d-i+1)$ forest of the subforest generated by S.*

PROOF Apply theorem 4.12 to the condition regarding the omission of layers of quantifiers in the definition of one-trivial inconsistency. □

Although the above formulation of one-trivial inconsistency might be slightly longer to state than the original formulation, it provides a graphical way of understanding the definition by showing how the different parts of a constituent must be related to each other for it not to be one-trivially inconsistent.

4.6 Another definition of trivial inconsistency

I have not been able to prove the completeness result Hintikka claims for one-trivial inconsistency. In this section, I give another definition of trivial inconsistency for which a similar completeness result does hold (and is proved in the next section).

Since there are different possible ways of defining trivial inconsistency, any definition of some other concept in terms of trivial inconsistency is relative to the particular definition of trivial inconsistency used. Discussing the significance of definitions which depend on the definition of trivial inconsistency, Hintikka [1965a] (p. 88) says: "That they depend on the conditions employed is true; but on the other hand it seems to me that in some rather elusive sense our conditions (A)–(E) of inconsistency are as strong as we can possibly hope natural conditions to be."

For one of the conditions referred to above (called (C)), there is a stronger natural condition which is also necessary for consistency.

The following condition is one of those used in the definition of one-trivial inconsistency:

- Each node N which has successors has, for each free individual term z in N, some branch which is a partial expansion of N with respect to z.

It can be strengthened to:

- Each node N which has successors has, for each free individual term z in N, some branch which is a partial expansion of *the conjunction of N and all the nodes that are above it* with respect to z.

We show that this is a necessary condition for consistency as follows:

An attributive constituent

$$\gamma^d(z_1,\ldots,z_k) \;=\; \gamma^0(z_1,\ldots,z_k)\wedge \bigwedge_{i\in I}(\exists x)\gamma_i^{d-1}(z_1,\ldots,z_k,x)\wedge(\forall x)\bigvee_{i\in I}\gamma_i^{d-1}(z_1,\ldots,z_k,x)$$

is equivalent to

$$\gamma^0(z_1,\ldots,z_k)\wedge \bigwedge_{i\in I}(\exists x)(\gamma^0(z_1,\ldots,z_k)\wedge\gamma_i^{d-1}(z_1,\ldots,z_k,x))\wedge(\forall x)\bigvee_{i\in I}(\gamma^0(z_1,\ldots,z_k)\wedge\gamma_i^{d-1}(z_1,\ldots,z_k,x)).$$

So $\gamma^d(z_1,\ldots,z_k)$ is equivalent to the formula obtained from it by, for each attributive constituent $\gamma^{d-j}(z_1,\ldots,z_{k+j})$ which occurs at any level of nesting in $\gamma^d(z_1,\ldots,z_k)$, replacing $\gamma^{d-j}(z_1,\ldots,z_{k+j})$ by $\gamma^0(z_1,\ldots,z_k) \wedge \gamma^0(z_1,\ldots,z_{k+1}) \wedge \ldots \wedge \gamma^0(z_1,\ldots,z_{k+j-1}) \wedge \gamma^{d-j}(z_1,\ldots,z_{k+j})$ where $\gamma^0(z_1,\ldots,z_k),\ldots,\gamma^0(z_1,\ldots,z_{k+j-1})$ are those nodes which occur above the root of $\gamma^{d-j}(z_1,\ldots,z_{k+j})$. Let the formula so obtained from $\gamma^d(z_1,\ldots,z_k)$ be called $\Gamma^d(z_1,\ldots,z_k)$. Now, for $\Gamma^d(z_1,\ldots,z_k)$ to be consistent, the subtrees generated by each node must all be consistent. These subtrees are of a form which is equivalent to

$$\gamma^0(z_1,\ldots,z_k) \wedge \ldots \wedge \gamma^0(z_1,\ldots,z_{k+j}) \wedge$$
$$\bigwedge_{i\in J}(\exists z_{k+j+1})\gamma_i^{d-j-1}(z_1,\ldots,z_{k+j+1}) \wedge (\forall z_{k+j+1}) \bigvee_{i\in J} \gamma_i^{d-j-1}(z_1,\ldots,z_{k+j+1})$$

which implies

$$\gamma^0(z_1,\ldots,z_k) \wedge \ldots \wedge \gamma^0(z_1,\ldots,z_{k+j}) \wedge \bigvee_{i\in J}(\gamma_i^{d-j-1}(z_1,\ldots,z_{k+j+1})(z_l/z_{k+j+1}))$$

for each $z_l \in \{z_1,\ldots,z_{k+j}\}$. And

$$\gamma^0(z_1,\ldots,z_k) \wedge \ldots \wedge \gamma^0(z_1,\ldots,z_{k+j}) \wedge \bigvee_{i\in J}(\gamma_i^{d-j-1}(z_1,\ldots,z_{k+j+1})(z_l/z_{k+j+1}))$$

can only be consistent if there is at least one $i \in J$ for which

$$\gamma^0(z_1,\ldots,z_k) \wedge \ldots \wedge \gamma^0(z_1,\ldots,z_{k+j}) \wedge \gamma_i^{d-j-1}(z_1,\ldots,z_{k+j+1})(z_l/z_{k+j+1})$$

is not directly contradictory. That is, $\gamma_i^{d-j-1}(z_1,\ldots,z_{k+j+1})$ is a partial expansion of $\gamma^0(z_1,\ldots,z_k) \wedge \ldots \wedge \gamma^0(z_1,\ldots,z_{k+j})$ with respect to z_l.

A similar argument holds for constituents.

We now show that the new condition is stronger than the old one.

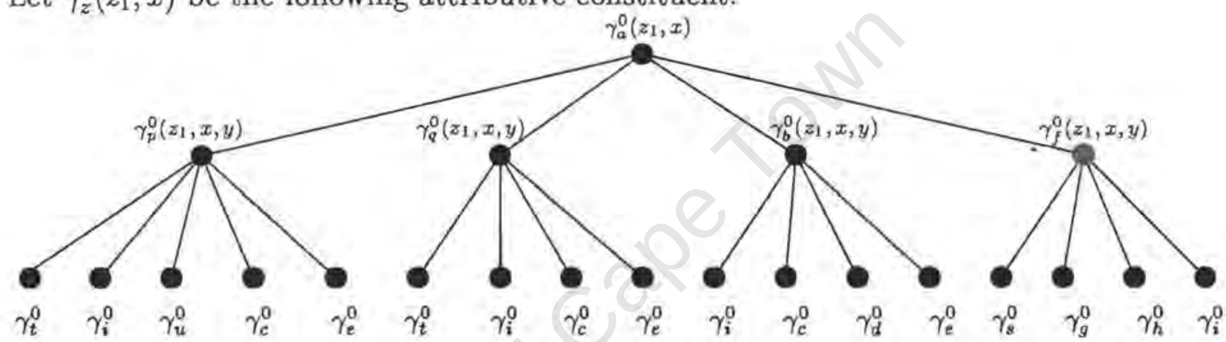
The following example is of an attributive constituent of depth 2 in which each node N with successors has, for each free individual term z in N, a branch which is a partial expansion of N with respect to z, but it does not satisfy the stronger condition.

Let

$$\begin{aligned}
\gamma_a^0(z_1,x) &= P(z_1,x) \wedge \neg P(x,z_1) \wedge P(x,x)\\
\gamma_b^0(z_1,x,y) &= P(z_1,y) \wedge \neg P(y,z_1) \wedge P(x,y) \wedge P(y,x) \wedge P(y,y)\\
\gamma_c^0(z_1,x,y,w) &= P(z_1,w) \wedge \neg P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge P(y,w) \wedge P(w,y) \wedge P(w,w)\\
\gamma_d^0(z_1,x,y,w) &= \neg P(z_1,w) \wedge \neg P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge P(y,w) \wedge P(w,y) \wedge P(w,w)\\
\gamma_e^0(z_1,x,y,w) &= P(z_1,w) \wedge P(w,z_1) \wedge \neg P(x,w) \wedge P(w,x) \wedge \neg P(y,w) \wedge P(w,y) \wedge P(w,w)\\
\gamma_f^0(z_1,x,y) &= P(z_1,y) \wedge P(y,z_1) \wedge \neg P(x,y) \wedge P(y,x) \wedge P(y,y)
\end{aligned}$$

$$
\begin{aligned}
\gamma_g^0(z_1,x,y,w) &= P(z_1,w) \wedge P(w,z_1) \wedge \neg P(x,w) \wedge P(w,x) \wedge P(y,w) \wedge P(w,y) \wedge P(w,w) \\
\gamma_h^0(z_1,x,y,w) &= P(z_1,w) \wedge \neg P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge P(y,w) \wedge \neg P(w,y) \wedge P(w,w) \\
\gamma_i^0(z_1,x,y,w) &= P(z_1,w) \wedge P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge P(y,w) \wedge P(w,y) \wedge P(w,w) \\
\gamma_p^0(z_1,x,y) &= \neg P(z_1,y) \wedge \neg P(y,z_1) \wedge P(x,y) \wedge P(y,x) \wedge P(y,y) \\
\gamma_q^0(z_1,x,y) &= P(z_1,y) \wedge P(y,z_1) \wedge P(x,y) \wedge P(y,x) \wedge P(y,y) \\
\gamma_s^0(z_1,x,y,w) &= \neg P(z_1,w) \wedge \neg P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge \neg P(y,w) \wedge \neg P(w,y) \wedge P(w,w) \\
\gamma_t^0(z_1,x,y,w) &= \neg P(z_1,w) \wedge \neg P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge P(y,w) \wedge P(w,y) \wedge P(w,w) \\
\gamma_u^0(z_1,x,y,w) &= P(z_1,w) \wedge P(w,z_1) \wedge P(x,w) \wedge P(w,x) \wedge \neg P(y,w) \wedge \neg P(w,y) \wedge P(w,w).
\end{aligned}
$$

Let $\gamma_z^2(z_1, x)$ be the following attributive constituent:



where the labels of the bottom layer of nodes are written without their parameters. In $\gamma_z^2(z_1, x)$, each node N with successors has, for each of its free individual terms z, a branch which is a partial expansion of N with respect to z:

- the subtree generated by $\gamma_b^0(z_1, x, y)$ is a partial expansion of $\gamma_a^0(z_1, x)$ with respect to x
- the subtree generated by $\gamma_f^0(z_1, x, y)$ is a partial expansion of $\gamma_a^0(z_1, x)$ with respect to z_1
- $\gamma_t^0(z_1, x, y, w)$ is a partial expansion of $\gamma_p^0(z_1, x, y)$ with respect to y
- $\gamma_i^0(z_1, x, y, w)$ is a partial expansion of $\gamma_p^0(z_1, x, y)$ with respect to x
- $\gamma_u^0(z_1, x, y, w)$ is a partial expansion of $\gamma_p^0(z_1, x, y)$ with respect to z_1
- $\gamma_i^0(z_1, x, y, w)$ is a partial expansion of $\gamma_q^0(z_1, x, y)$ with respect to y
- $\gamma_i^0(z_1, x, y, w)$ is a partial expansion of $\gamma_q^0(z_1, x, y)$ with respect to x
- $\gamma_i^0(z_1, x, y, w)$ is a partial expansion of $\gamma_q^0(z_1, x, y)$ with respect to z_1
- $\gamma_c^0(z_1, x, y, w)$ is a partial expansion of $\gamma_b^0(z_1, x, y)$ with respect to y
- $\gamma_d^0(z_1, x, y, w)$ is a partial expansion of $\gamma_b^0(z_1, x, y)$ with respect to x

- $\gamma_e^0(z_1, x, y, w)$ is a partial expansion of $\gamma_b^0(z_1, x, y)$ with respect to z_1
- $\gamma_g^0(z_1, x, y, w)$ is a partial expansion of $\gamma_f^0(z_1, x, y)$ with respect to y
- $\gamma_h^0(z_1, x, y, w)$ is a partial expansion of $\gamma_f^0(z_1, x, y)$ with respect to x
- $\gamma_i^0(z_1, x, y, w)$ is a partial expansion of $\gamma_f^0(z_1, x, y)$ with respect to z_1.

But $\gamma_p^0(z_1, x, y)$ does not have any branch which is a partial expansion of $\gamma_a^0(z_1, x) \wedge \gamma_p^0(z_1, x, y)$ with respect to z_1 since:

$$\begin{aligned}
\gamma_a^0(z_1, x) \wedge \gamma_p^0(z_1, x, y) &= P(z_1, x) \wedge \neg P(x, z_1) \wedge P(x, x) \wedge \\
&\quad \neg P(z_1, y) \wedge \neg P(y, z_1) \wedge P(x, y) \wedge P(y, x) \wedge P(y, y) \\
\gamma_t^0(z_1, x, y, w)(z_1/w) &= \neg P(z_1, z_1) \wedge \neg P(z_1, z_1) \wedge P(x, z_1) \wedge P(z_1, x) \wedge P(y, z_1) \wedge \\
&\quad P(z_1, y) \wedge P(z_1, z_1) \\
\gamma_i^0(z_1, x, y, w)(z_1/w) &= P(z_1, z_1) \wedge P(z_1, z_1) \wedge P(x, z_1) \wedge P(z_1, x) \wedge P(y, z_1) \wedge \\
&\quad P(z_1, y) \wedge P(z_1, z_1) \\
\gamma_u^0(z_1, x, y, w)(z_1/w) &= P(z_1, z_1) \wedge P(z_1, z_1) \wedge P(x, z_1) \wedge P(z_1, x) \wedge \neg P(y, z_1) \wedge \\
&\quad \neg P(z_1, y) \wedge P(z_1, z_1) \\
\gamma_c^0(z_1, x, y, w)(z_1/w) &= P(z_1, z_1) \wedge \neg P(z_1, z_1) \wedge P(x, z_1) \wedge P(z_1, x) \wedge P(y, z_1) \wedge \\
&\quad P(z_1, y) \wedge P(z_1, z_1) \\
\gamma_e^0(z_1, x, y, w)(z_1/w) &= P(z_1, z_1) \wedge P(z_1, z_1) \wedge \neg P(x, z_1) \wedge P(z_1, x) \wedge \neg P(y, z_1) \wedge \\
&\quad P(z_1, y) \wedge P(z_1, z_1).
\end{aligned}$$

Having seen that the new condition is necessary for consistency and is stronger than the old condition, we now use it in a definition of trivial inconsistency, which I will call *two-trivial inconsistency.*

A constituent or attributive constituent of depth d is *two-trivially inconsistent* if it does not satisfy the following conditions

- each node N which has successors has, for each free individual term z in N, some branch which is a partial expansion of the conjunction of N and all the nodes that are above it with respect to z,
- in each principal subtree, omitting the different layers of quantifiers all give the same result.

The second condition in the above definition is the same as that in the definition of one-trivial inconsistency. Expressing it in terms of the tree-structure of constituents and attributive constituents, we get the following equivalent (shown by theorem 4.12) formulation of two-trivial inconsistency:

A constituent or attributive constituent of depth d is two-trivially inconsistent if it does not satisfy the following conditions

- each node N which has successors has, for each free individual term z in N, some branch which is a partial expansion of the conjunction of N and all the nodes that are above it with respect to z,
- for each sibling-set S of level i, for each successor-set T of S, the contracted copy of the subforest generated by T is the top-$(d-i+1)$ forest of the subforest generated by S.

That one-trivial inconsistency and two-trivial inconsistency are not equivalent is shown by the above example. To show that $\gamma_z^2(z_1,x)$ is not one-trivially inconsistent, in addition to the existence of partial expansions shown above, we have the following:

- the contracted copy of $\{\gamma_t^0(z_1,x,y,w), \gamma_i^0(z_1,x,y,w), \gamma_u^0(z_1,x,y,w), \gamma_c^0(z_1,x,y,w), \gamma_e^0(z_1,x,y,w)\}$ is $\{\gamma_p^0(z_1,x,y), \gamma_q^0(z_1,x,y), \gamma_b^0(z_1,x,y), \gamma_f^0(z_1,x,y)\}$
- the contracted copy of $\{\gamma_t^0(z_1,x,y,w), \gamma_i^0(z_1,x,y,w), \gamma_c^0(z_1,x,y,w), \gamma_e^0(z_1,x,y,w)\}$ is $\{\gamma_p^0(z_1,x,y), \gamma_q^0(z_1,x,y), \gamma_b^0(z_1,x,y), \gamma_f^0(z_1,x,y)\}$
- the contracted copy of $\{\gamma_i^0(z_1,x,y,w), \gamma_c^0(z_1,x,y,w), \gamma_d^0(z_1,x,y,w), \gamma_e^0(z_1,x,y,w)\}$ is $\{\gamma_p^0(z_1,x,y), \gamma_q^0(z_1,x,y), \gamma_b^0(z_1,x,y), \gamma_f^0(z_1,x,y)\}$
- the contracted copy of $\{\gamma_s^0(z_1,x,y,w), \gamma_g^0(z_1,x,y,w), \gamma_h^0(z_1,x,y,w), \gamma_i^0(z_1,x,y,w)\}$ is $\{\gamma_p^0(z_1,x,y), \gamma_q^0(z_1,x,y), \gamma_b^0(z_1,x,y), \gamma_f^0(z_1,x,y)\}$.

