Notre Dame Journal of Formal Logic
Volume 33, Number 1, Winter 1992

Surface Reasoning

WILLIAM C. PURDY

Abstract Surface reasoning is defined to be deduction conducted in a surface language in terms of certain primitive logical relations. A surface language is a spoken or written natural language (in this paper, English), in contrast to a "base language" or "deep structure" sometimes hypothesized to explain natural language phenomena. The primitive logical relations are inclusion, exclusion, and overlap between classes of entities. A language and a calculus for representing surface reasoning is presented. Then a paradigm for reasoning in this calculus is developed. This paradigm is similar to but more general than syllogistic. Reasoning is represented as construction of fragments (subposets) of lattices. Elements of the lattices are expressions denoting classes of individuals. Strategies to streamline the reasoning process are proposed.

1 Introduction The underlying premise of this paper is that the relations of inclusion, exclusion, and overlap are primitive and important constructs of human reasoning; that the surface language is adequate to express and manipulate these constructs; and that disparate logics and complex transformations linking them to the surface language are not necessary to explain language understanding and reasoning. The *surface language*, used in relation to a natural language such as English, is the spoken or written language, in contrast to a "base language" or "deep structure" sometimes hypothesized to explain natural language phenomena. *Surface reasoning* is defined to be deduction conducted in the surface language, in terms of the primitive relations, inclusion, exclusion, and overlap, involving surface information (in the sense of Hintikka [6]). It is hypothesized that surface reasoning captures the essence of human reasoning. Therefore its study is a fruitful approach to understanding and implementing cognitive agents.

This contrasts with the *de facto* standard approach to automated reasoning, viz., deduction conducted in a disparate language, clausal-form logic, employing unification and resolution, typically performed in depth-first order. It is not suggested that surface reasoning can supplant conventional forms of formal

Received September 25, 1990; revised June 6, 1991

logic, but rather that surface reasoning provides a better description of natural language understanding and reasoning as it is practiced by humans. Indeed, as the reasoning process reaches greater depth, the conventional techniques of formal logical analysis seem more appropriate.

Inclusion, exclusion, and overlap are formalized as the categorical statements A, E, and I, respectively, of syllogistic. Syllogistic is therefore an ideal system in which to represent surface reasoning. However, syllogistic is limited in scope to monadic logic. A previous paper (Purdy [9]) introduced a polyadic logic which shares the characteristics of syllogistic. The present paper extends this logic and investigates surface reasoning in the context of the extended logic.

The principal results presented in this paper are the following.

(i) $\mathcal{L}_N$, a language in which to represent surface reasoning, is defined. This language shares many characteristics with natural language, including the use of certain generalized quantifiers.
(ii) The language is axiomatized and completeness properties are given.
(iii) Derived rules of inference are obtained. Principal among these are the Monotonicity Rules.
(iv) A paradigm for surface reasoning similar to syllogistic is developed. Graphical domains are described in which reasoning finds a natural representation. The reasoning process is represented as construction (or search) of fragments of these domains.
(v) Global strategies to streamline reasoning within this paradigm are proposed. These strategies impose global restrictions or preferences regarding the inference rules to be used.
(vi) A local strategy is described. This strategy provides guidance in selecting the direction in which construction is to proceed.

2 *Informal definition of* $\mathcal{L}_N$ This section introduces the concepts underlying the language $\mathcal{L}_N$. A formal definition is given in the next section.

2.1 Boolean operators Every expression of $\mathcal{L}_N$ denotes a set of n-tuples over a domain $\mathfrak{D}$, where $n \in \omega$ is the *arity* of the expression. (For simplicity, an expression and its denotation are represented by the same symbol.) For example, the predicate **like** denotes a set of 2-tuples $\langle p, o \rangle$, such that p likes o. Hence the logical types of $\mathcal{L}_N$ are indexed by ω. Boolean operators $\cap$ and $^-$ denote the operations intersection and complementation on these sets. When $n = 0$, the operations correspond to conjunction and negation, respectively. Intersection of a set of n-tuples and a set of m-tuples $(n \leq m)$ is defined:

$$X \cap Y := \{\langle y_1, \ldots, y_m \rangle : \langle y_1, \ldots, y_n \rangle \in X \wedge \langle y_1, \ldots, y_m \rangle \in Y\}.$$

When $n = m$, intersection is said to be *homogeneous*. Complementation of a set of n-tuples is defined:

$$\bar{Y} := \{\langle y_1, \ldots, y_n \rangle : \langle y_1, \ldots, y_n \rangle \notin Y\}.$$

2.2 The image operator The image operator embodies the overlap relation. It appears in the works of DeMorgan, Peirce, and Frege, but was given its modern form, $Y``X$, by Whitehead and Russell (see Quine [10]). For a set of 2-tuples $Y$ and a set of 1-tuples $X$, the image of $X$ by $Y$ is defined $Y``X := \{\langle y \rangle : \exists x(\langle y, x\rangle \in Y \wedge \langle x \rangle \in X)\}$. $\mathfrak{L}_N$ incorporates a variation appropriate to natural language, **some**XY, which denotes $\breve{Y}``X$. Thus **some boy like** denotes the set of things that some boy likes.

For two unary expressions X and Y, the denotation of **some**XY particularizes to $\{\langle\ \rangle : \exists x(\langle x \rangle \in Y \wedge \langle x \rangle \in X)\}$, that is, 'some X is Y' (overlap). Similarly, $\overline{\mathbf{some}XY}$ renders 'no X is Y' (exclusion), and $\overline{\mathbf{some}X\overline{Y}}$ renders 'no X is non-Y' or 'all X is Y' (inclusion). Thus all three primitive logical relations can be expressed by means of the image operator.

The image operator readily generalizes to an operator on an $(n+1)$-ary expression and a unary expression, yielding an n-ary expression: **some**XY denotes $\{\langle y_1, \ldots, y_n\rangle : \exists x(\langle x \rangle \in X \wedge \langle x, y_1, \ldots, y_n\rangle \in Y)\}$. Thus if **give** is a ternary predicate taking a donor, an object, and a recipient, **some boy give** denotes the set of pairs $\langle o, r\rangle$ such that some boy gives o to r.

The image operator also can be generalized to operate on an n-ary expression and an m-ary expression $(m \leq n)$, yielding an $(n-m)$-ary expression, but this seems inappropriate for application to natural language.

The appropriateness of the image construction defined above to natural language is supported by considerations related to θ-theory and case theory. Verbs are viewed as relations whose arguments are bound by noun phrases. Continuing the previous example, 'some boy gives a flower to a girl' is rendered by **some girl some flower some boy give**. The first argument of **give** is bound by **boy**, the second argument by **flower**, and the third argument by **girl**. Notice that arguments are filled from right to left, while scope is given from left to right. (The subexpression with the widest scope is on the left.)

someXY asserts that the overlap between X and Y is nonempty, but leaves its nature otherwise unspecified. The image operator can be further refined to cover a broader range of natural language quantifiers. Consider the refinement **k**XY with the denotation:

$$\{\langle y_1, \ldots, y_n\rangle : card(\{x : \langle x \rangle \in X \wedge \langle x, y_1, \ldots, y_n\rangle \in Y\}) \geq k\}.$$

Then **k**XY asserts that the overlap is not only nonempty, but also of cardinality at least k.

Refining the image still further, define **most**XY to have the denotation:

$$\{\langle y_1, \ldots, y_n\rangle : card(\{x : \langle x \rangle \in X \wedge \langle x, y_1, \ldots, y_n\rangle \in Y\}) \\ > card(\{x : \langle x \rangle \in X \wedge \langle x, y_1, \ldots, y_n\rangle \notin Y\})\}.$$

Then **most**XY asserts an overlap between X and Y that is nonempty and involves the majority of X.

Thus refined, the image operator allows most of the generalized quantifiers of natural language to be represented in $\mathfrak{L}_N$, either directly or by abbreviation.

2.3 The selection operator A *selection operator* is a tuple $\langle k_1, \ldots, k_m\rangle$ of positive integers, which operates on predicates of arity m. It plays the role played by bound variables in predicate logic. Its action can be illustrated by the following example.

$$\mathbf{some}X_4\mathbf{some}X_3\mathbf{some}X_2\mathbf{some}X_1\langle 4,1,3\rangle R^3 \equiv \mathbf{some}X_3\mathbf{some}X_1\mathbf{some}X_4R^3.$$

Selection operators subsume Quine's predicate functors, Inv, inv, Ref, Pad (Quine [12]):

$$\begin{aligned} \text{Inv} &= \langle n,1,\ldots,n-1\rangle \\ \text{inv} &= \langle 2,1,3,\ldots,n\rangle \\ \text{Ref} &= \langle 1,1,2,\ldots,n-1\rangle \\ \text{Pad} &= \langle 2,3,\ldots,n+1\rangle. \end{aligned}$$

Unlike Quine's predicate functors, selection operators are restricted to operate on predicate symbols only. This restriction does not affect expressiveness. It is motivated by appropriateness to natural language.

Selection operators are used to render natural language constructs such as:

Passive: $\langle 2,1\rangle$ **kiss** = **be-en kiss by**
Reflexive: $\langle 1,1\rangle$ **shave** = **shave self.**

In the first example of this subsection, the selection operator selected three of four arguments, ignoring the remaining one. Such selection operators will be termed *vacuous*. In the subsequent examples, the selection operators are nonvacuous. This seems to be the case in all natural language applications. It would be possible to restrict selection operators to be nonvacuous. However, vacuous selection operators will be allowed so that $\mathcal{L}_N$ will be equivalent to predicate calculus. It would also be possible to distinguish selection operators that: (i) only permute arguments, (ii) only identify arguments, and (iii) only skip arguments, using them in various combinations. In the interest of simplicity, this possibility will not be pursued. However, *restriction* of selection operators and its impact on reasoning strategies will be considered in a later section.

2.4 Singular terms and the identity relation The identity is defined to be a reflexive, symmetric, transitive, binary relation that obeys Leibniz' Law, that identicals are indiscernible and indiscernibles are identical. It is well known that the identity can be axiomatized in a first-order language as follows (e.g., Mendelson [7]):

(i) $\forall x(x = x)$
(ii) $\forall x \forall y((x = y) \rightarrow (\phi(x,x) \rightarrow \phi(x,y))$, where ϕ ranges over all well-formed expressions.

It is also well known that a first-order axiomatization constrains the denotation of the identity to be an equivalence relation only, while the intuitive expectation is that the identity denotes the diagonal relation, $\{\langle x,x\rangle : x \in \mathfrak{D}\}$. Two remedies are available, both having the same result. First, the identity can be declared part of the logical vocabulary and required to denote the diagonal relation in every interpretation. Second, consideration can be restricted to so-called *normal* interpretations, in which the identity is interpreted as the diagonal relation. Hence the identity justifiably is considered somewhat troublesome.

Sommers [14] denies the need for the identity relation. He claims that by endowing singular terms, such as **Socrates**, with "wild quantity," that is, taking **some Socrates is wise** to be equivalent to **all Socrates is wise** one can enjoy the benefits of the identity without the identity. Since this suggestion is followed in $\mathfrak{L}_N$, and since Sommers' position is surrounded by controversy and perhaps misunderstanding, some discussion is warranted.

The nonlogical vocabulary of $\mathfrak{L}_N$ includes a supply of predicate symbols $R_i^n (i \in \omega)$ of all arities $n \in (\omega - \{0\})$. In addition, a supply of *singular predicate symbols* S_i^{n+1} $(i \in \omega)$ of arity $n + 1$ for all $n \in \omega$ are provided. Singular predicates have a *quasi-logical* status in the following sense. While the denotation of a singular predicate may vary from one interpretation to another, it must possess the following property in every interpretation.

$$\forall x_1, \ldots, x_n \in \mathfrak{D}(\exists x \in \mathfrak{D}(\langle x_1, \ldots, x_n, x\rangle \in S_i^{n+1} \\ \wedge \forall x' \in \mathfrak{D}(\langle x_1, \ldots, x_n, x'\rangle \in S_i^{n+1} \rightarrow (x = x')))).$$

The *singular expressions* of $\mathfrak{L}_N$ are defined as follows:

(i) each unary singular predicate symbol S^1 is a singular expression
(ii) if S^{n+1} is an $(n + 1)$-ary singular predicate symbol, and $S_1, \ldots, S_n$ are singular expressions, then $\mathbf{some}S_n \cdots \mathbf{some}S_1 S^{n+1}$ is a singular expression
(iii) nothing else is a singular expression.

The quasi-logical status of singular predicates, along with the axiom schema $\mathbf{some}S\overline{X} \equiv \overline{\mathbf{some}SX}$, where S ranges over singular expressions and X ranges over $(n + 1)$-ary $(n \in \omega)$ expressions of $\mathfrak{L}_N$, ensures that $\mathbf{some}S_iS_j$ is always interpreted as identity of the denotations of singular expressions S_i and S_j. For example, if **Socrates** is a unary singular predicate and **father** is a binary singular predicate, then **some Socrates father** is a singular expression that denotes the (unique) individual which is the father of Socrates.

This device is not unique to $\mathfrak{L}_N$ or Sommers' Term Calculus. It is available to first-order predicate calculus as well. Seeing how might clarify this matter. Let the predicate calculus be equipped with the usual nonlogical constants, with the following slight modification. Function symbols and constants are replaced by singular predicate symbols defined as above, with the same quasi-logical property. Singular expressions are defined as follows:

(i) if S^1 is a unary singular predicate symbol, then $\lambda x(S^1(x))$ is a singular expression
(ii) if S^{n+1} is an $(n + 1)$-ary singular predicate symbol, and $S_1, \ldots, S_n$ are singular expressions, then $\lambda x(\exists x_1 \cdots \exists x_n(S^{n+1}(x_1, \ldots, x_n, x) \wedge S_1(x_1) \wedge \cdots \wedge S_n(x_n))$ is a singular expression
(iii) nothing else is a singular expression.