But $\gamma_z^2(z_1,x)$ is two-trivially inconsistent because, as we have seen above, $\gamma_p^0(z_1,x,y)$ does not have any branch which is a partial expansion of $\gamma_a^0(z_1,x) \wedge \gamma_p^0(z_1,x,y)$ with respect to z_1.

The above example shows that an attributive constituent can be two-trivially inconsistent without being one-trivially inconsistent. From the point of view of the completeness theorem, we would like to know whether there is an infinite sequence of constituents where the successor of each element is in its expansion such that no element of the sequence is one-trivially inconsistent, but some element is two-trivially inconsistent. I don't know if there is such a sequence. If there is, it would provide a counterexample to the completeness result Hintikka claims.

4.7 The completeness theorem of the theory of distributive normal forms

As for the rest of this chapter, this section deals only with languages without function symbols or identity, though they may contain constant symbols. We first give some definitions.

A *disproof method* for first-order logic is a method of proving the inconsistency of formulas (as opposed to a proof method which proves the validity of formulas). A disproof method is *complete* if for every inconsistent formula, it can show the formula to be inconsistent. As we will later see, from any complete disproof method for first-order logic, we can obtain a complete proof method. And from any complete proof method, we can obtain a complete disproof method.

A disjunction $\bigvee_{a\in A} \delta_a^d(z_1,\ldots,z_k)$ is said to be *trivially inconsistent* if each disjunct $\delta_a^d(z_1,\ldots,z_k)$ is trivially inconsistent.

Every formula can be expressed as a disjunction of constituents. It is inconsistent iff each disjunct in this disjunction is inconsistent. If any constituent $\delta^d(z_1,\ldots,z_k)$ is trivially inconsistent, then it is inconsistent. Otherwise, we can expand $\delta^d(z_1,\ldots,z_k)$ to some disjunction $\bigvee_{a\in A} \delta_a^{d+1}(z_1,\ldots,z_k)$. If each disjunct $\delta_a^{d+1}(z_1,\ldots,z_k)$ is trivially inconsistent then $\delta^d(z_1,\ldots,z_k)$ is inconsistent. If not, then we can omit the trivially inconsistent disjuncts $\delta_a^{d+1}(z_1,\ldots,z_k)$ and expand those which are not trivially inconsistent to depth $d+2$. In this way we get a disproof method for first-order logic which can be summarized as follows:

Given a formula F of depth d with free individual terms $z_1,\ldots,z_k$,

- convert F to distributive normal form $\bigvee_{a\in A} \delta_a^d(z_1,\ldots,z_k)$
- repeat:
    - omit all trivially inconsistent constituents
    - if no constituents remain, stop (since F has been found to be inconsistent)
    - expand to depth 1-greater than current depth.

The completeness theorem of the theory of distributive normal forms states that for each inconsistent formula F of depth d, there is some depth $d+e$ at which the expansion of F is trivially inconsistent, and thus that the above disproof method is complete.

To prove this completeness theorem we make use of *model sets*. In this context it is convenient to use a definition of *formula* which is slightly different from the one given earlier in that the symbols taken as primitive and defined are different; but it is equivalent in the sense that, including abbreviations, all the formulas are the same and have the same semantics. Every formula is equivalent to one in which all negation signs immediately precede an atomic formula. Such an equivalent formula can be found by replacing any formula of the form of one of the formulas on the left-hand side of the following equivalences by its equivalent formula on the right-hand side:

- $\neg(\exists x)A \iff (\forall x)\neg A$
- $\neg(\forall x)A \iff (\exists x)\neg A$
- $\neg(A \wedge B) \iff \neg A \vee \neg B$
- $\neg(A \vee B) \iff \neg A \wedge \neg B$
- $\neg\neg A \iff A$

Thus it is possible to consider all negations of non-atomic formulas as abbreviations of formulas in which the only negations are of atomic formulas. However, this makes it necessary to consider both quantifiers as primitive, and also both logical connectives $\wedge$ and $\vee$ as primitive. This way of defining formulas is used in the definition of model sets. The *sentences* obtained in this way can be defined explicitly (Hintikka [1955], p. 53) by:

- An atomic formula over the set of constants is a sentence, called an *atomic sentence*.
- If A is an atomic sentence then $\neg A$ is a sentence.
- If B and C are sentences then $B \wedge C$ and $B \vee C$ are sentences.
- If B is a sentence containing the constant a but not the variable x then $(\exists x)B(x/a)$ and $(\forall x)B(x/a)$ are sentences, where $B(x/a)$ is the result of substituting x for each occurrence of a in B.

All other sentences can be considered as abbreviations of a sentence by the above definition (by the equivalences above). The more general form of the notation for substitution used in the above definition is: If A is a formula and each of u and v is either a constant or a variable, then $A(u/v)$ denotes the formula obtained by substituting u for each occurrence of v in A.

Note that in constituents and attributive constituents of the second kind, the only negations are of atomic formulas.

induction containing of the $B(a/x)$ for into In order to use induction on the formation of sentences, all the sentences of the form $B(a/x)$ are considered as *subsentences* of $(\exists x)B$ and of $(\forall x)B$ (following the use of *subformulas* in Hintikka [1955]). The full definition of subsentence is:

- The sentence $\neg A$ has the subsentence A.

- The sentences $B \wedge C$ and $B \vee C$ have the subsentences B and C.
- The sentences $(\exists x)B$ and $(\forall x)B$ have the subsentences $B(a/x)$ for each constant a.
- A subsentence of any subsentence of A is a subsentence of A.

We will be interested in a certain kind of model, defined in Hintikka [1955], which I will call a *model with names*. A *model with names* for a language $\mathcal{L}$ is a triple $\mathcal{M} = \langle D, I_1, I_2 \rangle$ where D is a set whose elements are called individuals, $I_1 : \mathbf{Const}_{\mathcal{L}} \to D$ is a surjection, and I_2 is a function that maps, for each $n \in \mathbb{N}$, every n-ary predicate symbol of $\mathcal{L}$ to some n-ary relation on D. If we let $I = I_1 \cup I_2$ then $\langle D, I \rangle$ is a model as previously defined. So this definition can be considered as the special case of the previous one for which each individual e has some constant which is mapped by the interpretation to e. (There may be more than one constant which is mapped to some given individual.) A constant a such that $I_1(a) = e$ may be considered as a "name" for e, which is the reason for choosing the terminology "model with names".

The truth values of all sentences are determined by a model with names $\mathcal{M} = \langle D, I_1, I_2 \rangle$ by the relation $\models_*$ defined by:

- $\mathcal{M} \models_* P(a_1, \ldots, a_n)$ iff $\langle I_1(a_1), \ldots, I_1(a_n) \rangle \in I_2(P)$ for an atomic sentence $P(a_1, \ldots, a_n)$
- $\mathcal{M} \models_* \neg A$ iff $\mathcal{M} \not\models_* A$ for an atomic sentence A
- $\mathcal{M} \models_* A \wedge B$ iff $\mathcal{M} \models_* A$ and $\mathcal{M} \models_* B$
- $\mathcal{M} \models_* A \vee B$ iff $\mathcal{M} \models_* A$ or $\mathcal{M} \models_* B$
- $\mathcal{M} \models_* (\exists x)A$ iff $\mathcal{M} \models_* A(a/x)$ for some constant a
- $\mathcal{M} \models_* (\forall x)A$ iff $\mathcal{M} \models_* A(a/x)$ for each constant a

If $\mathcal{M} \models_* A$ then A is called *true in* $\mathcal{M}$, otherwise A is called *false in* $\mathcal{M}$.

The above definition is equivalent to the earlier definition of the truth-value of a formula in a model under a valuation for the special case of formulas that are sentences in models with names. The only difference, other than which symbols are taken as primitive in defining formulas, is that here the definition for sentences containing quantifiers uses substitution in the sentence rather than in a valuation. This can only be done because there is a "name" for each individual. In particular, if $(\exists x)A$ is a sentence, then $(\langle D, I_1, I_2 \rangle \models_* A(a/x)$ for some constant $a)$ iff $(\langle D, I_1 \cup I_2 \rangle, v \models A$ for some $v : \{x\} \to D)$, since if $\langle D, I_1, I_2 \rangle \models_* A(a/x)$ then put $v(x) = I_1(a)$, then $\langle D, I_1 \cup I_2 \rangle, v \models A$; and if $\langle D, I_1 \cup I_2 \rangle, v \models A$ then there is some constant a such that $I_1(a) = v(x)$ because I_1 is a surjection, so for this a, $\langle D, I_1, I_2 \rangle \models_* A(a/x)$. Similarly, $(\langle D, I_1, I_2 \rangle \models_* A(a/x)$ for each constant $a)$ iff $(\langle D, I_1 \cup I_2 \rangle, v \models A$ for each $v : \{x\} \to D)$.

A set $\mathbf{F}$ of sentences which satisfies the conditions

1. if $A \in \mathbf{F}$ then $\neg A \notin \mathbf{F}$
2. if $(A \wedge B) \in \mathbf{F}$ then $A \in \mathbf{F}$ and $B \in \mathbf{F}$
3. if $(A \vee B) \in \mathbf{F}$ then $A \in \mathbf{F}$ or $B \in \mathbf{F}$
4. if $(\exists x)A \in \mathbf{F}$ then $A(a/x) \in \mathbf{F}$ for some constant a
5. if $(\forall x)A \in \mathbf{F}$ then $A(a/x) \in \mathbf{F}$ for each constant a which occurs in any of the formulas in $\mathbf{F}$

is called a *model set of sentences* (Hintikka [1955]). The above condition 1. is called the *consistency condition* of model sets, and the conditions 2. to 5. are called the *closure conditions* of model sets.

Lemma 4.14 *For any model with names $\mathcal{M} = \langle D, I_1, I_2\rangle$ for a language $\mathcal{L}$ and any $a, b \in$ $\mathbf{Const}_{\mathcal{L}}$ and A a formula of $\mathcal{L}$ whose only free variable is x, if $I_1(a) = I_1(b)$ then $\mathcal{M} \models_* A(a/x)$ iff $\mathcal{M} \models_* A(b/x)$.*

PROOF Let A be a formula whose only free variable is x. Then A is of the form $B(x/c)$ for some sentence B containing some constant c. And $A(a/x)$ is $B(a/c)$ and $A(b/x)$ is $B(b/c)$. So we need to show that for each sentence B containing the constant c, $\mathcal{M} \models_* B(a/c)$ iff $\mathcal{M} \models_* B(b/c)$. We use induction on the formation of the sentence B. If B is $P(a_1, \ldots, a_n)$ where some of the $a_1, \ldots, a_n$ are c, then $B(a/c)$ is $P(b_1, \ldots, b_n)$ where b_i is a if a_i is c and a_i otherwise. And $B(b/c)$ is $P(c_1, \ldots, c_n)$ where c_i is b if a_i is c and a_i otherwise.

$$\begin{aligned} \mathcal{M} \models_* P(b_1, \ldots, b_n) \quad &\text{iff} \quad \langle I_1(b_1), \ldots, I_1(b_n)\rangle \in I_2(P) \\ &\text{iff} \quad \langle I_1(c_1), \ldots, I_1(c_n)\rangle \in I_2(P) \text{ since for each } i,\ I_1(b_i) = I_1(c_i) \\ &\text{iff} \quad \mathcal{M} \models_* P(c_1, \ldots, c_n). \end{aligned}$$

Thus $\mathcal{M} \models_* B(a/c)$ iff $\mathcal{M} \models_* B(b/c)$. If B is $\neg P(a_1, \ldots, a_n)$, then for b_i and c_i as above,

$$\begin{aligned} \mathcal{M} \models_* \neg P(b_1, \ldots, b_n) \quad &\text{iff} \quad \langle I_1(b_1), \ldots, I_1(b_n)\rangle \notin I_2(P) \\ &\text{iff} \quad \langle I_1(c_1), \ldots, I_1(c_n)\rangle \notin I_2(P) \text{ since for each } i,\ I_1(b_i) = I_1(c_i) \\ &\text{iff} \quad \mathcal{M} \models_* \neg P(c_1, \ldots, c_n). \end{aligned}$$

Thus $\mathcal{M} \models_* B(a/c)$ iff $\mathcal{M} \models_* B(b/c)$. Assume that for each subsentence C of B that contains c, $\mathcal{M} \models_* C(a/c)$ iff $\mathcal{M} \models_* C(b/c)$. If B is $D_1 \wedge D_2$ then at least one D_i ($i \in \{1, 2\}$) contains c. $B(a/c)$ is $D_1(a/c) \wedge D_2(a/c)$ and $B(b/c)$ is $D_1(b/c) \wedge D_2(b/c)$. For each D_i, if D_i contains c then by hypothesis $\mathcal{M} \models_* D_i(a/c)$ iff $\mathcal{M} \models_* D_i(b/c)$, and if D_i doesn't contain c then $D_i(a/c)$ is D_i is $D_i(b/c)$. Thus for each i, $\mathcal{M} \models_* D_i(a/c)$ iff $\mathcal{M} \models_* D_i(b/c)$. So $\mathcal{M} \models_* D_1(a/c) \wedge D_2(a/c)$ iff $\mathcal{M} \models_* D_1(b/c) \wedge D_2(b/c)$. So $\mathcal{M} \models_* B(a/c)$ iff $\mathcal{M} \models_* B(b/c)$. If B is $D_1 \vee D_2$, then for each i, $\mathcal{M} \models_* D_i(a/c)$ iff $\mathcal{M} \models_* D_i(b/c)$ (for the same reasons as above). So $\mathcal{M} \models_* D_1(a/c) \vee D_2(a/c)$ iff $\mathcal{M} \models_* D_1(b/c) \vee D_2(b/c)$. So $\mathcal{M} \models_* B(a/c)$ iff

$\mathcal{M} \models_* B(b/c)$. If B is $(\exists y)D(y/d)$ then D is a sentence containing distinct constants c and d. By hypothesis, $\mathcal{M} \models_* D(a/c)$ iff $\mathcal{M} \models_* D(b/c)$. Thus $\mathcal{M} \models_* (\exists y)D(a/c)(y/d)$ iff $\mathcal{M} \models_* (\exists y)D(b/c)(y/d)$. So, since $B(a/c)$ is $(\exists y)D(a/c)(y/d)$ and $B(b/c)$ is $(\exists y)D(b/c)(y/d)$, $\mathcal{M} \models_* B(a/c)$ iff $\mathcal{M} \models_* B(b/c)$. If B is $(\forall y)D(y/d)$ then by hypothesis $\mathcal{M} \models_* D(a/c)$ iff $\mathcal{M} \models_* D(b/c)$. Thus $\mathcal{M} \models_* (\forall y)D(a/c)(y/d)$ iff $\mathcal{M} \models_* (\forall y)D(b/c)(y/d)$. So, since $B(a/c)$ is $(\forall y)D(a/c)(y/d)$ and $B(b/c)$ is $(\forall y)D(b/c)(y/d)$, $\mathcal{M} \models_* B(a/c)$ iff $\mathcal{M} \models_* B(b/c)$. □

Theorem 4.15 (Hintikka [1955]) *Every model set of sentences is satisfiable.*

PROOF Let $\mathbf{F}$ be a model set of sentences of a language $\mathcal{L}$. We show that $\mathbf{F}$ is satisfied in a model with names. Let D be the set of all constants that occur in any of the sentences in $\mathbf{F}$. Let e be any element of D. Define $I_1 : \mathbf{Const}_{\mathcal{L}} \to D$ by

$$I_1(a) = \begin{cases} a & \text{for } a \in D \\ e & \text{for } a \notin D. \end{cases}$$

For each $n \in \mathbb{N}$, for each n-ary predicate symbol P, let $I_2(P) = \{\langle a_1, \ldots, a_n \rangle \mid P(a_1, \ldots, a_n) \in \mathbf{F}\}$. Then $\mathcal{M} = \langle D, I_1, I_2 \rangle$ is a model with names. To show that $\mathcal{M}$ satisfies each element of $\mathbf{F}$: If $P(a_1, \ldots, a_n) \in \mathbf{F}$ then $\langle a_1, \ldots, a_n \rangle \in I_2(P)$ by the definition of I_2, so $\langle I_1(a_1), \ldots, I_1(a_n) \rangle \in I_2(P)$ by the definition of I_1, so $\mathcal{M} \models_* P(a_1, \ldots, a_n)$ by the definition of $\models_*$. If $\neg P(a_1, \ldots, a_n) \in \mathbf{F}$ then $P(a_1, \ldots, a_n) \notin \mathbf{F}$ by the consistency condition of model sets, so $\langle a_1, \ldots, a_n \rangle \notin I_2(P)$ by the definition of I_2, so $\langle I_1(a_1), \ldots, I_1(a_n) \rangle \notin I_2(P)$ by the definition of I_1, so $\mathcal{M} \not\models_* P(a_1, \ldots, a_n)$ by the definition of $\models_*$, so $\mathcal{M} \models_* \neg P(a_1, \ldots, a_n)$ by the definition of $\models_*$. Assume that for any sentence A in $\mathbf{F}$, each subsentence of A which is also in $\mathbf{F}$ is true in $\mathcal{M}$. If $(A \wedge B) \in \mathbf{F}$ then $A \in \mathbf{F}$ and $B \in \mathbf{F}$ by a closure condition of model sets, so $\mathcal{M} \models_* A$ and $\mathcal{M} \models_* B$ by the induction hypothesis, so $\mathcal{M} \models_* A \wedge B$ by the definition of $\models_*$. If $(A \vee B) \in \mathbf{F}$ then $A \in \mathbf{F}$ or $B \in \mathbf{F}$ by a closure condition of model sets, so $\mathcal{M} \models_* A$ or $\mathcal{M} \models_* B$ by the induction hypothesis, so $\mathcal{M} \models_* A \vee B$ by the definition of $\models_*$. If $(\exists x)A \in \mathbf{F}$ then $A(a/x) \in \mathbf{F}$ for some constant a by a closure condition of models sets, so $\mathcal{M} \models_* A(a/x)$ by the induction hypothesis, so $\mathcal{M} \models_* (\exists x)A$ by the definition of $\models_*$. If $(\forall x)A \in \mathbf{F}$ then for each $a \in D$, $A(a/x) \in \mathbf{F}$ by a closure condition of models sets, so for each $a \in D$, $\mathcal{M} \models_* A(a/x)$ by the induction hypothesis. For each $a \in \mathbf{Const}_{\mathcal{L}} \setminus D$, $I_1(a) = e$, and $e \in D$ with $I_1(e) = e$ (by the definition of D), so $\mathcal{M} \models_* A(e/x)$, so by lemma 4.14 $\mathcal{M} \models_* A(a/x)$ for each $a \in \mathbf{Const}_{\mathcal{L}} \setminus D$. So $\mathcal{M} \models_* A(a/x)$ for each $a \in \mathbf{Const}_{\mathcal{L}}$, so $\mathcal{M} \models_* (\forall x)A$. □

In Hintikka [1965a], model sets are defined slightly differently to above. Their elements may contain free variables, but for convenience, the same variable will not be used both free (in some element) and bound (in some element) in the same model set. An individual term is said to occur free in a set $\mathbf{X}$ of formulas if it occurs free in some element of $\mathbf{X}$, and occurs bound in $\mathbf{X}$ if it occurs bound in some element of $\mathbf{X}$. For proving the completeness theorem, we need to embed attributive constituents containing free variables into model sets, so we use this modified definition of model set (from Hintikka [1965a], p. 80, 82):

A set $\mathbf{X}$ of formulas which satisfies the conditions

1. if $A \in \mathbf{X}$ then $\neg A \notin \mathbf{X}$
2. if $(F_1 \wedge \ldots \wedge F_n) \in \mathbf{X}$ then for each $i \in \{1, \ldots, n\}$, $F_i \in \mathbf{X}$
3. if $(F_1 \vee \ldots \vee F_n) \in \mathbf{X}$ then for some $i \in \{1, \ldots, n\}$, $F_i \in \mathbf{X}$
4. if $(\exists x)F \in \mathbf{X}$ then $F(a/x) \in \mathbf{X}$ for some individual term a
5. if $(\forall x)F \in \mathbf{X}$ then $F(a/x) \in \mathbf{X}$ for each individual term a which occurs free in $\mathbf{X}$
6. no variable occurs both free and bound in $\mathbf{X}$

is called a *model set.* It is not necessary to include the condition that no variable occur both free and bound in the definition of a model set, but the only model sets we deal with are ones satisfying this condition and it is convenient to include it in the definition. If $\mathbf{X}$ is a model set and there are sufficiently many constants to do so, then we can define a set $\mathbf{Y}$ by: for each variable which occurs free in $\mathbf{X}$ substitute some constant, a different constant for each variable and none of them in $\mathbf{X}$. Then $\mathbf{Y}$ is a model set of sentences, so $\mathbf{Y}$ is satisfiable. $\mathbf{X}$ is satisfiable iff $\mathbf{Y}$ is satisfiable, so $\mathbf{X}$ is satisfiable. If there are not sufficient constants to form such a set $\mathbf{Y}$ then in some extended language which has sufficient constants (at most countably many can be needed), such a set $\mathbf{Y}$ can be formed. Then $\mathbf{Y}$ is satisfiable, so $\mathbf{X}$ is satisfiable in the extended language, so $\mathbf{X}$ is satisfiable in its original language. Thus every model set is satisfiable.

For embedding attributive constituents of the second kind in, Hintikka [1965a] (p. 80–81) defines a notion closely related to a model set called a *constitutive model set.* I use a definition which allows it to contain constituents in addition to attributive constituents.

A set $\mathbf{X}$ of constituents and/or attributive constituents of the second kind which satisfies the following conditions is called a *constitutive model set*:

1. If there is some $F \in \mathbf{X}$ which contains (unnegated) an atomic formula A all of whose individual terms are free in F, then no element of $\mathbf{X}$ contains $\neg A$.
2. If $F \in \mathbf{X}$ and F is of the form $\gamma^d(z_1, \ldots, z_k)$ or $\delta^d(z_1, \ldots, z_k)$ for some $d \geq 1$ then for each $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})$ in F, there is some individual term x such that $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})(x/z_{k+1})$, or a formula obtained from it by a change of bound variables, is in $\mathbf{X}$.
3. If $F \in \mathbf{X}$ and F is of the form $\gamma^d(z_1, \ldots, z_k)$ or $\delta^d(z_1, \ldots, z_k)$ for some $d \geq 1$ then for every individual term x which occurs free in any formula in $\mathbf{X}$, there is some $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})$ in F such that $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})(x/z_{k+1})$, or a formula obtained from it by a change of bound variables, is in $\mathbf{X}$.

4. No variable occurs both free and bound in $\mathbf{X}$.

Lemma 4.16 *Every constitutive model set is satisfiable.*

PROOF Since changing the bound variables in a formula results in an equivalent formula, every constitutive model set is satisfiable iff every set $\mathbf{X}$ of constituents and/or attributive constituents of the second kind which satisfies the following conditions is satisfiable:

1. If there is some $F \in \mathbf{X}$ which contains (unnegated) an atomic formula A all of whose individual terms are free in F, then no element of $\mathbf{X}$ contains $\neg A$.
2. If $F \in \mathbf{X}$ and F is of the form $\gamma^d(z_1, \ldots, z_k)$ or $\delta^d(z_1, \ldots, z_k)$ for some $d \geq 1$ then for each $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})$ in F, there is some individual term x such that $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})(x/z_{k+1}) \in \mathbf{X}$.
3. If $F \in \mathbf{X}$ and F is of the form $\gamma^d(z_1, \ldots, z_k)$ or $\delta^d(z_1, \ldots, z_k)$ for some $d \geq 1$ then for every individual term x which occurs free in any formula in $\mathbf{X}$, there is some $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})$ in F such that $\gamma^{d-1}(z_1, \ldots, z_k, z_{k+1})(x/z_{k+1}) \in \mathbf{X}$.
4. No variable occurs both free and bound in $\mathbf{X}$.

Let $\mathbf{X}$ be such a set. We show that $\mathbf{X}$ is satisfiable by embedding it in a model set. Let $\mathbf{Y}$ be the smallest set of formulas that satisfies the following conditions:

- $\mathbf{X} \subseteq \mathbf{Y}$.
- If $(F_1 \wedge \ldots \wedge F_n) \in \mathbf{X}$ then for each $i \in \{1, \ldots, n\}$, $F_i \in \mathbf{Y}$.
- If $F \in \mathbf{X}$ and F is of the form $\gamma^d(z_1, \ldots, z_k)$ or $\delta^d(z_1, \ldots, z_k)$ for some $d \geq 1$ then for each individual term x which occurs free in any element of $\mathbf{X}$, $\bigvee_{i \in I} \gamma_i^{d-1}(z_1, \ldots, z_{k+1})$ $(x/z_{k+1}) \in \mathbf{Y}$ where the attributive constituents $\gamma_i^{d-1}(z_1, \ldots, z_{k+1})$ are those which occur in F.

To show that $\mathbf{Y}$ is a model set: Note that $\mathbf{Y}$ has the same free individual terms as $\mathbf{X}$, and the same bound variables as $\mathbf{X}$, so no variable occurs both free and bound in $\mathbf{Y}$.

- If $A \in \mathbf{Y}$ and $\neg A \in \mathbf{Y}$ for some atomic formula A, then A and $\neg A$ are each either elements of $\mathbf{X}$ or conjuncts of elements of $\mathbf{X}$. Either way, all individual terms in A are free in these particular elements of $\mathbf{X}$, contradicting the definition of a constitutive model set.
- All conjuncts of elements of $\mathbf{Y}$ are in $\mathbf{Y}$ by its definition.

- If some disjunction is an element of $\mathbf{Y}$ then it is of the form $\bigvee_{i\in I}\gamma_i^{d-1}(z_1,\ldots,z_k,x)$ and there is some element of $\mathbf{X}$ of the form $\gamma^d(z_1,\ldots,z_k)$ or $\delta^d(z_1,\ldots,z_k)$ which contains the corresponding attributive constituents $\gamma_i^{d-1}(z_1,\ldots,z_k,z_{k+1})$, and x is free in $\mathbf{X}$ (since the free individual terms of $\mathbf{Y}$ are the same as those of $\mathbf{X}$). So, by the definition of constitutive model set, there is some $i \in I$ for which $\gamma_i^{d-1}(z_1,\ldots,z_k,x) \in \mathbf{X}$. So for this i, $\gamma_i^{d-1}(z_1,\ldots,z_k,x) \in \mathbf{Y}$.

- If $(\exists x)F \in \mathbf{Y}$ then F is some attributive constituent $\gamma^{d-1}(z_1,\ldots,z_k,x)$ which occurs in some element of $\mathbf{X}$ of the form $\gamma^d(z_1,\ldots,z_k)$ or $\delta^d(z_1,\ldots,z_k)$, so there is some individual term a such that $\gamma^{d-1}(z_1,\ldots,z_k,x)(a/x) \in \mathbf{X}$, so $\gamma^{d-1}(z_1,\ldots,z_k,x)(a/x) \in \mathbf{Y}$.

- If $(\forall x)F \in \mathbf{Y}$ then F is of the form $\bigvee_{\in I}\gamma_i^{d-1}(z_1,\ldots,z_k,x)$ where there is some element of $\mathbf{X}$ of the form $\gamma^d(z_1,\ldots,z_k)$ or $\delta^d(z_1,\ldots,z_k)$ which contains the attributive constituents $\gamma_i^{d-1}(z_1,\ldots,z_k,x)$. So by the definition of $\mathbf{Y}$, $\bigvee_{\in I}\gamma_i^{d-1}(z_1,\ldots,z_k,x) \in \mathbf{Y}$. □

We also use the following well-known result about trees. The proof here is from Smullyan [1968].

Lemma 4.17 (König's Lemma) *Every tree which is finitely branching and infinite has at least one infinite path.*

Proof Let $\mathcal{T}$ be a tree which is finitely branching and infinite. Call a node *good* if there are infinitely many nodes below it and *bad* if there are only finitely many below it. Then since $\mathcal{T}$ is infinite, its root is good. Now, if all successors of a node are bad, then the node itself is bad (since $\mathcal{T}$ is finitely branching). So every good node has at least one good successor. So $\mathcal{T}$ has an infinite path. □

Hintikka [1965a] (p. 75–85) claims the following completeness result:

For an inconsistent formula F of depth d, there is some depth $d+e$ such that the expansion of F to depth $d+e$ is one-trivially inconsistent.

However, there is an error in the attempted proof that is presented. *Ibid.* p. 85 says: "... there must be in (7) an a-constituent (9) of depth $d-1$ such that the bough of (7) determined by (9) is strongly symmetric with respect to x and a_i. Consider then ... the formula obtained from (9) by replacing x by a_i. ... this formula is identical with the result of reducing (7) with respect to (9) and therefore also identical with the result of omitting one layer of quantifiers from (7)."

Here, (7) is of the form $\gamma^d(a_1,\ldots,a_k)$; (9) is of the form $\gamma^{d-1}(a_1,\ldots,a_k,x)$; a_i is one of $a_1,\ldots,a_k$; and "the bough of (7) determined by (9) is strongly symmetric with respect to x and a_i" means that (9) is a partial expansion with respect to a_i of the root of (7). The claim that the result of replacing x by a_i in (9) is the same as the result of reducing (7) with respect to (9) means that replacing x by a_i in the root of (9) gives the root of (7), and replacing x by a_i in the subforest generated by the sibling-set of level 2 in (9) has the same effect as deleting all conjuncts which contain x from this subforest. But the first part of this claim, that substituting a_i for x in the root of (9) gives the root of (7), does not hold except for $i = k$, because a formula of the form $\gamma^0(a_1, \ldots, a_k, x)(a_i/x)$ contains all atomic formulas over $a_1, \ldots, a_k$ which contain a_i (with some repetitions) and none which don't contain a_i; and a formula of the form $\gamma^0(a_1, \ldots, a_k)$ contains all atomic formulas over $a_1, \ldots, a_k$ which contain a_k and none which don't contain a_k.

However, we do get the following similar result (lemma 4.10):

If $\gamma_b^{d-1}(z_1, \ldots, z_k, x)$ is a partial expansion to depth $d-1$ of $\delta_a^0(z_1, \ldots, z_k)$ with respect to z_i for some $i \in \{1, \ldots, k\}$, then substituting z_i for x in the root of $\gamma_b^{d-1}(z_1, \ldots, z_k, x)$ gives an attributive constituent $\gamma^0(\{z_1, \ldots, z_k\}, z_i)$ which is a part of $\delta_a^0(z_1, \ldots, z_k)$, and substituting z_i for x in the subforest generated by the sibling-set of level 2 in $\gamma_b^{d-1}(z_1, \ldots, z_k, x)$ has the same effect as deleting all conjuncts which contain x from this subforest.

I have not been able to prove the completeness theorem as stated by Hintikka, nor have I found any counterexample to it. But using two-trivial inconsistency, we do get a completeness theorem. I use a similar method of proof to the attempted proof found in Hintikka [1965a]. I first describe the method that will be used and give some preparatory explanation.

For a formula of depth d, there are in general many expansions to a greater depth e, but they can differ only in inconsistent elements. For constituents and attributive constituents, we define a particular expansion as follows: The *expansion*$_1$ of a constituent $\delta^d(z_1, \ldots, z_k)$ to a greater depth e contains exactly the constituents $\delta^e(z_1, \ldots, z_k)$ of which $\delta^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree (or, top-d forest in the case of no free individual terms), and which are not two-trivially inconsistent. Similarly, the *expansion*$_1$ of an attributive constituent $\gamma^d(z_1, \ldots, z_k)$ to a greater depth e contains exactly the attributive constituents $\gamma^e(z_1, \ldots, z_k)$ of which $\gamma^d(z_1, \ldots, z_k)$ is the top-$(d+1)$ tree (or, top-d forest in the case of no free individual terms), and which are not two-trivially inconsistent. By theorems 4.4 and 4.8, and the fact that the consistent elements of an expansion are unique, the expansion$_1$ of a constituent to some greater depth really is an expansion. Note that all expansion$_1$s can be found algorithmically.

The completeness theorem states that if a constituent C is inconsistent then there is some depth at which its expansion$_1$ is empty. This is equivalent to showing that if there is no depth at which its expansion$_1$ is empty, then C is consistent. If the constituent C is two-trivially inconsistent, then there is nothing to show. Otherwise, define the tree $\mathcal{T}$ as follows: the root is C, and the successors of each node X are the constituents in the expansion$_1$ of X to a depth 1-greater than itself. (Thus each node implies all the nodes that are above it.) Since every expansion has only a finite number of elements (i.e. disjuncts), $\mathcal{T}$ is finitely branching. If there is no depth at which the expansion$_1$ of C is empty, then $\mathcal{T}$ is infinite, so by König's Lemma (lemma 4.17), $\mathcal{T}$ contains an infinite path. (Conversely, if

$\mathcal{T}$ contains an infinite path, then there is no depth at which the expansion$_1$ of C is empty.) So, we must show that if C is the first element of an infinite sequence (or equivalently, if C occurs in an infinite sequence) of constituents where each element except the first is in the expansion$_1$ of its predecessor P to depth 1-greater than that of P, then C is consistent.