The following axiom schema is needed.

$$\exists x(S(x) \wedge \phi(x)) \leftrightarrow \forall x(S(x) \rightarrow \phi(x)),$$

where S ranges over singular expressions and ϕ ranges over well-formed formulas. Now it is easy to see that $\exists x(S_i(x) \wedge S_j(x))$ is always interpreted as identity of the denotations of the singular expressions S_i and S_j.

Several observations are in order. First, Sommers' treatment of identity is not radically different from the conventional treatment. Both require a restriction placed on the notion of interpretation. The axiom schema required by Sommers' treatment is quite similar to the axiom for Leibniz' Law, except that the former entails the reflexivity property.

Second, Sommers' treatment does not provide all the benefits of the identity. For example, $\exists x(X(x) \wedge \forall y(X(y) \rightarrow y = x))$ can be expressed only by a schema (i.e., infinite conjunction) such as **something** $X \cap \overline{(\mathbf{some}SX \cap \mathbf{some}X\bar{S})}$. Whenever identity of quantified variables is expressed in predicate calculus, the translation in $\mathcal{L}_N$ must be a schema.

Third, Sommers' treatment nonetheless seems particularly suited to variable-free logics for natural language. The use of schematic variables representing singular expressions is quite similar to the use of pronouns in natural language. More important, Sommers' treatment makes identity a special case of the more general principle of monotonicity.

2.5 Discussion $\mathcal{L}_N$ is intended to be a vehicle for study and implementation of surface reasoning. Other contenders include predicate calculus, Logic with Generalized Quantifiers (Barwise [1]), Predicate Functor Logic (Quine [11]), Term Calculus (Sommers [15]), combinatory logic (Steedman [16]), and various varieties of lambda calculus. Each can claim adherents.

In its favor $\mathcal{L}_N$ can claim a combination of features possessed by none of them:

- Structural similarity to surface English
  - a 'well-translatable' (Čulík [5]) English fragment suitable for reasoning can be defined (for an example, see [9])
  - includes 'singular relations,' i.e., singular predicates of arity greater than one
  - includes the generalized quantifiers of natural language
  - implicitly many-sorted
- Appropriate to the task of representing surface reasoning
  - a reasoning paradigm similar to syllogistic can be defined
  - has the expressiveness of predicate calculus, but not more expressiveness than necessary for its intended task
- Simplicity
  - variable-free
  - no special identity relation
  - simple type structure.

3 Formal definition of $\mathcal{L}_N$ The vocabulary of $\mathcal{L}_N$ consists of the following (let $\omega_+ := \omega - \{0\}$):

1. Predicate symbols $\mathcal{P} = (\bigcup_{n\in\omega_+} \mathcal{R}_n) \cup (\bigcup_{n\in\omega_+} \mathcal{S}_n)$, where $\mathcal{R}_n = \{R^n_i : i \in \omega\}$ and $\mathcal{S}_n = \{S^n_i : i \in \omega\}$.
2. Selection operators $\{\langle k_1, \dots, k_n\rangle : n \in \omega_+,\ k_i \in \omega_+,\ 1 \le i \le n\}$.
3. Quantifiers **some**, $\{\mathbf{k} : \mathbf{k} \in \omega_+\}$, and **most**.
4. Boolean operators $\cap$ and $^-$.
5. Parentheses (and).

$\mathcal{L}_N$ is partitioned into sets of n-ary expressions for $n \in \omega$. These sets are defined to be the smallest satisfying the following conditions.

1. For all $n \in \omega_+$, each $S_i^n \in \mathcal{S}_n$ is an n-ary expression.
2. For all $n \in \omega_+$, each $R_i^n \in \mathcal{R}_n$ is an n-ary expression.
3. For each predicate symbol $P \in \mathcal{P}$ of arity m, $\langle k_1, \ldots, k_m \rangle P$ is a n-ary expression where $n = max\ (k_i)_{1 \leq i \leq m}$.
4. If X is an n-ary expression then $\overline{(X)}$ is an n-ary expression.
5. If X is an m-ary expression and Y is an l-ary expression then $(X \cap Y)$ is an n-ary expression where $n = max(l, m)$.
6. If X is a unary expression and Y is an $(n + 1)$-ary expression then $(\mathbf{some}XY)$ is an n-ary expression.
7. If X is a unary expression and Y is an $(n + 1)$-ary expression then $(\mathbf{k}XY)$ is an n-ary expression for each $\mathbf{k} \in \omega_+$.
8. If X is a unary expression and Y is an $(n + 1)$-ary expression then $(\mathbf{most}XY)$ is an n-ary expression.

In the sequel, superscripts and parentheses are dropped whenever no confusion can result. Metavariables are used as follows: S^n ranges over $\mathcal{S}_n$; R^n ranges over $\mathcal{R}_n$; P ranges over $\mathcal{P}$; X, Y, Z, W, V range over $\mathcal{L}_N$; X^n, Y^n, Z^n, W^n, V^n range over n-ary expressions of $\mathcal{L}_N$; and S ranges over singular expressions of $\mathcal{L}_N$. Applying subscripts to these symbols does not change their ranges.

An *interpretation* of $\mathcal{L}_N$ is a pair $\mathcal{I} = \langle \mathcal{D}, \mathcal{F} \rangle$ where $\mathcal{D}$ is a nonempty set, and $\mathcal{F}$ is a mapping defined on $\mathcal{P}$ satisfying:

1. for each $S_i^1 \in \mathcal{S}_1$, $\mathcal{F}(S_i^1) = \{\langle d \rangle\}$ for some (not necessarily unique) $d \in \mathcal{D}$,
2. for each $S_i^{n+1} \in \mathcal{S}_{n+1}$ $(n > 0)$, $\mathcal{F}(S_i^{n+1}) \subseteq \mathcal{D}^{n+1}$, such that $\forall \langle d_1, \ldots, d_n \rangle : \exists d : \langle d_1, \ldots, d_n, d \rangle \in \mathcal{F}(S_i^{n+1})$ and $\forall d' : \langle d_1, \ldots, d_n, d' \rangle \in \mathcal{F}(S_i^{n+1})$ implies $d = d'$, and
3. for each $R^n \in \mathcal{R}_n$, $\mathcal{F}(R^n) \subseteq \mathcal{D}^n$.