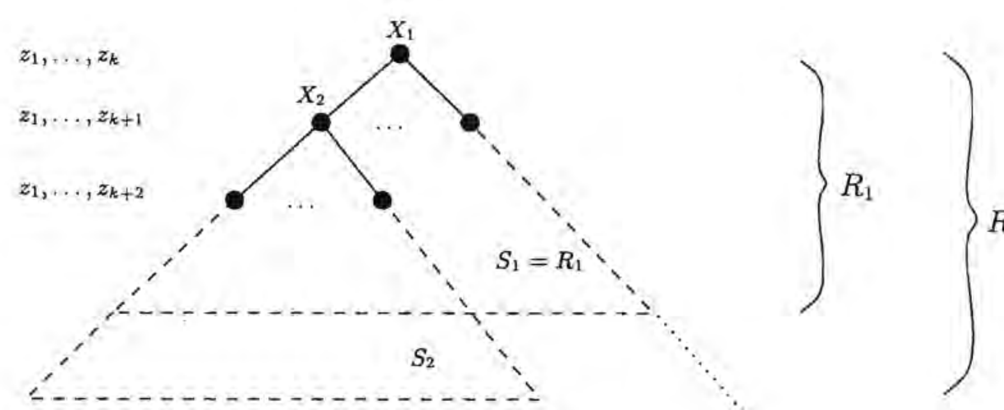
If the constituent which we must show to be consistent is of the form $\delta^d(z_1, \ldots, z_k)$, then $z_1, \ldots, z_k$ may include both variables and constants. By the definition of constituents, the bound variables used in an infinite sequence $\langle \delta^d(z_1, \ldots, z_k), \delta^{d+1}(z_1, \ldots, z_k), \ldots \rangle$ of constituents of increasing depth are all the variables not in $z_1, \ldots, z_k$. Instead of using all the variables here, we could use some infinite set such that its complement in the set of variables is also infinite. Let $z_{k+1}, z_{k+2}, \ldots$ be an infinite set of variables (in their alphabetic order) not in $z_1, \ldots, z_k$ which has an infinite complement $y_1, y_2, \ldots$ in the set of variables not in $z_1, \ldots, z_k$. Throughout the proof, formulas which differ from constituents or attributive constituents only in their bound variables will also be considered as constituents or attributive constituents respectively. In the sequence of constituents of increasing depth which we are given (which will be called $\mathbf{R}$), we will take the bound variables to be $z_{k+1}, z_{k+2}, \ldots$. We can do this since constituents defined using any infinite set of variables are equivalent. We show that the first element of $\mathbf{R}$ is consistent by embedding a formula which is equivalent to it in a constitutive model set. To do this, we form a sequence $\mathbf{S}$ whose first element is the first element of $\mathbf{R}$ and whose other elements are attributive constituents which occur in the corresponding elements of $\mathbf{R}$. We then form a sequence $\mathbf{Q}$ which is satisfiable iff $\mathbf{S}$ is, and whose first element is the same as the first element of $\mathbf{S}$ except for a change of bound variables (and thus is equivalent to it), and embed $\mathbf{Q}$ in a constitutive model set. In defining this constitutive model set, I use the notation (defined on p. 97) $\gamma^d(\{z_1, \ldots, z_k\}, z_i)$ for an attributive constituent of depth d whose free individual terms are $z_1, \ldots, z_k$ (not necessarily in this order) where the last free individual term is z_i.

Theorem 4.18 (Completeness theorem of the theory of distributive normal forms, cf. Hintikka [1965a]) *For an inconsistent formula F of depth d, there is some depth $d+e$ such that the expansion of F to depth $d+e$ is two-trivially inconsistent.*

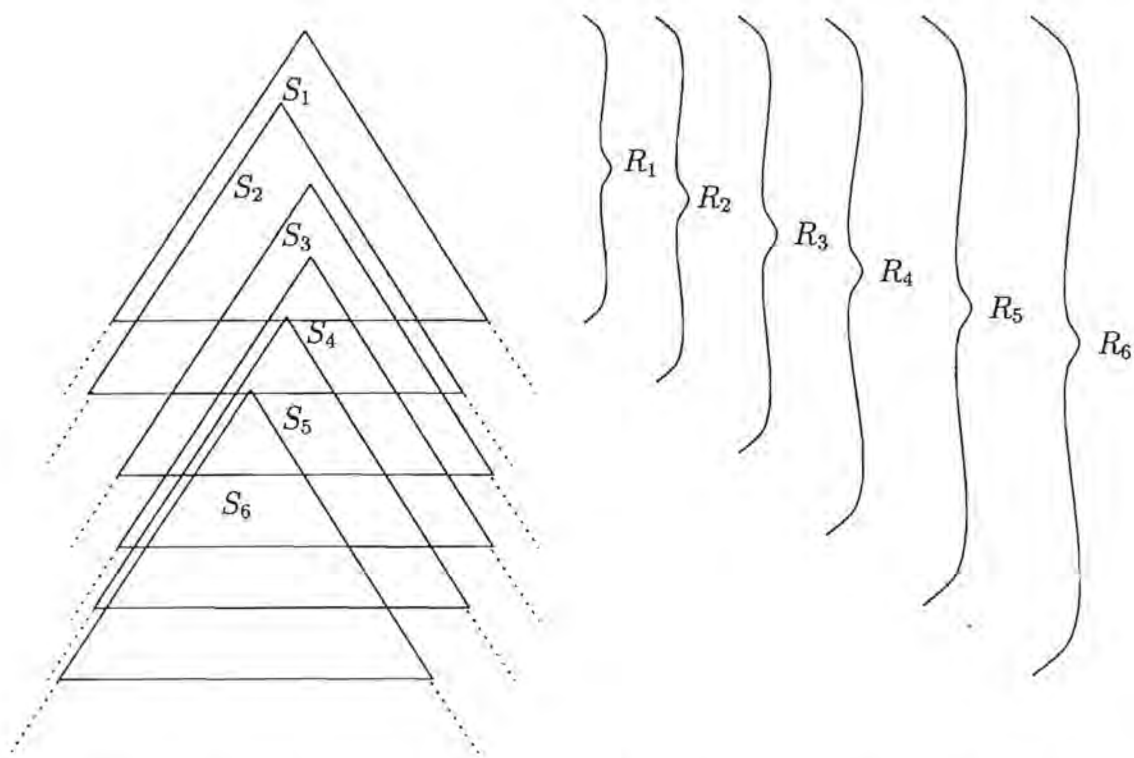
PROOF As we have seen above, we need to show that any constituent which can occur as the first element of an infinite sequence of constituents, where each element except the first is in the expansion$_1$ of its predecessor to 1-greater depth, is consistent. Let $\mathbf{R} = \langle R_1, R_2, \ldots \rangle$ be a sequence of constituents where for each $i \in \{2, 3, \ldots\}$, R_i is in the expansion$_1$ of R_{i-1} to depth 1-greater than that of R_{i-1}, and where R_1 is of the form $\delta^d(z_1, \ldots, z_k)$ and the bound variables in the elements of $\mathbf{R}$ are $z_{k+1}, z_{k+2}, \ldots$ and the variables not in any element of $\mathbf{R}$ are $y_1, y_2, \ldots$. We need to show that R_1 is consistent. We define a sequence $\mathbf{S} = \langle S_1, S_2, \ldots \rangle$ where S_1 is R_1 and for each $i \in \{2, 3, \ldots\}$, S_i is an attributive constituent which has one more free variable than S_{i-1}. For each element $S_i \in \mathbf{S}$, the root of S_i will be called X_i and the set of branches of X_i in S_i will be called $\mathbf{T}_i$. $\mathbf{T}_i$ will be partitioned into *ranks* which are themselves linearly ordered. Note that R_1 is the top-$(d+1)$ tree of R_2; R_2 is the top-$(d+2)$ tree of R_3; and in general R_i is the top-$(d+i)$ tree of R_{i+1}. Now, S_1 is

R_1 and X_1 is its root. $\mathbf{T}_1$ is the set of branches in S_1 of X_1. Let all the elements of $\mathbf{T}_1$ be of the same rank. The remaining elements of $\mathbf{S}$ are defined as follows: Given some S_i with the elements of $\mathbf{T}_i$ partitioned into some ranks, S_{i+1} is obtained by: Choose any element T of the highest rank in $\mathbf{T}_i$. Let X_{i+1} be the root of T and S_{i+1} be the subtree in R_{i+1} generated by X_{i+1}. And the ranks of elements of $\mathbf{T}_{i+1}$ are obtained by: Because R_{i+1} is not two-trivially inconsistent, the contracted copy of the subforest generated by the successor-set of X_{i+1} in R_{i+1} is the subforest generated by the successor-set of X_i in R_i. So, for each element $A = \gamma^{d-1}(z_1, \ldots, z_{k+i})$ of $\mathbf{T}_i$, there is some element $B = \gamma^{d-1}(z_1, \ldots, z_{k+i+1})$ of $\mathbf{T}_{i+1}$ such that the contracted copy of B is A; and for each $B \in \mathbf{T}_{i+1}$, the contracted copy of B is an element of $\mathbf{T}_i$. Any such B for which the contracted copy of B is A will be called a *correlate* of A. The particular element of $\mathbf{T}_i$ which is the top-d tree of S_{i+1} will be said to be *associated* with S_{i+1}. For each element $T \in \mathbf{T}_i$ except that one which is associated with S_{i+1}, choose some correlate $U \in \mathbf{T}_{i+1}$. For each rank in $\mathbf{T}_i$ (except the highest in the case it is a singleton), the set of chosen correlates of the elements of that rank form a rank in $\mathbf{T}_{i+1}$, and these ranks in $\mathbf{T}_{i+1}$ are ordered such that rank is preserved by correlates, and the set of elements of $\mathbf{T}_{i+1}$ which are not chosen correlates of elements of $\mathbf{T}_i$ form a rank which is the lowest in $\mathbf{T}_{i+1}$. (We don't actually need to use the axiom of choice to get the sequence $\mathbf{S}$ since for each choice that must be made, the set of attributive constituents from which the choice is made has, by its definition, some order, so we can just choose the first one. This proof doesn't depend on which particular sequence $\mathbf{S}$ is used.)

So X_2 is a successor in S_1 of X_1, and S_2 is the subtree in R_2 generated by X_2. The relation between these formulas can be represented graphically as follows:



And the sequence $\mathbf{S} = \langle S_1, S_2, \ldots \rangle$ has the following form:



By the definition of a correlate, if $U = \gamma^{d-1}(z_1, \ldots, z_{k+i+1})$ is a correlate of $T = \gamma^{d-1}(z_1, \ldots, z_{k+i})$ then T is the contracted copy of U. And, by the definition of being associated, if $T = \gamma^{d-1}(z_1, \ldots, z_{k+i})$ is associated with $S_{i+1} = \gamma^d(z_1, \ldots, z_{k+i})$ then T is the result of omitting a layer (any one) of quantifiers from S_{i+1}. If we now consider all the sets $\mathbf{T}_i$, we see that increasing i by 1 decreases the number of elements of the highest rank by 1, until that rank disappears, and creates a new rank lower than all the ones there were previously. Since each $\mathbf{T}_i$ is finite, each rank eventually disappears. Thus, for each i, for each $A_i \in \mathbf{T}_i$, there is some finite sequence $A_i, A_{i+1}, A_{i+2}, \ldots, A_{i+n}$ where each element other than the first is a correlate of its predecessor such that A_{i+n} is associated with S_{i+n+1}. (This is because when the highest rank disappears, we have the required sequences for each element of $\mathbf{T}_1$; when the second highest rank disappears, we have the required sequences for each element of $\mathbf{T}_2$; and so on.) Thus, for each i, for each $A \in \mathbf{T}_i$, there is some S_{i+n+1} such that omitting one layer of quantifiers from S_{i+n+1} and then taking the contracted copy n times results in A.

We now form a sequence $\mathbf{Q} = \langle Q_1, Q_2, \ldots \rangle$ of constituents in which each Q_i is very much like S_i, but with certain changes. Its bound variables are changed so that in $\mathbf{Q}$ no variable will be used both free and bound. Note that the individual terms used in $\mathbf{S}$ are $z_1, z_2, \ldots, z_k, z_{k+1}, z_{k+2}, \ldots$, and the variables not used in $\mathbf{S}$ are $y_1, y_2, \ldots$. Each element S_i of $\mathbf{S}$ is of the form $\gamma^d(z_1, \ldots, z_n)$ (or $\delta^d(z_1, \ldots, z_n)$ for $i = 1$) for $n = k + i - 1$, and the variables which occur bound in S_i are $z_{n+1}, \ldots, z_{n+d}$. Let Q_i be $X_1 \wedge \ldots \wedge X_{i-1} \wedge S_i$ with the bound variables $z_{n+1}, \ldots, z_{n+d}$ changed to $y_1, \ldots, y_d$ (simultaneously). Thus, considering the extra conjuncts $X_1, \ldots, X_{i-1}$ to be part of the root of Q_i, the form of Q_i

is $\delta^d(z_1, \ldots, z_{k+i-1})$. And $\mathbf{Q} = \langle Q_1, Q_2, \ldots \rangle$ is satisfiable iff $\mathbf{S}$ is. In $\mathbf{Q}$, $z_1, z_2, \ldots$ only occur free and $y_1, y_2, \ldots$ only occur bound, so these lists will be called respectively the free individual terms and the bound variables of $\mathbf{Q}$. The relation between the elements of $\mathbf{Q}$ is similar to that between the elements of $\mathbf{S}$. The $\mathbf{Q}$-*contracted copy* of some attributive constituent $\gamma^e(z_1, \ldots, z_j)$ whose bound variables are $y_1, \ldots, y_e$ is the result of deleting all atomic formulas containing z_{j-1} from $\gamma^e(z_1, \ldots, z_j)$. (Note that if taking $\mathbf{Q}$-contracted copies of subforests of elements of $\mathbf{Q}$ were to correspond to taking contracted copies of subforests of elements of $\mathbf{S}$ (in the sense of how the different elements of $\mathbf{Q}$ and $\mathbf{S}$ are related to each other), then the definition of $\mathbf{Q}$-contracted copies would have to include, after deleting the atomic formulas containing z_{j-1}, substituting z_{j-1} for z_j, but it does not.) To $\mathbf{Q}$-*omit* a layer of quantifiers is to omit it, but in the renaming of bound variables to use only $y_1, y_2, \ldots$ in their order. $\mathbf{Q}$-omitting layers of quantifiers in elements of $\mathbf{Q}$ corresponds to omitting layers of quantifiers in elements of $\mathbf{S}$ (in the sense of how the different elements of $\mathbf{Q}$ and $\mathbf{S}$ are related to each other). We have seen above that for each element $S_i = \gamma^d(z_1, \ldots, z_{k+i-1})$ of $\mathbf{S}$, for each $\gamma^{d-1}(z_1, \ldots, z_{k+i})$ in S_i, there is some $S_{i+n+1} \in \mathbf{S}$ such that by omitting a layer of quantifiers from S_{i+n+1} and taking contracted copies n times we get $\gamma^{d-1}(z_1, \ldots, z_{k+i})$. Thus for each $Q_i \in \mathbf{Q}$, for each $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)$ in Q_i, there is some $Q_{i+n+1} \in \mathbf{Q}$ such that from the part of Q_{i+n+1} of the form $\gamma^d(z_1, \ldots, z_{k+i+n})$, $\mathbf{Q}$-omitting a layer of quantifiers and then taking $\mathbf{Q}$-contracted copies n times we get $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)(z_{k+i+n}/y_1)$. Stating this in slightly less detail: for each $\delta^d(z_1, \ldots, z_j) \in \mathbf{Q}$, for each $\gamma^{d-1}(z_1, \ldots, z_j, y_1)$ in $\delta^d(z_1, \ldots, z_j)$, there is some $Q_n \in \mathbf{Q}$ and some z_h such that by $\mathbf{Q}$-omitting a layer of quantifiers from a part of Q_n and taking $\mathbf{Q}$-contracted copies a number of times we get $\gamma^{d-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$. Now, let $\mathbf{Y}$ be the smallest set of formulas which satisfies the following conditions:

- $\mathbf{Q} \subseteq \mathbf{Y}$.
- The result of $\mathbf{Q}$-omitting any layer of quantifiers from an element of $\mathbf{Y}$ is in $\mathbf{Y}$.
- If $\gamma^e(z_1, \ldots, z_j) \in \mathbf{Y}$ then the $\mathbf{Q}$-contracted copy of $\gamma^e(z_1, \ldots, z_j)$ is in $\mathbf{Y}$.
- If $\delta^e(z_1, \ldots, z_j) \in \mathbf{Y}$ then each part $\gamma^e(\{z_1, \ldots, z_j\}, z_i)$ of $\delta^e(z_1, \ldots, z_j)$ is in $\mathbf{Y}$.

$\mathbf{Y}$ is a set of constituents and attributive constituents which are not two-trivially inconsistent because: The omission of a layer of quantifiers from a constituent or attributive constituent which is not two-trivially inconsistent is a constituent or attributive constituent which is not two-trivially inconsistent (by the definition of two-trivial inconsistency), and similarly for the $\mathbf{Q}$-omission of a layer of quantifiers. And the contracted copy of an attributive constituent which is not two-trivially inconsistent is an attributive constituent (by definition of contracted copy). To show that it is not two-trivially inconsistent: If a node N in an attributive constituent A has a branch B which is a partial expansion with respect to z of the conjunction of N and the nodes above it, then in the contracted copy of A, the contracted copy of B is a partial expansion with respect to z of the contracted copy of the conjunction of N and the nodes above it. And for a sibling-set S in an attributive constituent A, if for a successor-set T of S, the contracted copy of the subforest generated by T is the top-n forest of the subforest generated by S, then in the contracted copy of A, the contracted copy of the subforest that corresponds to that generated by T is the top-n forest of the subforest that corresponds to that generated by S. Similarly, the **Q**-contracted copy of an attributive constituent which is not two-trivially inconsistent is an attributive constituent which is not two-trivially inconsistent. And if a constituent is not two-trivially inconsistent then each part of that constituent is also not two-trivially inconsistent. Thus **Y** is a set of constituents and attributive constituents (which are not two-trivially inconsistent). We now show that **Y** is a constitutive model set.

If any element F of **Y** contains an atomic formula A, negated or unnegated, all of whose individual terms are free in F, then there must be some element of **Q** which contains A in the same way, and then there must be some element of **S** which contains A in the same way. So we need only verify the first condition for constitutive model sets for **S**. Any atomic formula $\alpha(z_1, \ldots, z_j)$ which occurs in some element of **S** (negated or unnegated) occurs only in that element whose root is of the form $\gamma^0(z_1, \ldots, z_j)$. So no atomic formula all of whose individual terms are free occurs in more than one element of **S**. So no atomic formula occurs both negated and unnegated in elements of **S**, nor therefore in elements of **Q**, nor therefore in elements of **Y**.

To show that for each element $A \in \mathbf{Y}$ of the form $\gamma^e(z_1, \ldots, z_j)$ or $\delta^e(z_1, \ldots, z_j)$, for each $\gamma^{e-1}(z_1, \ldots, z_j, y_1)$ in A, there is some z_h such that $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1) \in \mathbf{Y}$: We first show that for each $Q_i = \delta^d(z_1, \ldots, z_{k+i-1}) \in \mathbf{Q}$, for each $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)$ in Q_i, there is some z_h such that $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)(z_h/y_1) \in \mathbf{Y}$. We have seen above that for each $Q_i \in \mathbf{Q}$, for each $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)$ in Q_i, there is some $Q_n \in \mathbf{Q}$ and some z_h such that by **Q**-omitting a layer of quantifiers from a part of Q_n and taking **Q**-contracted copies a number of times we get $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)(z_h/y_1)$. By the definition of **Y**, $\gamma^{d-1}(z_1, \ldots, z_{k+i-1}, y_1)(z_h/y_1) \in \mathbf{Y}$. Now, if for some $A \in \mathbf{Y}$, for each $\gamma^{e-1}(z_1, \ldots, z_j, y_1)$ in A, there is some z_h such that $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1) \in \mathbf{Y}$, then for A with a layer of quantifiers **Q**-omitted, each branch of its root with an appropriate substitution is in **Y** because these branches are $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$ with a layer of quantifiers **Q**-omitted; and similarly for the **Q**-contracted copy of A, the required branches of the root with some substitution are in **Y** because they are the **Q**-contracted copies of $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$. And if the branches of some constituent A with some substitution are in **Y**, then a similar condition holds for each part of A since the branches of the parts of A are the same as the branches of A. Thus the required condition holds for all elements of **Y**.

To show that for each element $A \in \mathbf{Y}$ of the form $\gamma^e(z_1, \ldots, z_j)$ or $\delta^e(z_1, \ldots, z_j)$, for each free individual term z_h of **Y**, there is some $\gamma^{e-1}(z_1, \ldots, z_j, y_1)$ in A for which $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1) \in \mathbf{Y}$: We first show that for each $Q_i = \delta^d(z_1, \ldots, z_j) \in \mathbf{Q}$, for each z_h, there is some $\gamma^{d-1}(z_1, \ldots, z_j, y_1)$ in Q_i for which $\gamma^{d-1}(z_1, \ldots, z_j, y_1)(z_h/y_1) \in \mathbf{Y}$. Let $Q_i = \delta^d(z_1, \ldots, z_j)$ be an element of **Q** and z_h a free individual term of **Y**. If $h > j$

then consider the first element $Q_n \in \mathbf{Q}$ which contains z_h. In $\mathbf{T}_{n-1}$ is some T_{n-1} which is associated with S_n, and T_{n-1} is a correlate of some $T_{n-2} \in \mathbf{T}_{n-2}$, and T_{n-2} is a correlate of some $T_{n-3} \in \mathbf{T}_{n-3}$, and so on until we reach some $T_i \in \mathbf{T}_i$. Now T_i is the result of omitting a layer of quantifiers from S_n and then taking contracted copies a number of times. Let $\gamma^{d-1}(z_1, \ldots, z_j, y_1)$ be that branch of the root of Q_i which corresponds to T_i, then $\gamma^{d-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$ is the result of $\mathbf{Q}$-omitting a layer of quantifiers from Q_n and then taking $\mathbf{Q}$-contracted copies a number of times. So $\gamma^{d-1}(z_1, \ldots, z_j, y_1)(z_h/y_1) \in \mathbf{Y}$. If $h \leq j$ then there is some $\gamma_b^{d-1}(z_1, \ldots, z_j, y_1)$ in Q_i which is a partial expansion of the root $\delta_a^0(z_1, \ldots, z_j)$ of Q_i with respect to z_h to depth $d-1$ (since the root of Q_i is the conjunction of the root of S_i and those nodes above it in R_i, and R_i is not two-trivially inconsistent). So by lemma 4.10, substituting z_h for y_1 in $\gamma_b^{d-1}(z_1, \ldots, z_j, y_1)$ gives an attributive constituent whose root is a part of $\delta_a^0(z_1, \ldots, z_j)$ and whose subforest generated by the sibling-set of level 2 is the result of deleting all conjuncts which contain y_1 from the subforest generated by the sibling-set of level 2 of $\gamma_b^{d-1}(z_1, \ldots, z_j, y_1)$. Thus $\gamma_b^{d-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$ can be obtained by changing the bound variables of the result of omitting a layer of quantifiers from a part of Q_i. So, by the definition of $\mathbf{Y}$, $\gamma_b^{d-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$ with a change of bound variables is in $\mathbf{Y}$. Now, if for some $A \in \mathbf{Y}$, for each z_h, there is some $\gamma^{e-1}(z_1, \ldots, z_j, y_1)$ in A for which $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1) \in \mathbf{Y}$, then for A with a layer of quantifiers $\mathbf{Q}$-omitted, for each z_h, $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$ with a layer of quantifiers $\mathbf{Q}$-omitted is the required element of $\mathbf{Y}$; and similarly for the $\mathbf{Q}$-contracted copy of A, for each z_h, the corresponding $\mathbf{Q}$-contracted copy of $\gamma^{e-1}(z_1, \ldots, z_j, y_1)(z_h/y_1)$ is the required element of $\mathbf{Y}$. And if the branches of some constituent A with some substitution are in $\mathbf{Y}$, then a similar condition holds for each part of A since the branches of the parts of A are the same as the branches of A. Thus the required condition holds for all elements of $\mathbf{Y}$.

So $\mathbf{Y}$ is a constitutive model set. Thus Q_1 is consistent, thus S_1 is consistent. That is, R_1 is consistent. □

Recall that a disproof method is a method of proving the inconsistency of formulas. A proof of the inconsistency of F will be called a *disproof* of F.

We have the following disproof method for first-order logic:

Method 1 Given a formula F of depth d with free individual terms $z_1, \ldots, z_k$,

- convert F to distributive normal form $\bigvee_{a \in A} \delta_a^d(z_1, \ldots, z_k)$
- repeat:
    - omit all two-trivially inconsistent constituents
    - if no constituents remain, halt
    - expand to depth 1-greater than current depth.

The distributive normal forms and expansions meant here are the particular ones that are found by the algorithms given earlier. A *disproof* of F consists of a halting execution of the above algorithm applied to F.

A proof or disproof method need not be an algorithm, but in the case of method 1 it is. Since

1. if a formula is two-trivially inconsistent then it is inconsistent, and
2. if a formula is inconsistent then it has some expansion which is two-trivially inconsistent (by theorem 4.18),

method 1 is a complete disproof method. We use it to define the following proof method:

Method 2 A *proof* of F is a disproof by method 1 of $\neg F$.

Since method 1 is a complete disproof method and the valid formulas are exactly those whose negations are inconsistent, method 2 is a complete proof method. Thus the completeness theorem of the theory of distributive normal forms gives one way of showing the completeness of first-order logic.

The way of obtaining a complete proof method from a complete disproof method used above works for any complete disproof method. Similarly, from any complete proof method, we can obtain a complete disproof method.

We use theorem 4.18 to show that we can not determine algorithmically how many of the constituents of the different forms are consistent.

Theorem 4.19 *For a language with at least one n-ary predicate symbol for some $n \geq 2$, there is no algorithm for finding how many constituents of each form are consistent.*

PROOF By the completeness theorem 4.18, if we could find how many constituents of the form $\delta^d(z_1, \ldots, z_k)$ are inconsistent, we could expand all of the constituents of the form $\delta^d(z_1, \ldots, z_k)$ to greater depths until the number of constituents found to be two-trivially inconsistent is the number of inconsistent constituents. Then all remaining constituents of the form $\delta^d(z_1, \ldots, z_k)$ would be consistent. If we could do this for the constituents of each form, we would have an algorithm for deciding whether or not any constituent is consistent. But there is no such algorithm by theorem 4.9. □

Chapter 5

Some proposed applications of distributive normal forms

This chapter gives some examples of how distributive normal forms have been used in inductive logic and verisimilitude.

5.1 Inductive logic

Carnap describes the *inductive logic* which he developed as a theory of probability which can be used for inductive reasoning and determining rational decisions (see e.g. Carnap and Jeffrey [1971], chapter 1). a reasoning for of that every and right, the often, of thinking is valid by It was motivated by normative decision theory which is a theory of rational decision making. Such a theory is described by *ibid.* (p. 7–9) as follows: A person X at a time T must choose among possible acts $A_1, \ldots, A_m$ (for some $m \in \mathbb{N}$). The possible states of the part of nature relevant for this decision are $W_1, \ldots, W_n$ (for some $n \in \mathbb{N}$). X knows these states, but does not know which one of them is the actual state. The outcome of carrying out act A_i in state W_j, denoted $O_{i,j}$, is uniquely determined by A_i and W_j and X knows how it is determined. X has a utility function U_X (which expresses the desirability of the possible outcomes) and knows this function. The *value* of A_i for X at time T is

$$V_{X,T}(A_i) = \sum_{j=1}^{n}(U_X(O_{i,j})P(W_j))$$

where $P(W_j)$ is the "probability" of W_j. If X can determine the probability of each possible state W_j then this probability is used in the above equation. But in general we can not expect X to know these probabilities, and $P(W_j)$ is instead taken to be the "degree of belief" of X in the statement "W_j is the actual state". The choice of any A_i for which

$V_{X,T}(A_i)$ is maximal is called a *rational decision.* (Such a choice is said to follow the *Bayesian rule* of decision making.)

We can ask what kind of functions P can be used to model degree of belief (also called *credence*) in statements. To try to find functions to model *rational credence* (and some related concepts) is the subject of inductive logic. This has been attempted for some classes of first-order languages.

A number of different intuitive considerations and particular definitions for inductive logic have been studied. A discussion and comparison of different approaches can be found, for example, in Hintikka and Suppes [1966] and Lakatos [1968]. In this section, I will just consider one of Hintikka's definitions which uses constituents and distributive normal forms and the definitions of Carnap on which it is based.

One condition of rationality for a credence function on which everyone seems to agree is that it must be a probability measure (or weak probability measure) as defined below.

Let Ω be a set. A set $\mathcal{F}$ of subsets of Ω is a *σ-field over* Ω if

- $\emptyset, \Omega \in \mathcal{F}$
- if $E \in \mathcal{F}$ then $\Omega \setminus E \in \mathcal{F}$
- if for each $n \in \mathbb{N}$, $E_n \in \mathcal{F}$ then $\bigcap_{n\in\mathbb{N}} E_n \in \mathcal{F}$
- if for each $n \in \mathbb{N}$, $E_n \in \mathcal{F}$ then $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{F}$.

If the last two conditions in the above definition are replaced by:

- if $E_1, E_2 \in \mathcal{F}$ then $E_1 \cap E_2 \in \mathcal{F}$
- if $E_1, E_2 \in \mathcal{F}$ then $E_1 \cup E_2 \in \mathcal{F}$

then we get a definition for a *field over* Ω. Any subset of a field $\mathcal{F}$ which is itself a field is called a *subfield* of $\mathcal{F}$. Let Ω be a set and $\mathcal{A}$ a set of subsets of Ω. The *field [σ-field] over Ω generated by $\mathcal{A}$* is the intersection of all the fields [σ-fields] over Ω which are supersets of $\mathcal{A}$.

Let $\mathcal{F}$ be a σ-field over Ω. A function $P : \mathcal{F} \to \mathbb{R}$ is a *probability measure on* $\mathcal{F}$ if

- for each $E \in \mathcal{F}$, $0 \leq P(E) \leq 1$
- $P(\emptyset) = 0$, $P(\Omega) = 1$
- for every disjoint sequence $\{E_n\}_{n\in\mathbb{N}}$ of elements of $\mathcal{F}$, $P(\bigcup_{n\in\mathbb{N}} E_n) = \sum_{n\in\mathbb{N}} P(E_n)$.

If the last condition in the above definition is replaced by:

- if $E_1, E_2 \in \mathcal{F}$ and $E_1 \cap E_2 = \emptyset$ then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

then we get a definition for a *weak probability measure.*

If a probability measure P as above also satisfies

- if $E \in \mathcal{F}$, $A \subseteq E$, $P(E) = 0$ then $A \in \mathcal{F}$

then it is called *complete.*

If P is a complete probability measure on a σ-field $\mathcal{F}$ over Ω, then

- if $E \in \mathcal{F}$, $P(E) = 0$, $A \subseteq E$ then $A \in \mathcal{F}$ and $P(A) = 0$
- if $E \in \mathcal{F}$, $P(E) = 1$, $\Omega \supseteq A \supseteq E$ then $A \in \mathcal{F}$ and $P(A) = 1$.

Probability measures are used to model *random experiments* or *games of chance*, that is situations in which some procedure is carried out which has a number of different possible outcomes. Each time the experiment is performed exactly one *outcome* occurs. An *event* is a subset of the set of all possible outcomes. It can be regarded as a conjecture about the result of the experiment. The event $\emptyset$ is called *the impossible event.* We won't necessarily consider all sets of outcomes as events, but only some which are convenient.

A *probability space* is a triple $\langle \Omega, \mathcal{F}, P \rangle$ where

- Ω is a set, the elements of which are called *outcomes*
- $\mathcal{F}$ is a σ-field over Ω, whose elements are called *events*
- $P : \mathcal{F} \to \mathbb{R}$ is a complete probability measure on $\mathcal{F}$ for which $P(E)$ is called the *probability* of E.