Let $\alpha = \langle d_1, d_2, \ldots \rangle \in \mathcal{D}^\omega$ (a sequence of individuals). Then $X \in \mathcal{L}_N$ *is satisfied by* α *in* $\mathcal{I}$ (written $\mathcal{I} \vDash_\alpha X$) iff one of the following holds:

1. $X \in \mathcal{P}$ with arity n and $\langle d_1, \ldots, d_n \rangle \in \mathcal{F}(X)$
2. $X = \langle k_1, \ldots, k_m \rangle P$ where $P \in \mathcal{P}$ with arity m and $\langle d_{k_1}, \ldots, d_{k_m} \rangle \vDash_\alpha P$
3. $X = \overline{Y}$ and $\mathcal{I} \nvDash_\alpha Y$
4. $X = Y \cap Z$ and $\mathcal{I} \vDash_\alpha Y$ and $\mathcal{I} \vDash_\alpha Z$
5. $X = \mathbf{some} Y^1 Z^{n+1}$ and for some $d \in \mathcal{D}$, $\langle d \rangle \vDash_\alpha Y^1$ and $\langle d \rangle \vDash_\alpha Z^{n+1}$
6. $X = \mathbf{k} Y^1 Z^{n+1}$ and $card(\{d \in \mathcal{D} : \langle d \rangle \vDash_\alpha Y^1 \text{ and } \langle d \rangle \vDash_\alpha Z^{n+1}\}) \geq k$
7. $X = \mathbf{most} Y^1 Z^{n+1}$ and $card(\{d \in \mathcal{D} : \langle d \rangle \vDash_\alpha Y^1 \text{ and } \langle d \rangle \vDash_\alpha Z^{n+1}\}) > card(\{d \in \mathcal{D} : \langle d \rangle \vDash_\alpha Y^1 \text{ and } \langle d \rangle \nvDash_\alpha Z^{n+1}\})$,

where $\mathcal{I} \nvDash_\alpha X$ is an abbreviation for not $(\mathcal{I} \vDash_\alpha X)$ and $\langle d_{i_1}, \ldots, d_{i_n} \rangle \vDash_\alpha X$ is an abbreviation for $\mathcal{I} \vDash_{\langle d_{i_1}, \ldots, d_{i_n}, d_1, d_2, \ldots \rangle} X$.

X is true in $\mathcal{I}$ (written $\mathcal{I} \vDash X$) iff $\mathcal{I} \vDash_\alpha X$ for every $\alpha \in \mathcal{D}^\omega$. *X is valid* (written $\vDash X$) iff X is true in every interpretation of $\mathcal{L}_N$. A 0-ary expression of $\mathcal{L}_N$ is called a *sentence*. A set Γ of sentences *is satisfied in* $\mathcal{I}$ iff each $X \in \Gamma$ is true in $\mathcal{I}$.

The following abbreviations are introduced to improve readability.

1. $\breve{R}^n := \langle n, \ldots, 1 \rangle R^n$
2. $X \cup Y := \overline{(\bar{X} \cap \bar{Y})}$
3. $X \subseteq Y := \overline{(X \cap \bar{Y})}$
4. $X \equiv Y := (X \subseteq Y) \cap (Y \subseteq X)$
5. $T :=$ **thing** $:= (S_0^1 \subseteq S_0^1)$
6. $\mathbf{some}X_n\mathbf{some}X_{n-1} \cdots \mathbf{some}X_1 Y := (\mathbf{some}X_n(\mathbf{some}X_{n-1} \cdots (\mathbf{some}X_1 Y) \cdots))$
7. $\mathbf{some}X^1 Y_n^2 \circ Y_{n-1}^2 \circ \cdots \circ Y_1^2 := (\mathbf{some} \cdots (\mathbf{some}(\mathbf{some}X^1 Y_n^2) Y_{n-1}^2) \cdots Y_1^2)$
8. $\mathbf{all}X^1 Y := \overline{\mathbf{some}X^1 \bar{Y}}$
9. $\mathbf{no}X^1 Y := \overline{\mathbf{some}X^1 Y}$
10. $!\mathbf{k}X^1 Y := \mathbf{k}X^1 Y \cap \overline{(\mathbf{k}+\mathbf{1})X^1 Y}$
11. $\bar{\mathbf{k}}X^1 Y := \overline{\mathbf{k}X^1 Y}$.

It is easy to see that:

1. $\mathfrak{I} \models_\alpha X \cup Y$ iff ($\mathfrak{I} \models_\alpha X$ or $\mathfrak{I} \models_\alpha Y$)
2. $\mathfrak{I} \models_\alpha X \subseteq Y$ iff ($\mathfrak{I} \models_\alpha X$ implies $\mathfrak{I} \models_\alpha Y$)
3. $\mathfrak{I} \models_\alpha X \equiv Y$ iff ($\mathfrak{I} \models_\alpha X$ iff $\mathfrak{I} \models_\alpha Y$)
4. $\mathfrak{I} \models_\alpha T$ for every $\mathfrak{I}$ and α
5. $\mathfrak{I} \models_\alpha \mathbf{some}X^1 Y_n^2 \circ \cdots \circ Y_1^2$ iff for some $a_1, \ldots, a_n \in \mathfrak{D} : \langle a_1, d_1 \rangle \models_\alpha Y_1^2$ and $\langle a_2, a_1 \rangle \models_\alpha Y_2^2$ and $\cdots$ and $\langle a_n, a_{n-1} \rangle \models_\alpha Y_n^2$ and $\langle a_n \rangle \models_\alpha X^1$ (thus $\circ$ denotes composition of relations in $\mathfrak{I}$)
6. $\mathfrak{I} \models_\alpha \mathbf{all}X^1 Y$ iff for all $d \in \mathfrak{D}$, $\langle d \rangle \models_\alpha X^1$ implies $\langle d \rangle \models_\alpha Y$
7. $\mathfrak{I} \models_\alpha \mathbf{no}X^1 Y$ iff for all $d \in \mathfrak{D}$, $\langle d \rangle \models_\alpha X^1$ implies $\langle d \rangle \not\models_\alpha Y$
8. $\mathfrak{I} \models_\alpha !\mathbf{k}X^1 Y$ iff $card(\{d \in \mathfrak{D} : \langle d \rangle \models_\alpha X^1$ and $\langle d \rangle \models_\alpha Y\}) = k$
9. $\mathfrak{I} \models_\alpha \bar{\mathbf{k}}X^1 Y$ iff $card(\{d \in \mathfrak{D} : \langle d \rangle \models_\alpha X^1$ and $\langle d \rangle \models_\alpha Y\}) < k$.

4 Axiomatization of $\mathfrak{L}_N$ The axiom schemata of $\mathfrak{L}_N$ are the following (see definitions of metavariables given in the previous section):

BT *Every schema that can be obtained from a tautologous Boolean wff by uniform substitution of nullary metavariables of $\mathfrak{L}_N$ for sentential variables, $\cap$ for $\wedge$, and $^-$ for $\neg$.*

C $\mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \langle k_1, \ldots, k_m \rangle P \equiv \mathbf{some}S_{i_{k_m}} \cdots \mathbf{some}S_{i_{k_1}} P$, *where P is of arity m and $n = max\ (k_j)_{1 \le j \le m}$.*

S $\mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \overline{(\mathbf{some}SX^{n+1})} \equiv \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \overline{\mathbf{some}SX^{n+1}}$.

D $\mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} (X^m \cap Y^l) \equiv (\mathbf{some}S_{i_m} \cdots \mathbf{some}S_{i_1} X^m \cap \mathbf{some}S_{i_l} \cdots \mathbf{some}S_{i_1} Y^l)$ *where $n = max(l, m)$.*

EG $(\mathbf{some}SX^1 \cap \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1}\mathbf{some}SY^{n+1}) \subseteq \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \mathbf{some}X^1 Y^{n+1}$.

KG1 $\mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1}\mathbf{some}X^1 Y^{n+1} \equiv \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \mathbf{1}X^1 Y^{n+1}$.

KG2 $(\mathbf{some}SX^1 \cap \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1}\mathbf{some}SY^{n+1} \cap \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \mathbf{k}(X^1 \cap \bar{S})Y^{n+1}) \subseteq \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} (\mathbf{k}+\mathbf{1})X^1 Y^{n-1}$ *for each* $\mathbf{k} \in \omega_+$.

MG1 $(\mathbf{some}SX^1 \cap \mathbf{all}S_{i_n} \cdots \mathbf{all}S_{i_1}\mathbf{all}X^1 Y^{n+1}) \subseteq \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \mathbf{most}X^1 Y^{n+1}$.

MG2 $(\mathbf{some}S_i X^1 \cap \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1}\mathbf{some}S_j Y^{n+1} \cap \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \mathbf{most}(X^1 \cap \bar{S_i} \cap \bar{S_j})Y^{n+1}) \subseteq \mathbf{some}S_{i_n} \cdots \mathbf{some}S_{i_1} \mathbf{most}X^1 Y^{n+1}$.

The inference rules of $\mathfrak{L}_N$ are the following:

MP From X^0 and $X^0 \subseteq Y^0$ infer Y^0.

EI From $\overline{(Z^0 \cap \mathbf{some}S^1X^1 \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{some}S^1Y^{n+1})}$, where S^1 is a unary singular predicate symbol which does not occur in X^1, Y^{n+1}, Z^0, or $S_{i_1},\ldots,S_{i_n}$, infer $\overline{(Z^0 \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{some}X^1Y^{n+1})}$.

KI From $\overline{(Z^0 \cap \mathbf{some}S^1X^1 \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{some}S^1Y^{n+1} \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{k}(X^1 \cap \overline{S^1})Y^{n+1})}$, where S^1 is a unary singular predicate symbol which does not occur in X^1, Y^{n+1}, Z^0, or $S_{i_1},\ldots,S_{i_n}$, infer $\overline{(Z^0 \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}(\mathbf{k}+\mathbf{1})X^1Y^{n+1})}$ for each $\mathbf{k} \in \omega_+$.

MI From $\overline{(Z^0 \cap \mathbf{some}S_i^1X^1 \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{some}S_i^1Y^{n+1} \cap (\mathbf{all}S_{i_n}\cdots\mathbf{all}S_{i_1}\mathbf{all}X^1Y^{n+1} \cup \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{most}(X^1 \cap \overline{S_i^1} \cap \overline{S_j^1})Y^{n+1}))}$, where S_i^1 and S_j^1 are unary singular predicate symbols which do not occur in X^1, Y^{n+1}, Z^0, or $S_{i_1},\ldots,S_{i_n}$, infer $\overline{(Z^0 \cap \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{most}X^1Y^{n+1})}$.

The set $\mathfrak{T}$ of theorems of $\mathfrak{L}_N$ is the smallest set containing the axioms and closed under MP, EI, KI, and MI.

Axiom S can also be written

S $\mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{all}SX^{n+1} \equiv \mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{some}SX^{n+1}$.

In view of this "wild quantity" of singular expressions, $\mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{all}SX^{n+1}$ and $\mathbf{some}S_{i_n}\cdots\mathbf{some}S_{i_1}\mathbf{some}SX^{n+1}$ will usually be written simply $S_{i_n}\cdots S_{i_1}SX^{n+1}$.

The following theorem establishes the soundness of this axiomatization:

Theorem 1 *$X \in \mathfrak{T}$ only if $\models X$.*

Proof: Let $\mathfrak{F}$ be extended to singular expressions by the inductive definition: $\mathfrak{F}(S_{i_n}\cdots S_{i_1}S^{n+1}) = \{\langle d\rangle\}$ where for $1 \le j \le n$, $\mathfrak{F}(S_{i_j}) = \{\langle d_{i_j}\rangle\}$, and $\langle d_{i_1},\ldots, d_{i_n}, d\rangle \in \mathfrak{F}(S^{n+1})$. Observe that by the definition of satisfaction, $\langle \mathfrak{F}(S_{i_1}),\ldots, \mathfrak{F}(S_{i_n})\rangle \models X^n$ iff $\langle \mathfrak{F}(S_{i_2}),\ldots,\mathfrak{F}(S_{i_n})\rangle \models S_{i_1}X^n$ iff $\cdots$ iff $\mathfrak{I} \models S_{i_n}\cdots S_{i_1}X^n$. From this observation and the definition of validity, it is not difficult to show that the axioms are valid and that validity is preserved by the inference rules. Details will be given only for MG1, MG2, and MI. In this proof, $\mathfrak{C}_1 := \{d : \langle d\rangle \models X\}$, $\mathfrak{C}_1' := \{d : \langle d\rangle \models X \cap \overline{S_j^1}\}$, $\mathfrak{C}_1'' := \{d : \langle d\rangle \models X \cap \overline{S_j^1} \cap \overline{S_i^1}\}$, and $\mathfrak{C}_2 := \{d : \langle d, \mathfrak{F}(S_{i_1}), \ldots, \mathfrak{F}(S_{i_n})\} \models Y\}$. Understanding the arguments of this theorem can be facilitated by translating them to van Benthem's *tree of numbers* (van Benthem [4] and Westerståhl [18]).

Claim (i) *$MG1$ is valid.*

Proof: Let $\mathfrak{I}$ be an arbitrary interpretation of $\mathfrak{L}_N$. $\mathfrak{I} \models SX \cap S_{i_n}\cdots S_{i_1}\mathbf{all}XY$ iff $\mathfrak{I} \models SX$ and $\mathfrak{I} \models S_{i_n}\cdots S_{i_1}\mathbf{all}XY$ iff $\langle \mathfrak{F}(S)\rangle \models X$ and $\forall d \in \mathfrak{D} : \langle d\rangle \models X$ implies $\langle d, \mathfrak{F}(S_{i_1}), \ldots, \mathfrak{F}(S_{i_n})\rangle \models Y$. Therefore $card(\mathfrak{C}_1 \cap \mathfrak{C}_2) \ge 1$ and $card(\mathfrak{C}_1 \cap \overline{\mathfrak{C}_2}) = 0$. Hence $\mathfrak{I} \models S_{i_n}\cdots S_{i_1}\mathbf{most}XY$.