If in some probability space $\langle \Omega, \mathcal{F}, P \rangle$, $P(E) = 1$ then the event E is said to occur *almost certainly.* In terms of the experiment that is modelled, if $P(E) = 0$ then E is not necessarily impossible, but occurs only a *negligible* proportion of the time when the experiment is repeated a large number of times. More generally, if an experiment is repeated n times and the outcome is ω_i for $r(n)$ of them, then the *frequency* of ω_i in the n repetitions is $\frac{r(n)}{n}$. If, as n increases, the ratios $\frac{r(n)}{n}$ approach some number p_i, then p_i is called the *probability* of ω_i. This is modelled in the probability space by letting $P(\omega_i) = p_i$. This is called the *frequency interpretation of probability.*

One of the simplest examples of a probability space is the following:

Let $\Omega = \{\omega_i\}_{i \in I}$ where I is a finite set $\{1, \ldots, n\}$ or I is $\mathbb{N}$, and let $\{p_i\}_{i \in I}$ be such that each $p_i \geq 0$ and $\sum_{i \in I} p_i = 1$. Then let $\mathcal{F} = \mathcal{P}(\Omega)$ and define $P : \mathcal{F} \to \mathbb{R}$ by

$$P(E) = \sum_{\{i \mid \omega_i \in E\}} p_i.$$

Then $\langle \Omega, \mathcal{F}, P \rangle$ is a probability space.

function) (represented proposition). $L.\ c : Sent(L) \times$e,e$_1$, $inconsistentcanbemeasureconfirmati$

If the probabilities of events are given by some probability measure (or weak probability measure) P, then we can define a function which gives the probabilities of events *conditional* on other events. If A, B are events and $P(B) > 0$, then the *conditional probability* of A given B is

$$P'(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

The function P' is called a *conditional probability function*.

If P is a probability measure which models rational credence then the conditional probability function P' models *rational conditional credence* where $P'(A \mid B)$ is what the rational degree of belief in A would be if the evidence were B. For those probability measures used to model rational credence, the outcomes are models (as defined in chapter 1) and the events are *propositions*, where in this context a *proposition* is a set of models.

An inductive logic is relative to a first-order language. We will consider monadic first-order languages which contain:

- a non-zero countable number of constants: a_1, a_2, $\ldots$;
- a non-zero countable number of unary predicates: $P_1^1, P_2^1, \ldots, P_1^2, P_2^2, \ldots, \ldots, P_1^n, P_2^n, \ldots$. The set of predicates is partitioned into a finite number of *families*: $F^1 = \{P_1^1, P_2^1, \ldots\}$, $F^2 = \{P_1^2, P_2^2, \ldots\}$, ..., $F^n = \{P_1^n, P_2^n, \ldots\}$; and the set of families $\{F^1, F^2, \ldots, F^n\}$ is denoted $\mathcal{F}$.

For such a language $\mathcal{L}$, let $\mathcal{Z}_{\mathcal{L}}$ be the set of models $\langle D, I \rangle$ for $\mathcal{L}$ such that

- $D = \mathbf{Const}_{\mathcal{L}}$ and I restricted to $\mathbf{Const}_{\mathcal{L}}$ is a permutation of $\mathbf{Const}_{\mathcal{L}}$;
- for each family F^m of $\mathcal{L}$, for each $a_i \in D$, there is exactly one predicate $P_j^m \in F^m$ such that $a_i \in I(P_j^m)$ (that is, the interpretations of the predicates in F^m form a partition of D).

For each sentence S of $\mathcal{L}$, let E_S be the set of models in $\mathcal{Z}_{\mathcal{L}}$ in which S is true.

For a language $\mathcal{L}$, the set $\mathcal{E}_\mathcal{L}$ of *propositions of* $\mathcal{L}$ is the σ-field over $\mathcal{Z}_\mathcal{L}$ generated by $\{E_S \mid S$ is a sentence of $\mathcal{L}\}$.

The proposition E_S is said to *correspond* to S, and S is said to *describe* E_S. If S is an *atomic sentence* (i.e. of the form $P_j^m(a_i)$) then E_S is called an *atomic proposition*. The set of *molecular propositions of* $\mathcal{L}$ is the field over $\mathcal{Z}_\mathcal{L}$ generated by $\{E_S \mid S$ is an atomic sentence of $\mathcal{L}\}$. Or equivalently, a molecular proposition of $\mathcal{L}$ is a proposition E_S where S is a sentence of $\mathcal{L}$ which does not contain identity or quantifiers. The set $\mathcal{E}_\mathcal{L}$ can also be equivalently defined as the σ-field over $\mathcal{Z}_\mathcal{L}$ generated by $\{E_S \mid S$ is an atomic sentence of $\mathcal{L}\}$. (Both of these equivalences are shown in Carnap [1971].)

Carnap [1971] considers credence functions P that are weak probability measures and satisfy some further conditions. The first is that the only molecular proposition H for which $P(H) = 0$ is $H = \emptyset$ (which is a strengthening of the condition $P(\emptyset) = 0$). The second involves the way the credence function changes in time. It requires that the credence function at a later time depend only on the credence function at an earlier time and the evidence obtained between those two times, where this dependence is of a particular form. The credence function at time T_n will be denoted P_n and at some later time T_{n+1} will be denoted P_{n+1}. If E is the evidence obtained between T_n and T_{n+1} then P_{n+1} must be the function defined by

$$P_{n+1}(H) = \frac{P_n(E \cap H)}{P_n(E)} = P'_n(H \mid E). \tag{5.1}$$

(The above function P_{n+1} is only defined for evidence E such that $P_n(E) \neq 0$. This is the case for any evidence that can be obtained provided that the only molecular proposition H for which $P(H) = 0$ is $H = \emptyset$.) We now consider a sequence of credence functions at different times, starting with some time T_0, where each change is obtained by the formula (5.1). For each $i \in \{1, \ldots, n\}$, let the evidence obtained between T_{i-1} and T_i be denoted E_i. Let $K_n = \bigcap_{i=1}^{n} E_i$. Then, by repeatedly applying the definition (5.1), we get

$$P_n(H) = P'_0(H \mid K_n).$$

P_0 is an *initial credence function*. The *conditional initial credence function* P'_0 is also called a *credibility function*. A conditional credence function is also called a *confirmation function*.

So, to determine the values of the different possible acts (and hence the possible rational decisions), we need some credibility function P'_0. It can either be defined in terms of a corresponding initial credence function P_0, or can be taken as primitive. Carnap [1971] (p. 38–40) defines both initial credence functions and credibility functions and shows that for those which satisfy the axiom of *regularity* (defined below), each can be defined in terms of the other.

A *credibility function* for the language $\mathcal{L}$ is a function $\mathcal{C} : (\mathcal{E}_\mathcal{L} \times (\mathcal{E}_\mathcal{L} \setminus \{\emptyset\})) \to \mathbb{R}$ which satisfies the following conditions:

- $0 \leq \mathcal{C}(H \mid E) \leq 1$
- if $E \subseteq H$ then $\mathcal{C}(H \mid E) = 1$
- if $E \cap G \cap H = \emptyset$ then $\mathcal{C}(G \cup H \mid E) = \mathcal{C}(G \mid E) + \mathcal{C}(H \mid E)$
- if $E \cap H \neq \emptyset$ then $\mathcal{C}(H \cap G \mid E) = \mathcal{C}(H \mid E)\mathcal{C}(G \mid E \cap H)$.

An *initial credence function* for the language $\mathcal{L}$ is a function $\mathcal{M} : \mathcal{E}_{\mathcal{L}} \to \mathbb{R}$ which is a weak probability measure.

If $\mathcal{C}$ is a credibility function for $\mathcal{L}$ and E is a nonempty element of $\mathcal{E}_{\mathcal{L}}$ then the function $\mathcal{C}_E : \mathcal{E}_{\mathcal{L}} \to \mathbb{R}$ defined by

$$\mathcal{C}_E(A) = \mathcal{C}(A \mid E)$$

is is an initial credence function for $\mathcal{L}$.

A credibility function $\mathcal{C}$ and an initial credence function $\mathcal{M}$ for $\mathcal{L}$ are *related* to each other if: for all $H \in \mathcal{E}_{\mathcal{L}}$ and all nonempty $E \in \mathcal{E}_{\mathcal{L}}$,

$$\mathcal{C}(H \mid E) = \frac{\mathcal{M}(E \cap H)}{\mathcal{M}(E)}.$$

This is equivalent to the condition that $\mathcal{M} = \mathcal{C}_{\mathcal{Z}_{\mathcal{L}}}$. So, for any credibility function for $\mathcal{L}$, there is exactly one related initial credence function for $\mathcal{L}$.

If $\mathcal{M}$ is an initial credence function which satisfies the condition that for every nonempty molecular proposition H in which only finite families are involved, $\mathcal{M}(H) > 0$, then $\mathcal{M}$ is called *regular*.

If $\mathcal{C}$ is a credibility function which satisfies the condition that for all molecular propositions E and H involving only finite families such that $E \cap H \neq \emptyset$, $\mathcal{C}(H \mid E) > 0$, then $\mathcal{C}$ is called *regular*.

An initial credence function $\mathcal{M}$ is regular iff it is related to a regular credibility function, and a credibility function $\mathcal{C}$ is regular iff it is related to a regular initial credence function. Thus there is a one-to-one correspondence between regular initial credence functions and regular credibility functions for a language $\mathcal{L}$.

The conditions given so far for initial credence functions and credibility functions (including regularity) are also required for all the other credence functions and conditional credence functions if they are to be considered rational. The following condition, called *symmetry*, applies only to initial credence functions and credibility functions.

If X is any nonempty countable set and π is a permutation of X, then π is called a *finite permutation* if the number of elements $x \in X$ for which $\pi(x) \neq x$ is finite.

If π is a permutation of $\mathbf{Const}_{\mathcal{L}}$, then the *induced mapping between models* $\pi' : \mathcal{Z}_{\mathcal{L}} \to \mathcal{Z}_{\mathcal{L}}$ is the function defined by $\pi'(\langle D, I \rangle) = \langle D, I' \rangle$ where

$$\begin{aligned} I'(a_i) &= \pi(I(a_i)) \\ I'(P_j^m) &= \{a_i \mid \pi^{-1}(a_i) \in I(P_j^m)\}. \end{aligned}$$

And the *induced mapping between propositions* $\pi'' : \mathcal{E}_{\mathcal{L}} \to \mathcal{E}_{\mathcal{L}}$ is the function defined by

$$\pi''(A) = \{\pi'(Z) \mid Z \in A\}.$$

Two models $Z, Z' \in \mathcal{Z}_{\mathcal{L}}$ are *isomorphic* if there is some permutation $\pi : \mathbf{Const}_{\mathcal{L}} \to \mathbf{Const}_{\mathcal{L}}$ such that $Z' = \pi'(Z)$. Two propositions $A, B \in \mathcal{E}_{\mathcal{L}}$ are *isomorphic* if there is some permutation $\pi : \mathbf{Const}_{\mathcal{L}} \to \mathbf{Const}_{\mathcal{L}}$ such that $A = \pi''(B)$.

A credibility function $\mathcal{C}$ is *symmetric* if for each finite permutation $\pi : \mathbf{Const}_{\mathcal{L}} \to \mathbf{Const}_{\mathcal{L}}$, for each H and nonempty E in $\mathcal{E}_{\mathcal{L}}$, $\mathcal{C}(H \mid E) = \mathcal{C}(\pi''(H) \mid \pi''(E))$. An initial credence function $\mathcal{M}$ is *symmetric* if for each finite permutation $\pi : \mathbf{Const}_{\mathcal{L}} \to \mathbf{Const}_{\mathcal{L}}$, for each H in $\mathcal{E}_{\mathcal{L}}$, $\mathcal{M}(H) = \mathcal{M}(\pi''(H))$. If a credibility function $\mathcal{C}$ and an initial credence function $\mathcal{M}$ are related then $\mathcal{C}$ is symmetric iff $\mathcal{M}$ is symmetric.

The *basic system of inductive logic* described by Carnap [1971] uses credibility functions that are regular and symmetric, or equivalently initial credence functions that are regular and symmetric.

We now consider a language with just one family of predicates. Let these predicates be $P_1, \ldots, P_k$ (for some finite $k \geq 2$). Thus the universe is partitioned into k *cells*. We consider the specific type of evidence which gives, for some finite number n of individuals, the cells to which they belong. This kind of evidence is called a *sample* of size n. The *representative function* f of some credibility function $\mathcal{C}$ gives, for each cell, the probability that the next individual to be observed belongs to that cell. A representative function f can be considered as consisting of k representative functions $f_1, \ldots, f_k$ where each f_i gives the probability that the next individual belongs to the ith cell. If $\mathcal{C}$ is symmetric, then its representative function f depends on the observed individuals only through the total numbers which have been observed in each cell, and not on the order in which they were observed. For the derivation of his λ-continuum of inductive methods, Carnap assumes the stronger condition that each representative function f_i depends on the sample only through its size n and the number n_i of individuals which have been observed in the ith cell. Carnap [1973] shows that with this assumption, the representative functions f_i are of the form (for $k > 2$)

$$f_i(n_i, n) = \frac{n_i + \frac{\lambda}{k}}{n + \lambda}$$

where $\lambda \in \mathbb{R}^+$; and for $k = 2$ the functions f_i are also of the above form if it is assumed that f_i is a linear function of n_i. The values of λ and $f_i(0, 1)$ uniquely determine each other through the equation

$$f_i(0,1) = \frac{\frac{\lambda}{k}}{1+\lambda}.$$

A disadvantage of Carnap's λ-continuum is that for an infinite universe, all confirmation functions based on it assign, on any finite evidence, a value of 0 to all constant-free sentences which are not logically true. Hintikka [1966] introduces a *two-dimensional continuum of inductive methods* which does not have this disadvantage and includes Carnap's λ-continuum as a limit case.

We consider a monadic first-order language which contains a finite number of predicates $P_1, \ldots, P_k$, but it is *not* assumed that these predicates partition the universe. Recall, from section 1.6, the definition of attributive constituents

$$A_i(x) = (\pm)P_1(x) \wedge \ldots \wedge (\pm)P_k(x)$$

and constituents

$$C_i = (\pm)(\exists x)A_1(x) \wedge \ldots \wedge (\pm)(\exists x)A_K(x)$$

where $K = 2^k$. A constituent $(\pm)(\exists x)A_1(x) \wedge \ldots \wedge (\pm)(\exists x)A_K(x)$ can also be written in an equivalent form as $(\exists x)A_{i_1}(x) \wedge \ldots \wedge (\exists x)A_{i_c}(x) \wedge (\forall x)(A_{i_1}(x) \vee \ldots \vee A_{i_c}(x))$. (Since we are dealing only with languages with a finite number of predicates, we use the attributive constituents and constituents relative to all the predicates.) Now, the attributive constituents $A_1(x), \ldots, A_K(x)$ partition the universe into K cells. We consider evidence which says, for n individuals, which attributive constituents they satisfy. Let the attributive constituents which are satisfied by any of these n individuals be $A_{i_1}(x), \ldots, A_{i_c}(x)$; and for each $j \in \{1, \ldots, c\}$, let n_j be the number of the observed individuals which satisfy $A_{i_j}(x)$. So $\sum_{j=1}^{c} n_j = n$.

A probability measure p and related conditional probability function p' which can be used for credence and confirmation functions respectively will be defined. Carnap approached the problem of defining such functions by considering the case of defining the probability that the next individual to be observed belongs to a particular cell, given a sample of n individuals (called the problem of *singular predictive inference*). Hintikka [1966] considers instead the question of (given a sample of n individuals) which constituent should be the most probable (called the problem of *universal inference*). It is assumed that constituents which contain (positively) the same number of attributive constituents have the same probability on no evidence, and also that p' satisfies the condition that for n sufficiently large relative to K:

- If E is a sample of n individuals which satisfy $A_{i_1}(x), \ldots, A_{i_c}(x)$ and C is the constituent $(\exists x)A_{i_1}(x) \wedge \ldots \wedge (\exists x)A_{i_c}(x) \wedge (\forall x)(A_{i_1}(x) \vee \ldots \vee A_{i_c}(x))$, then for any constituent D which contains more than c attributive constituents, $p'(C \mid E) > p'(D \mid E)$.

Once $p(C)$ is defined for constituents C, we get (by the additivity of p) for every constant-free sentence F,

$$p(F) = \sum_{i \in I} p(C_i)$$

where $\bigvee_{i \in I} C_i$ is the distributive normal form of F. But for evidence which contains constants, p' is not yet defined.