Claim (ii) *$MG2$ is valid.*

Proof: Let $\mathcal{I}$ be an arbitrary interpretation of $\mathcal{L}_N$. As above, $\mathcal{I} \models S_i X \cap S_{i_n} \cdots S_{i_1} S_i Y \cap S_{i_n} \cdots S_{i_1} \mathbf{most}(X \cap \overline{S_i} \cap \overline{S_j})Y$ iff $\langle \mathcal{F}(S_i) \rangle \models X$, $\langle \mathcal{F}(S_i), \mathcal{F}(S_{i_1}), \ldots, \mathcal{F}(S_{i_n}) \rangle \models Y$, and $card(\mathcal{C}_1'' \cap \mathcal{C}_2) > card(\mathcal{C}_1'' \cap \overline{\mathcal{C}_2})$. Since $\mathcal{F}(S_i) \notin \mathcal{C}_1''$ but $\mathcal{F}(S_i) \in \mathcal{C}_1'$, $card(\mathcal{C}_1' \cap \mathcal{C}_2) = card(\mathcal{C}_1'' \cap \mathcal{C}_2) + 1 > card(\mathcal{C}_1'' \cap \overline{\mathcal{C}_2}) + 1 = card(\mathcal{C}_1' \cap \overline{\mathcal{C}_2}) + 1$. Therefore, for any value of $\mathcal{F}(S_j)$, $card(\mathcal{C}_1 \cap \mathcal{C}_2) \geq card(\mathcal{C}_1' \cap \mathcal{C}_2) > card(\mathcal{C}_1' \cap \overline{\mathcal{C}_2}) + 1 \geq card(\mathcal{C}_1 \cap \overline{\mathcal{C}_2})$. Hence, $\mathcal{I} \models S_{i_n} \cdots S_{i_1} \mathbf{most} XY$.

Claim (iii) *MI preserves validity.*

Proof: Suppose $\models \overline{(Z^0 \cap S_i^1 X \cap S_{i_n} \cdots S_{i_1} S_i^1 Y \cap (S_{i_n} \cdots S_{i_1} \mathbf{all} XY \cup S_{i_n} \cdots S_{i_1} \mathbf{most}(X \cap \overline{S_i^1} \cap \overline{S_j^1})Y))}$, where S_i^1 and S_j^1 do not occur in X, Y, Z^0, or $S_{i_1}, \ldots, S_{i_n}$, but there exist interpretations $\mathcal{I}$ such that $\mathcal{I} \models Z^0 \cap S_{i_n} \cdots S_{i_1} \mathbf{most} XY$. Thus $card(\mathcal{C}_1 \cap \mathcal{C}_2) > card(\mathcal{C}_1 \cap \overline{\mathcal{C}_2}) \geq 0$. Since S_i^1 is *fresh* (i.e., has no other occurrences), there are among the $\mathcal{I}$ interpretations $\mathcal{I}'$ such that $\mathcal{F}(S_i^1) \in \mathcal{C}_1 \cap \mathcal{C}_2$. Therefore, $\mathcal{I}' \models S_i^1 X$ and $\mathcal{I}' \models S_{i_n} \cdots S_{i_1} S_i^1 Y$. Now there are two cases to consider:

(a) $card(\mathcal{C}_1 \cap \overline{\mathcal{C}_2}) = 0$.
Then $\mathcal{I}' \not\models S_{i_n} \cdots S_{i_j} \mathbf{some} X\overline{Y}$, i.e., $\mathcal{I}' \models \overline{S_{i_n} \cdots S_{i_1} \mathbf{some} X\overline{Y}}$, which contradicts the assumption.

(b) $card(\mathcal{C}_1 \cap \overline{\mathcal{C}_2}) > 0$.
Then $card(\mathcal{C}_1 \cap \mathcal{C}_2) > 1$. Since S_j^1 is fresh, there are among the $\mathcal{I}'$ interpretations $\mathcal{I}''$ such that $\mathcal{F}(S_j^1) \in \mathcal{C}_1 \cap \overline{\mathcal{C}_2}$. Therefore, $card(\mathcal{C}_1'' \cap \mathcal{C}_2) = card(\mathcal{C}_1 \cap \mathcal{C}_2) - 1$ and $card(\mathcal{C}_1'' \cap \overline{\mathcal{C}_2}) = card(\mathcal{C}_1 \cap \overline{\mathcal{C}_2}) - 1$. Hence $\mathcal{I}'' \models S_{i_n} \cdots S_{i_1} \mathbf{most}(X \cap \overline{S_i^1} \cap \overline{S_j^1})Y$, which again contradicts the assumption.

The axiomatization is not complete however. Indeed the quantifier **most** cannot be axiomatized in a first-order language. This can be shown as follows. (See also [1].)

Suppose **most** is axiomatizable in $\mathcal{L}_N$. Let $X = \mathbf{most}TB$ and let Γ be a set of sentences such that for any interpretation $\mathcal{I}$ of $\mathcal{L}_N$, $\mathcal{I} \models X$ iff $\mathcal{I} \models \Gamma$. Let $n = \{0,1,2,\ldots,n-1\}$ and $\zeta_n = \{0,1,2,\ldots,\lfloor n/2 \rfloor + 1\}$. For each $n \in \omega_+$, define interpretation $\mathcal{I}_n = \langle n, \mathcal{F}_n \rangle$, where $\mathcal{F}_n(B) = \zeta_n$. Obviously, for each $n \in \omega_+$, $\mathcal{I}_n \models X$, and by hypothesis, $\mathcal{I}_n \models \Gamma$.

Now define $\mathcal{I} = \prod_{n \in \omega_+} \mathcal{I}_n / F$, where F is a nonprincipal ultrafilter (e.g., an extension of the Frechét filter to an ultrafilter). By Loš's Theorem (e.g., see Bell and Slomson [2]), $\mathcal{I} \models \Gamma$. Since F contains no singletons, $\mathbf{!k}TT$ cannot be satisfied in $\mathcal{I}$ for any k. Therefore $\mathcal{I}$ is infinite. Similarly, $\mathbf{!k}TB$ cannot be satisfied in $\mathcal{I}$ for any k, so $\mathcal{F}(B)$ is infinite. Hence both T and B denote infinite sets of the same cardinality in $\mathcal{I}$, viz., $2^{\aleph_0}$ ([2], Theorem 6.3.12). It follows that $\mathcal{I} \not\models \mathbf{most}TB$, a contradiction.

If the quantifier **most** were eliminated, the axiomatization of the remainder of $\mathcal{L}_N$ would be complete. The proof is a standard Henkin proof. Given a set of sentences Γ, a maximal consistent extension Γ^+ with witnesses is constructed. Witnesses are provided by $\mathbb{S}_1$. The singular expressions, modulo equivalence in Γ^+, comprise the domain of the interpretation.

Alternatively, if interpretations are restricted to some fixed finite upper bound (e.g., by adding the axiom $\overline{N}TT$), the axiomatization is complete. Of course, this is tantamount to accepting incompleteness. In any event, incompleteness does not negate the usefulness of the axiomatization for reasoning about natural language discourse.

*5 **Theorems*** The theorems presented in [9] can be generalized to apply to the extended language of this paper. Since the proofs closely follow those given in [9], the theorems will be stated without proof.

The main results are two monotonicity theorems. These theorems establish the monotonicity properties of quantifiers. They subsume the resolution principle. In addition, other properties of natural language quantifiers, including conservativity, are proved.

Before stating the first monotonicity theorem, some definitions are needed.

An occurrence of a subexpression Y in an expression W has *positive* (*negative*) *polarity* if that occurrence of Y lies in the scope of an even (odd) number of $^-$ operations in W, unless that occurrence of Y is a subexpression of V in $\mathbf{most}VZ$, in which case Y has *both* positive *and* negative polarity. This definition can be applied to expressions containing abbreviations by replacing the abbreviations with their definitions. The definition of polarity is summarized in Table 1. In this connection, singular expressions require special mention. Since $\mathbf{all}SX := \overline{\mathbf{some}S\overline{\overline{X}}} \equiv \overline{\overline{\mathbf{some}SX}} \equiv \mathbf{some}SX$, any governing occurrence of a singular expression can be taken to have *either* positive *or* negative polarity.

An occurrence of a subexpression Y^m, where $m \geq 1$, is *governed by X in W* if W is $\mathbf{some}XY^m$, $\mathbf{some}X\overline{Y^m}$, $\mathbf{some}X(Y^m \cap Z^l)$, $\mathbf{some}X(\overline{Y^m} \cap Z^l)$, $\mathbf{k}XY^m$, $\mathbf{k}X\overline{Y^m}$, $\mathbf{k}X(Y^m \cap Z^l)$, $\mathbf{k}X(\overline{Y^m} \cap Z^l)$, $\mathbf{most}XY^m$, $\mathbf{most}X\overline{Y^m}$, $\mathbf{most}X(Y^m \cap Z^l)$, $\mathbf{most}X(\overline{Y^m} \cap Z^l)$, or the complement of one of these expressions. An occurrence of Y^m is *governed by $X_n \cdots X_1$ in W*, where $1 \leq n \leq m$, if V is governed by X_n in W and that occurrence of Y^m is governed by $X_{n-1} \cdots X_1$ in V. An occurrence of Y^m in $\langle k_1, \ldots, k_m \rangle Y^m$ is *governed by $X_{k_m} \cdots X_{k_1}$ in W* if $\langle k_1, \ldots, k_m \rangle Y^m$ is governed by $X_n \cdots X_1$ in W, where $n = max(k_i)_{1 \leq i \leq m}$.

Table 1.

if V is:	and polarity of V is:	then polarity of X is:	and polarity of Y is:
$\mathbf{some}XY$	$+(-)$	$+(-)$	$+(-)$
$\mathbf{all}XY$	$+(-)$	$-(+)$	$+(-)$
$\mathbf{no}XY$	$+(-)$	$-(+)$	$-(+)$
$\mathbf{k}XY$	$+(-)$	$+(-)$	$+(-)$
$\mathbf{!k}XY$	$+(-)$	$+$*and*$-$ $(+$*and*$-)$	$+$*and*$-$ $(+$*and*$-)$
$\bar{\mathbf{k}}XY$	$+(-)$	$-(+)$	$-(+)$
$\mathbf{most}XY$	$+(-)$	$+$*and*$-$ $(+$*and*$-)$	$+(-)$
$X \cap Y$	$+(-)$	$+(-)$	$+(-)$
$X \subseteq Y$	$+(-)$	$-(+)$	$+(-)$
$X \cup Y$	$+(-)$	$+(-)$	$+(-)$
$X \equiv Y$	$+(-)$	$+$*and*$-$ $(+$*and*$-)$	$+$*and*$-$ $(+$*and*$-)$

Theorem 2 (First Monotonicity Theorem) *Let Y^m occur in W with positive* (*respectively, negative*) *polarity. Let* $(\mathbf{all}T)^m(Y^m \subseteq Z^l)$ (*respectively,* $(\mathbf{all}T)^m(Z^l \subseteq Y^m)$), *where $l \leq m$. Let W' be obtained from W by* (i) *substi-*

tuting Z^l for that occurrence of Y^m, (ii) *substituting $\langle k_1,.., k_l\rangle$ for selection operator $\langle k_1, \ldots, k_m\rangle$ on Y^m, if any, and* (iii) *eliminating all occurrences of governing subexpressions that no longer govern after the substitutions in* (i) *and* (ii). *Finally, let* **some**TX *for every governing subexpression X with an occurrence of negative polarity that was eliminated in* (iii). *Then* $(\mathbf{all}T)^h(W \subseteq W')$, *where h is the arity of W.*

Before the second monotonicity theorem can be presented, another definition is needed.

A subexpression Y^m will be said to *occur disjunctively* in expression W iff (i) $W = \mathbf{all}X_n \cdots \mathbf{all}X_1 Y^m \cup Z$, where $n \leq m$; or (ii) $W = \mathbf{all}X_n \cdots \mathbf{all}X_{k+1}(Z_1 \cup Z_2)$, where $0 \leq k \leq n$ and Y^m occurs disjunctively in Z_1.

Theorem 3 (Second Monotonicity Theorem) *Let Y^m occur disjunctively in W, governed by $X_k \cdots X_1$. Let* $(\mathbf{all}X_k \cdots \mathbf{all}X_1(Y^m \subseteq Z^l))$, *where $l \leq m$. Let W' be obtained from W by* (i) *substituting Z^l for that occurrence of Y^m*, (ii) *substituting $\langle k_1, \ldots, k_l\rangle$ for selection operator $\langle k_1, \ldots, k_m\rangle$ on Y^m, if any, and* (iii) *eliminating all occurrences of governing subexpressions that no longer govern after the substitutions in* (i) *and* (ii). *Finally, let* **some**TX *for every governing subexpression X that was eliminated in* (iii). *Then* $(\mathbf{all}T)^h(W \subseteq W')$, *where h is the arity of W.*

It is easy to see (from the equivalence $(Y^m \subseteq Z^l) \equiv (\overline{Y^m} \cup Z^l)$) that this theorem corresponds to the resolution principle in conventional logic. A corollary provides a rule corresponding to unit resolution.