The function $\pi : (\mathbb{N} \times \mathbb{R}^+) \to \mathbb{R}$ is defined by

$$\pi(\alpha, \lambda) = (\lambda)(1 + \lambda)(2 + \lambda) \ldots (\alpha - 1 + \lambda).$$

Let $E^c(n_1, \ldots, n_c)$ be evidence as above where each n_j is the number of observed individuals in $A_{i_j}(x)$, and let C^w be a constituent which contains w attributive constituents. Then, for an infinite universe, p is defined as follows (Hintikka [1966], p. 117–121):

$$p(C^w) = \frac{\pi(\alpha, \frac{w\lambda}{K})}{\sum_{i=0}^{K} \binom{K}{i} \pi(\alpha, \frac{i\lambda}{K})}$$

where $\alpha \in \mathbb{N}$, $\lambda \in \mathbb{R}^+$; and

$$p'(E^c(n_1, \ldots, n_c) \mid C^w) = \frac{\prod_{j=1}^{c} \pi(n_j, \frac{\lambda}{w})}{\pi(n, \lambda)}.$$

Then $p'(C^w \mid E^c(n_1, \ldots, n_c))$ is determined by the equation:

$$p'(C^w \mid E^c(n_1, \ldots, n_c)) = \frac{p(C^w) p'(E^c(n_1, \ldots, n_c) \mid C^w)}{\sum_{i=0}^{K-c} \binom{K-c}{i} p(C^{c+i}) p'(E^c(n_1, \ldots, n_c) \mid C^{c+i})}.$$

And

$$p'(F \mid E^c(n_1, \ldots, n_c)) = \sum_{i \in I} p'(C_i \mid E^c(n_1, \ldots, n_c))$$

where $\bigvee_{i \in I} C_i$ is the distributive normal form of F. Thus for infinite universes, p' (and p) are fully defined.

For any universe, let

$$p(E^c(n_1, \ldots, n_c)) = \frac{\sum_{i=0}^{K-c} \binom{K-c}{i} \pi(\alpha, \frac{(c+i)\lambda}{K}) \prod_{j=1}^{c} \pi(n_j, \frac{\lambda}{c+i})}{\sum_{i=0}^{K} (\binom{K}{i} \pi(\alpha, \frac{i\lambda}{K})) \pi(n, \lambda)},$$

then the representative function of p is determined by

$$p(A_{i_j}(x)(a_{n+1}/x) \mid E^c(n_1, \ldots, n_c)) = \frac{p(E^c(n_1, \ldots, n_j + 1, \ldots, n_c))}{p(E^c(n_1, \ldots, n_c))}$$

for $j \in \{1, \ldots, c\}$, and

$$p(A_{i_j}(x)(a_{n+1}/x) \mid E^c(n_1, \ldots, n_c)) = \frac{p(E^{c+1}(n_1, \ldots, n_c, 1))}{p(E^c(n_1, \ldots, n_c))}$$

for $j \notin \{1, \ldots, c\}$. For finite universes, p' (and p) are fully determined by the representative function. And the values of p and p' for infinite universes (as defined above) are the limits of the values for finite universes of size N as $N \to \infty$.

This definition and some of its consequences are discussed in Hintikka [1966] and Mondadori [1987]. I will not consider the question of how it compares to intuitive considerations regarding induction or to other definitions for inductive logic, other than to say that it can be used for a theory of inductive generalization (discussed in Hintikka [1965c] and Hintikka [1965b]), whereas Carnap's definitions can not.

Regarding the use of constituents in defining an inductive logic, Hintikka [1965c] (p. 279–283) indicates the possibility of using constituents in a similar way to that done in the monadic case to get an inductive logic for other first-order languages. This raises the following problem. One of the fundamental assumptions in inductive logic is that logically equivalent sentences should be assigned equal probabilities. Since each sentence S has a distributive normal form at its own depth and every greater depth, the probability measure p should give the distributive normal form of S at each depth the same value. In particular, if a constituent δ^d is equivalent to $\bigvee_{a \in A} \delta_a^{d+e}$, then $p(\delta^d)$ must equal $\sum_{a \in A} p(\delta_a^{d+e})$. This problem is mentioned in *ibid.* (p. 282), but I don't know if there has been any study of it.

Also, to satisfy the definition of being a probability measure, p must assign zero probability to all inconsistent constituents. But since whether or not a constituent is consistent is not generally decidable (shown in chapter 4), any probability measure defined in terms of only the consistent constituents is not computable. This is not necessarily considered to be a problem. For example, Hintikka [1965c] (p. 283, note 22) says: "The unsolvability of the problem of determining the degree of confirmation of an arbitrary generalization is no argument against my approach, however, but rather for it." But to make sense in the context of normative decision theory, the credence and confirmation functions used must be computable. However, inductive logic may be of interest from some points of view other than decision theory.

5.2 Verisimilitude

A theory of *verisimilitude* or *truthlikeness* is a theory of *closeness to the truth*, where both *closeness* and *truth* have been used with various different meanings. Verisimilitude, in this sense, was introduced by Karl Popper. His original aim was to justify the claim that science aims at truth by showing that it made sense to compare theories in terms of closeness to the truth, and thus that progress could be made by replacing a scientific theory with one that was closer to the truth (see e.g. Popper [1972]). It turned out to be rather difficult to compare scientific theories in terms of closeness to the truth, and various attempts were made to define verisimilitude for theories or sentences of propositional or first-order languages instead. The aims of a theory of verisimilitude for logical theories seem less clear than for scientific theories.

As in the previous section, we consider a monadic first-order language which contains a non-zero finite number of unary predicates $P_1, \ldots, P_k$, and a universe which is partitioned into $K = 2^k$ cells by the attributive constituents $A_1(x), \ldots, A_K(x)$. If $\mathcal{L}$ is a monadic language with k predicates and $\mathcal{M}$ is a model for $\mathcal{L}$, then exactly one constituent relative to the set of all predicates is true in $\mathcal{M}$. Verisimilitude for a monadic language is defined relative to a constituent which is taken to be the true constituent. This true constituent is *the truth* to which a distance is to be defined. This is achieved by first defining a distance between pairs of constituents. This distance is then used to define a distance between constant-free sentences and constituents, using the distributive normal form of the sentences. Then a sentence has greater verisimilitude than another sentence if its distance from the truth is smaller.

Various metrics can be defined on the set of constituents.

A constituent C_i can be represented by the set of attributive constituents $A_m(x)$ such that $(\exists x)A_m(x)$ is a conjunct of C_i, and an attributive constituent $A_m(x)$ by the set of its unnegated conjuncts. The same notation will be used for these representations as for the constituent or attributive constituent represented, and the context will make clear which is meant. The symmetric difference of sets X and Y will be denoted $X \triangle Y$.

Then the *normalized Hamming distance* between constituents C_i and C_j is given by

$$d_1(C_i, C_j) = \frac{1}{K} \sum_{m=1}^{K} \begin{cases} 0 & \text{if } A_m(x) \notin C_i \triangle C_j \\ 1 & \text{if } A_m(x) \in C_i \triangle C_j. \end{cases}$$

The metric d_1 in effect considers all attributive constituents (thus all kinds of individuals) as equally far from each other, which leads to unacceptable results in situations where there is a strong intuitive notion of similarity between individuals. (Some examples of this are given in Tichý [1978] p. 184–187, and Niiniluoto [1987] p. 315.)

The similarity between attributive constituents can be taken into account by first defining the normalized Hamming distance between attributive constituents:

$$e(A_i(x), A_j(x)) = \frac{1}{k}\sum_{m=1}^{k}\begin{cases} 0 & \text{if } P_m(x) \notin A_i(x) \triangle A_j(x) \\ 1 & \text{if } P_m(x) \in A_i(x) \triangle A_j(x). \end{cases}$$

A distance between constituents can then be defined in terms of e in various ways. For example, Niiniluoto [1987] (p. 315–316) defines the *Jyväskylä measure* as follows

$$d_J(C_i, C_j) = \sum_{m=1}^{K}\begin{cases} 0 & \text{if } A_m(x) \notin C_i \triangle C_j \\ \min_{A_u(x)\in C_i} e(A_u(x), A_m(x)) & \text{if } A_m(x) \in C_j \setminus C_i \\ \min_{A_u(x)\notin C_i} e(A_u(x), A_m(x)) & \text{if } A_m(x) \in C_i \setminus C_j. \end{cases}$$

The function d_J is not a metric since it is not symmetric, but it has been suggested (*ibid.* p. 316) that *distance from the truth* may depend on which constituent is true, and thus that d_J may still be considered as some kind of distance function. And *ibid.* (p. 316–317) defines the *weighted symmetric difference* as follows

$$d_w(C_i, C_j) = \sum_{m=1}^{K}\begin{cases} 0 & \text{if } A_m(x) \notin C_i \triangle C_j \\ \min_{A_u(x)\in C_i} e(A_u(x), A_m(x)) & \text{if } A_m(x) \in C_j \setminus C_i \\ \min_{A_u(x)\in C_j} e(A_u(x), A_m(x)) & \text{if } A_m(x) \in C_i \setminus C_j. \end{cases}$$

A number of other distance functions on constituents have been defined. (Examples can be found in Tichý [1976], Tichý [1978], Oddie [1981].)

A normalized distance function d on the set of constituents can be extended to a function giving the distance between constituents and constant-free sentences in a variety of ways. For example, if $\bigvee_{j\in J} C_j$ is the distributive normal form of F and C_i is a constituent, then (from Niiniluoto [1987], p. 214):

$$D_{\min}(C_i, F) = \min_{j\in J} d(C_i, C_j)$$

$$D_{\max}(C_i, F) = \max_{j\in J} d(C_i, C_j)$$

$$D_{\text{av}}(C_i, F) = \frac{1}{|J|}\sum_{j\in J} d(C_i, C_j)$$

$$D_{\text{sum}}(C_i, F) = \frac{\sum_{j\in J} d(C_i, C_j)}{\sum_{j\in\{1,\ldots,2^K\}} d(C_i, C_j)}$$

and *ibid.* (p. 216) gives the following distance functions which depend on parameters γ, γ':

$$D^{\gamma}_{\mathrm{mm}}(C_i, F) = \gamma D_{\min}(C_i, F) + (1-\gamma) D_{\max}(C_i, F)$$
$$D^{\gamma,\gamma'}_{\mathrm{ms}}(C_i, F) = \gamma D_{\min}(C_i, F) + \gamma' D_{\mathrm{sum}}(C_i, F)$$

where $0 \leq \gamma, \gamma' \leq 1$.

Given some distance function D between constituents and constant-free sentences, F has *greater verisimilitude* than G if

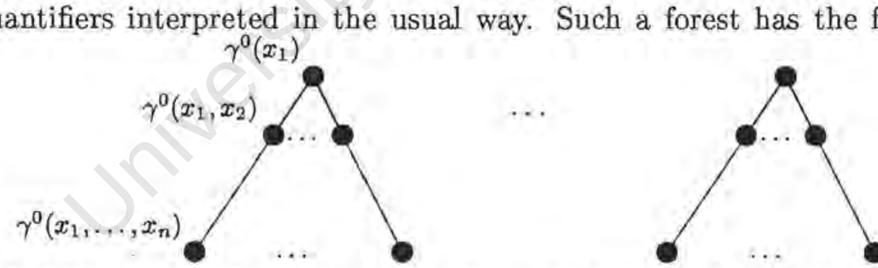
$$D(C_\star, F) < D(C_\star, G)$$

where $C_\star$ is the true constituent.

We now consider first-order languages with identity which have at least one predicate of arity greater than 1. The approach is similar to that in the monadic case, but all the definitions are relative to some fixed depth. Verisimilitude is defined for a language together with a complete theory (defined below) for that language. This complete theory is taken to be the true complete theory.

The Hamming distance between constituents of depth 0 is defined by: $e(\delta^0_i(z_1,\ldots,z_m), \delta^0_j(z_1,\ldots,z_m))$ is the number of conjuncts of $\delta^0_i(z_1,\ldots,z_m)$ which are not conjuncts of $\delta^0_j(z_1,\ldots,z_m)$.

For each constituent of the form $\delta^n()$, a set $\mathbf{S}_{\delta^n()}$ of depth-0 constituents is defined as follows: Because $\delta^n()$ has no free individual terms, it has the same form as a constituent of the corresponding language without identity, except that all quantifiers are interpreted exclusively. Thus $\delta^n()$ can be represented by the same forest as represents the corresponding constituent with the quantifiers interpreted in the usual way. Such a forest has the following form:



For each leaf of this forest, the conjunction of it and all the nodes above it is a formula of the form $\delta^0(x_1,\ldots,x_n)$. Let $\mathbf{S}_{\delta^n()}$ be the set of all these conjunctions $\delta^0(x_1,\ldots,x_n)$ of the nodes along a maximal path in $\delta^n()$. Note that $\mathbf{S}_{\delta^n()}$ is not really a representation of $\delta^n()$ because we can not get $\delta^n()$ back from $\mathbf{S}_{\delta^n()}$, as is shown by the following example: For the language with one binary predicate symbol P and identity, let

$$\begin{aligned}
\delta^2_a() &= (\exists x_1)\delta^1_b(x_1) \wedge (\exists x_1)\delta^1_c(x_1) \wedge (\forall x_1)(\delta^1_b(x_1) \vee \delta^1_c(x_1)) \\
\delta^1_b(x_1) &= P(x_1,x_1) \wedge (\exists x_2)(P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge P(x_2,x_2)) \wedge (\forall x_2)(P(x_1,x_2) \wedge \\
&\quad \neg P(x_2,x_1) \wedge P(x_2,x_2)) \\
\delta^1_c(x_1) &= P(x_1,x_1) \wedge (\exists x_2)(\neg P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2)) \wedge (\forall x_2)(\neg P(x_1,x_2) \wedge \\
&\quad P(x_2,x_1) \wedge P(x_2,x_2)).
\end{aligned}$$

Then $\mathbf{S}_{\delta^2_a()} = \{P(x_1,x_1) \wedge P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge P(x_2,x_2),\ P(x_1,x_1) \wedge \neg P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2)\}$. Let

$$\begin{aligned}
\delta^2_d() &= (\exists x_1)\delta^1_e(x_1) \wedge (\forall x_1)\delta^1_e(x_1) \\
\delta^1_e(x_1) &= P(x_1,x_1) \wedge (\exists x_2)(P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge P(x_2,x_2)) \wedge (\exists x_2)(\neg P(x_1,x_2) \wedge \\
&\quad P(x_2,x_1) \wedge P(x_2,x_2)) \wedge (\forall x_2)((P(x_1,x_2) \wedge \neg P(x_2,x_1) \wedge P(x_2,x_2)) \vee \\
&\quad (\neg P(x_1,x_2) \wedge P(x_2,x_1) \wedge P(x_2,x_2))).
\end{aligned}$$

Then $\mathbf{S}_{\delta^2_d()} = \mathbf{S}_{\delta^2_a()}$, although $\delta^2_a()$ and $\delta^2_d()$ are not equivalent. In fact, $\delta^2_a()$ is consistent and $\delta^2_d()$ is inconsistent. For a constituent $\delta^n()$, the set $\mathbf{S}_{\delta^n()}$ describes the possible sequences of n individuals that can be chosen from a universe without replacement (i.e. the same individual can not be chosen more than once) in a model in which $\delta^n()$ is true. Each individual is described by the predicates it satisfies on its own and together with the previous individuals in the sequence. But, as the above example shows, this set of sequences is not sufficient to determine the constituent.

The following definitions are from Tichý [1976]. The distance between two constituents $\delta^n_i()$ and $\delta^n_j()$ is defined in terms of $\mathbf{S}_{\delta^n_i()}$ and $\mathbf{S}_{\delta^n_j()}$.

A *linkage* between two sets is a surjection from the larger set onto the smaller one (if they are the same size, from either one onto the other). If l is a linkage between $\mathbf{S}_{\delta^n_i()}$ and $\mathbf{S}_{\delta^n_j()}$, the *breadth* of l is the average distance between linked elements:

$$B(l) = \frac{1}{|l|} \sum_{\langle \delta^0_u(x_1,\ldots,x_n),\delta^0_v(x_1,\ldots,x_n)\rangle \in l} e(\delta^0_u(x_1,\ldots,x_n), \delta^0_v(x_1,\ldots,x_n)).$$

Then, the distance between constituents of depth n is defined by:

$$d_n(\delta^n_i(), \delta^n_j()) = \text{the breadth of the narrowest linkage between } \mathbf{S}_{\delta^n_i()} \text{ and } \mathbf{S}_{\delta^n_j()}.$$

And the distance of a constant-free sentence F of depth n from a constituent $\delta^n_i()$ is given by

$$D_n(\delta^n_i(), F) = \frac{1}{|J|} \sum_{j \in J} d_n(\delta^n_i(), \delta^n_j())$$

where $\bigvee_{j \in J} \delta^n_j()$ is the distributive normal form$_0$ of F at depth n. This particular distance function D_n gives the average distance between a set of constituents and a constituent, and is a generalization of the function D_{av} defined above for the monadic case. Because d_n is defined in terms of the sets of sequences of individuals allowed by the constituents, a significant amount of information given by the constituents is ignored. This leads to some unsatisfactory results. Tichý [1978] gives a modified definition of *linkage* which takes into account the tree-structure of constituents and requires that a linkage between two constituents preserve the predecessor/successor relation of the trees or forests. There are also a number of other distance functions which have been defined on the sets of constituents of each depth (e.g. Niiniluoto [1987], p. 349). Once the distance functions on the constituents are defined, a distance function between sets of constituents and constituents can be defined in various ways, as in the monadic case.

The relative verisimilitude of constant-free sentences F and G of depth n can be defined as follows: F has *greater verisimilitude* than G if

$$D_n(\delta_\star^n(), F) < D_n(\delta_\star^n(), G)$$

where $\delta_\star^n()$ is the true constituent of depth n. And sentences of different depths can be compared by using the distance function D_n where n is the greater depth and the distributive normal form$_0$s of both sentences at depth n. (The relative verisimilitude of sentences of different depths is defined like this by Tichý [1976], p. 33.) Regarding the dependence of the distance functions on depth, it seems to me that we have two options. Either we can accept that in expanding sentences to a greater depth, their relative verisimilitude may be reversed (which is not surprizing since at different depths they are compared with different truths). Then equivalent sentences may have different verisimilitudes, so we can not compare the verisimilitude of a sentence with one of a greater depth by using its expansion to that greater depth. So we have no way of comparing the verisimilitudes of sentences of different depths. Or (if we want to use definitions relative to depth to compare sentences of different depths) we can try to find a definition for which the relative verisimilitude of sentences does not change as they are expanded to greater depths. As far as I know, this has not been done. But it seems that if we want to compare sentences of different depths, it would be better to use a definition which is not relative to depth.

At a fixed depth, the constituents are the strongest sentences, but when we consider all the depths together there are not any strongest sentences. For each model for a language, exactly one *complete theory* (as defined below) is true. If a distance function on complete theories is defined, then they can be used in a way similar to how constituents were used above.

A *theory* $\mathbf{T}$ of a first-order language $\mathcal{L}$ is a set of sentences of $\mathcal{L}$ that is closed under logical consequence, that is, if $\mathbf{T} \Longrightarrow X$ for some sentence X of $\mathcal{L}$, then $X \in \mathbf{T}$. A theory $\mathbf{T}$ is *complete* if for each sentence X, either $\mathbf{T} \Longrightarrow X$ or $\mathbf{T} \Longrightarrow \neg X$. A subset $\mathbf{B} \subseteq \mathbf{T}$ is said to *generate* the theory $\mathbf{T}$ if $\mathbf{T} = \{X \mid X \text{ is a sentence and } \mathbf{B} \Longrightarrow X\}$. A sequence of constituents $\{\delta^n() \mid n \in \mathbb{N}\}$ such that

$$\ldots \delta^{n+1}() \Longrightarrow \delta^n() \Longrightarrow \ldots \Longrightarrow \delta^1()$$

is called *monotone*. A theory is complete iff it is generated by a monotone sequence of constituents. (This follows from the proof of a similar result in Niiniluoto [1987], p. 77.) For any model, the true constituents form a monotone sequence. The theory (which is complete) generated by the set of true constituents of each depth will be *the truth* with which other theories will be compared.

If, for each $n \in \mathbb{N}$, d_n is a distance function on the constituents of depth n, then a distance function d on complete theories can be defined by (Niiniluoto [1987], p. 362)

$$d(\mathbf{T}_1, \mathbf{T}_2) = \lim_{n \to \infty} d_n(\delta^n_{i_1}(), \delta^n_{i_2}())$$

where $\mathbf{T}_1$ is generated by the set of constituents $\{\delta^n_{i_1}() \mid n \in \mathbb{N}\}$ and $\mathbf{T}_2$ by the set of constituents $\{\delta^n_{i_2}() \mid n \in \mathbb{N}\}$. The function d can then be extended to a distance function between arbitrary theories and complete theories (Niiniluoto [1987], p. 368–371).

If the functions d_n are metrics, the function d need only be a pseudometric (that is, we may have $d(\mathbf{T}_1, \mathbf{T}_2) = 0$ for some theories $\mathbf{T}_1 \neq \mathbf{T}_2$). For this reason, it may not be satisfactory to define comparative verisimilitude as follows: $\mathbf{T}_1$ has greater verisimilitude than $\mathbf{T}_2$ if

$$d(\mathbf{T}_\star, \mathbf{T}_1) < d(\mathbf{T}_\star, \mathbf{T}_2)$$

where $\mathbf{T}_\star$ is the true complete theory. And Niiniluoto [1987] (p. 363) gives the following definition instead: $\mathbf{T}_1$ has greater verisimilitude than $\mathbf{T}_2$ if there is some $n_0 \in \mathbb{N}$ such that

$$d_n(\delta^n_\star(), \delta^n_{i_1}()) < d_n(\delta^n_\star(), \delta^n_{i_2}())$$

for all $n \geq n_0$, where the true complete theory is generated by $\{\delta^n_\star() \mid n \in \mathbb{N}\}$, $\mathbf{T}_1$ by $\{\delta^n_{i_1}() \mid n \in \mathbb{N}\}$, and $\mathbf{T}_2$ by $\{\delta^n_{i_2}() \mid n \in \mathbb{N}\}$.

The kind of definitions mentioned above don't seem to be of much use for the problem of defining verisimilitude for scientific theories, particularly if we want some practical way of estimating (comparative) verisimilitude of theories. There are a number of reasons for this. For example, in all of the above definitions, the only way the verisimilitude of theories can be compared is by knowing the truth and making some suitable comparison with it of each of the theories to be compared. If, in some situation, we know what the truth is, then the idea of verisimilitude is not of much interest. We could just use the theory which is the truth, and not consider any others. But in practice we don't know what the truth is. The only way we have of finding out about it is by observation which is incomplete and inaccurate. Niiniluoto [1987] (chapter 7) has proposed a way of estimating verisimilitude relative to evidence, given some definition for verisimilitude where the truth is known. However, it uses a probability measure (as defined in the previous section), where the corresponding conditional probability function is supposed to represent rational degree of belief in statements of the form "C_i is the true constituent" given some evidence. What is achieved is to convert some definition of verisimilitude to a probabilistic version, where the evidence is used to determine the probabilities, but the problem of defining a probability measure remains. In fact, Niiniluoto [1987] (p. 278) says "... the epistemic problem of truthlikeness is equally difficult as the traditional problem of induction", which is of course true for Niiniluoto's particular approach to this problem, but there may be other approaches not involving probability measures. The problem of defining verisimilitude in a way that is applicable to scientific theories such that it can be estimated on evidence remains open.

Some remaining questions

1. Probably the most significant open problem brought to light by this thesis is the question of whether or not Hintikka's definitions of trivial inconsistency (particularly the one I have referred to as *one-trivial inconsistency*) are adequate for the completeness theorem of the theory of distributive normal forms to hold. A comparison of one-trivial and two-trivial inconsistency may be helpful. Although these two kinds of inconsistency are not equivalent, it is not known whether there is an infinite sequence of constituents where the successor of each element is in that element's expansion such that no element of the sequence is one-trivially inconsistent, but some element is two-trivially inconsistent. If there is, then the completeness result Hintikka claims is false, otherwise it is true.

2. In section 4.4, we saw that most constituents are inconsistent. Of these, many are trivially inconsistent, which means that they can be found to be inconsistent by a particular algorithm. It would be interesting to try to find a lower bound for the fraction of not-(trivially-inconsistent) constituents that are inconsistent, for some known definition of trivial inconsistency (such as one-trivial inconsistency or two-trivial inconsistency as defined in chapter 4). It might even be possible to find some results which apply to *all* algorithms for finding inconsistent constituents, regarding the fraction of inconsistent constituents (of suitably great depths) that the algorithm wouldn't find.

3. In considering the possible uses of distributive normal forms and constituents, it is clear that we can not expect applications which actually convert formulas to distributive normal form to use for some calculation or comparison. However, just the existence of a normal form for first-order logic could be useful for getting general results about all formulas. The only result of this nature that I know of is the completeness theorem of the theory of distributive normal forms, which provides a way of showing the completeness of first-order logic. Perhaps distributive normal forms can also be used to prove some general results which haven't been previously proved by other methods. In this regard, it might be interesting to consider what the existence of the distributive normal form means from an algebraic point of view.

Bibliography

Aho, A. V. and J. D. Ullman [1992]. *Foundations of Computer Science.* W. H. Freeman and Co.

Barwise, J., editor [1977]. *Handbook of Mathematical Logic,* Volume 90 of *Studies in Logic and the Foundations of Mathematics.* North-Holland.

Bell, J. L. and M. Machover [1977]. *A Course in Mathematical Logic.* North-Holland.

Bell, J. L. and A. B. Slomson [1971]. *Models and Ultraproducts: An Introduction.* Amsterdam, North-Holland.

Bogdan, R. J., editor [1987]. *Jaakko Hintikka,* Volume 8 of *Profiles: An Internation Series on Contemporary Philosophers and Logicians.* D. Reidel.

Boolos, G. and R. Jeffrey [1974]. *Computability and Logic.* Cambridge University Press.

Brink, C. [1989]. Verisimilitude: Views and reviews. *History and Philosophy of Logic* **10**, 181–201.

Carnap, R. [1956]. *Meaning and Necessity,* Enlarged edition. The University of Chicago Press.

Carnap, R. [1971]. A basic system of inductive logic, Part I. In Carnap and Jeffrey [1971].

Carnap, R. [1973]. Notes on probability and induction. *Synthese* **25**, 269–298. (from UCLA, Philosophy 242, Fall semester 1955, edited for publication by A. Benson).

Carnap, R. and R. C. Jeffrey, editors [1971]. *Studies in Inductive Logic and Probability,* Volume 1. University of California Press.

Chang, C. and R. C. Lee [1973]. *Symbolic Logic and Mechanical Theorem Proving.* Computer Science and Applied Mathematics. Academic Press.

Chang, C. C. and H. J. Keisler [1973]. *Model Theory,* Volume 73 of *Studies in Logic and the Foundations of Mathematics.* North-Holland.

Church, A. [1956]. *Introduction to Mathematical Logic,* Volume 1. Princeton University Press.

Cohen, D. I. A. [1991]. *Introduction to Computer Theory,* Revised edition. John Wiley & Sons, Inc.

Fine, K. [1975]. Normal forms in modal logic. *Notre Dame Journal of Formal Logic* **XVI**(2), 229–237.

Fitting, M. [1990]. *First-Order Logic and Automated Theorem Proving.* Texts and Monographs in Computer Science. Springer-Verlag.

Halmos, P. R. [1950]. *Measure Theory.* The University Series in Higher Mathematics. D. Van Nostrand Company, Inc.

Hamilton, A. G. [1988]. *Logic for Mathematicians*, Revised edition. Cambridge University Press.

Harel, D. [1992]. *The Spirit of Computing*, Second edition. Addison-Wesley.

Hilbert, D. and W. Ackermann [1950]. *Principles of Mathematical Logic.* Chelsea Publishing Company.

Hintikka, J. [1953]. Distributive normal forms in the calculus of predicates. *Acta Philosophica Fennica* **VI**, 1–71.

Hintikka, J. [1955]. Form and content in quantification theory. *Acta Philosophica Fennica* **VIII**, 7–55.

Hintikka, J. [1965a]. Distributive normal forms in first order logic. In J. Crossley and M. Dummett, editors, *Formal Systems and Recursive Functions (SLFM proceedings)*, pp. 48–91. North-Holland. Reprinted in Hintikka [1973d].

Hintikka, J. [1965b]. On a combined system of inductive logic. *Acta Philosophica Fennica* **18**, 21–30.

Hintikka, J. [1965c]. Towards a theory of inductive generalization. In Y. Bar-Hillel, editor, *Logic, Methodology and Philosophy of Science: Proceedings of the 1964 International Congress*, pp. 274–288. North-Holland.

Hintikka, J. [1966]. A two-dimensional continuum of inductive methods. In Hintikka and Suppes [1966].

Hintikka, J. [1970]. Surface information and depth information. In Hintikka and Suppes [1970].

Hintikka, J. [1973a]. An analysis of analyticity. In Hintikka [1973d].

Hintikka, J. [1973b]. Are logical truths tautologies? In Hintikka [1973d].

Hintikka, J. [1973c]. Logic in philosophy — philosophy of logic. In Hintikka [1973d].

Hintikka, J. [1973d]. *Logic, Language-Games and Information.* Oxford, Clarendon Press.

Hintikka, J. [1973e]. On the different ingredients of an empirical theory. In P. Suppes, L. Henkin, A. Joja, and G. C. Moisil, editors, *Logic, Methodology and Philosophy of Science IV*, pp. 313–322. North-Holland.

Hintikka, J. and P. Suppes, editors [1966]. *Aspects of Inductive Logic.* Studies in Logic and the Foundations of Mathematics. North-Holland.

Hintikka, J. and P. Suppes, editors [1970]. *Information and Inference.* D. Reidel.

Hodel, R. E. [1995]. *An Introduction to Mathematical Logic.* PWS Publishing Company.

Hughes, G. E. and M. J. Cresswell [1968]. *An Introduction to Modal Logic.* London, Methuen and Co Ltd.

Hughes, G. E. and D. G. Londey [1965]. *The Elements of Formal Logic.* Methuen & Co Ltd.

Jeffrey, R. C., editor [1980]. *Studies in Inductive Logic and Probability,* Volume 2. University of California Press.

Kieseppä, I. A. [1996a]. On the aim of the theory of verisimiltude. *Synthese* **107**, 421–438.

Kieseppä, I. A. [1996b]. *Truthlikeness for Multidimensional, Quantitative Cognitive Problems,* Volume 254 of *Synthese Library.* Kluwer Academic Press.

Kingman, J. F. C. and S. J. Taylor [1966]. *Introduction to Measure and Probability.* Cambridge University Press.

Kneebone, G. T. [1963]. *Mathematical Logic and the Foundations of Mathematics.* D. Van Nostrand Company, Ltd.

Lakatos, I., editor [1968]. *The Problem of Inductive Logic, Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, volume 2.* Studies in Logic and the Foundations of Mathematics. North-Holland.

Mendelson, E. [1987]. *Introduction to Mathematical Logic,* Third edition. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth, Inc.

Mondadori, M. [1987]. Hintikka's inductive logic. In Bogdan [1987].

Niiniluoto, I. [1987]. *Truthlikeness,* Volume 185 of *Synthese Library.* D. Reidel.

Oddie, G. [1981]. Verisimilitude reviewed. *The British Journal for the Philosophy of Science* **32**, 237–265.

Padmanabhan, T. [1993]. *Structure Formation in the Universe.* Cambridge University Press.

Popper, K. R. [1972]. *Objective Knowledge.* Oxford University Press.

Rantala, V. [1987]. Constituents. In Bogdan [1987].

Robbin, J. W. [1969]. *Mathematical Logic: A First Course.* University Mathematics Series. W. A. Benjamin, Inc.

Rogers, R. [1971]. *Mathematical Logic and Formalized Theories.* North-Holland.

Saarinen, E., R. Hilpinen, I. Niiniluoto, and M. P. Hintikka, editors [1979]. *Essays in Honour of Jaakko Hintikka,* Volume 124 of *Synthese Library: Studies in Epistemology, Logic, Methodology, and Philosophy of Science.* D. Reidel Publishing Co.

Shoenfield, J. R. [1967]. *Mathematical Logic*. Addison-Wesley Series in Logic. Addison-Wesley.

Smullyan, R. M. [1968]. *First-Order Logic*, Volume 43 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag.

Tichý, P. [1976]. Verisimilitude redefined. *The British Journal for the Philosophy of Science* **27**, 25–42.

Tichý, P. [1978]. Verisimilitude revisited. *Synthese* **38**, 175–196.

van Benthem, J. [1983]. *Modal Logic and Classical Logic*, Volume III of *INDICES Monographs in Philosophical Logic and Formal Linguistics*. Bibliopolis.

Index

Symbol table

The following symbols are not defined in the contexts in which they are used:

$\mathbf{2}$	the 2-element Boolean algebra $\{0,1\}$ with 1 as the top and 0 as the bottom
$\emptyset$	the empty set
$\in$	element
$\subseteq$	subset
$\supseteq$	superset
$\cap$	binary intersection
$\cup$	binary union
$\setminus$	set difference
$\bigcap$	intersection
$\bigcup$	union
$\|X\|$ for a set X	the cardinality of X
$\mathbb{N}$	the set of natural numbers $\{1,2,\ldots\}$
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^+$	$\{x \in \mathbb{R} \mid x > 0\}$
$\mathcal{P}(X)$ for a set X	the powerset of X
$\sum$	sum
$\prod$	product

The following symbols are defined with their first uses:

$\neg$	p. 4
$\vee$	p. 4
$\wedge$	p. 4
$\rightarrow$	p. 4
$\leftrightarrow$	p. 4
$\models$	p. 5
$\Longrightarrow$	p. 6
$\Longleftrightarrow$	p. 6
$\bigvee$	p. 6