Corollary 4 (Cancellation Rule) *Let Y^m occur disjunctively in W, governed by $X_k \cdots X_1$. Let* $\mathbf{all}X_k \cdots \mathbf{all}X_1 \overline{Y^m}$. *Let W' be obtained from W by deleting that occurrence of Y^m and all occurrences of governing subexpressions that no longer govern. Let* **some**TX *for every governing subexpression X that was deleted. Then* $(\mathbf{all}T)^h(W \subseteq W')$, *where h is the arity of W.*

A second corollary establishes a distributivity property.

Corollary 5 (Distributivity Rule) *Let Y^m occur disjunctively in W, governed by $X_k \cdots X_1$. Let* $(\mathbf{all}X_l \cdots \mathbf{all}X_1 Z^l))$, *where $l \leq m$. Let W' be obtained from W by replacing that occurrence of Y^m with $(Z^l \cap Y^m)$. Then* $(\mathbf{all}T)^h(W \subseteq W')$, *where h is the arity of W.*

The final theorems establish the property of conservativity and the rules for conversion in the case of unary expressions.

Theorem 6 (Conservativity) (schema)
(i) $(\mathbf{all}T)^{m-1}\mathbf{some}XY^m \equiv (\mathbf{all}T)^{m-1}\mathbf{some}X(Y^m \cap X)$
(ii) $(\mathbf{all}T)^{m-1}\mathbf{all}XY^m \equiv (\mathbf{all}T)^{m-1}\mathbf{all}X(Y^m \cap X)$
(iii) $(\mathbf{all}T)^{m-1}\mathbf{k}XY^m \equiv (\mathbf{all}T)^{m-1}\mathbf{k}X(Y^m \cap X)$
(iv) $(\mathbf{all}T)^{m-1}\mathbf{most}XY^m \equiv (\mathbf{all}T)^{m-1}\mathbf{most}X(Y^m \cap X)$.

Theorem 7 (Conversion) *For unary expressions X and Y*:
(i) $\mathbf{some}XY \equiv \mathbf{some}YX$
(ii) $\mathbf{all}XY \equiv \mathbf{all}(\overline{Y})\overline{X}$
(iii) $\mathbf{k}XY \equiv \mathbf{k}YX$.

6 A paradigm for surface reasoning Reasoning is viewed as theorem proving, using either direct or indirect proof methods. The objective of this section is to develop a paradigm for reasoning in $\mathfrak{L}_N$ that resembles syllogistic (monadic) reasoning, i.e., reasoning about inclusion, exclusion, and overlap of classes of individuals. To this end, a graphical domain is defined in which these relations can be naturally represented. But first a standard form for problem statements is defined.

Any sentences of $\mathfrak{L}_N$ can be *purified* [12], that is, put in a form in which all quantifiers have minimum scope. The procedure is well-known, using DeMorgan's laws (instances of Axiom BT) and the following lemmas, which follow directly from the First Monotonicity Theorem, the Distributivity Rule, and the Conservativity Theorem.

Lemma 8 (schema) $\mathbf{some}X_n \cdots \mathbf{some}X_1(Y^n \cap Z^0) \equiv (\mathbf{some}X_n \cdots \mathbf{some}X_1 Y^n \cap Z^0)$.

Lemma 9 (schema) $\mathbf{all}X_n \cdots \mathbf{all}X_1(Y^l \cap Z^m) \equiv (\mathbf{all}X_l \cdots \mathbf{all}X_1 Y^l \cap \mathbf{all}X_m \cdots \mathbf{all}X_1 Z^m)$, *where* $n = max(l, m)$.

Lemma 10 (schema) $\mathbf{some}XY^n \equiv \mathbf{some}T(X \cap Y^n)$.

After purification, the prime subexpressions all have the form $\mathbf{some}T(Z_1 \cap \cdots \cap Z_g)$ or $\mathbf{all}T(V_1 \cup \cdots \cup V_h)$. Putting the result of purification in disjunctive normal form yields a disjunction of expressions of the form $\mathbf{some}TX_1 \cap \cdots \cap \mathbf{some}TX_k \cap \mathbf{all}TY_1 \cap \cdots \cap \mathbf{all}TY_l$, where the X_i are conjunctions of prime subexpressions and the Y_j are disjunctions of prime subexpressions. A set $\Gamma = \{\mathbf{some}TX_1, \ldots, \mathbf{some}TX_k, \mathbf{all}TY_1, \ldots, \mathbf{all}TY_l\}$ of sentences comprising such a disjunct, or a set of sentences equivalent to these under Axiom S and Lemma 10, will be called a *standard form*. Sentences of the form SX are ambiguous with regard to their position in Γ. To remove this ambiguity, the convention will be adopted that SX is always interpreted as $\mathbf{all}SX$ or $\mathbf{all}T(S \subseteq X)$. Obviously, any problem (i.e., finite set of sentences) can be stated as a disjunction of standard forms. Indeed most problems involved in natural language reasoning can be stated as a single standard form.

The subset $\Gamma_+ = \{\mathbf{some}TX_1, \ldots, \mathbf{some}TX_k\}$ will be called the *positive part*, and the subset $\Gamma_- = \{\mathbf{all}TY_1, \ldots, \mathbf{all}TY_l\}$ the *negative part*, of Γ. Often the positive part will consist of a single element. The positive part represents a *lower bound*, *LB*, on the models of Γ in that at least the denotations of the X_i are asserted to be nonempty. Similarly the negative part represents an *upper bound*, *UB*, on the models of Γ in that at most the denotations of the Y_j are asserted to be nonempty. Therefore if Γ has a model, then each $X_i \in LB$ must be nonempty and contained in each $Y_j \in UB$.

Let $\Gamma \subseteq \mathfrak{L}_N$ be a consistent set of sentences. The relation $\sqsubseteq_\Gamma$, or simply $\sqsubseteq$ when no confusion can result, is defined: $X^l \sqsubseteq Y^m :\Leftrightarrow \Gamma \vdash (\mathbf{all}T)^n(X^l \subseteq Y^m)$, where $n = max(l, m)$. It is easy to see that $\sqsubseteq$ is a quasi-order on $\mathfrak{L}_N$. Moreover, if $\approx$ is defined $X \approx Y :\Leftrightarrow (X \sqsubseteq Y) \cap (Y \sqsubseteq X)$, then $\sqsubseteq$ is a partial order on $\mathfrak{L}_N/\!\approx$. The poset $\mathbf{L}_\Gamma = \langle \mathfrak{L}_N/\!\approx, \sqsubseteq \rangle$ is the *Lindenbaum algebra of* Γ. It can be shown (e.g., Bell and Slomson [3]) that $\mathbf{L}_\Gamma$ is a Boolean lattice with greatest and least elements $|\mathbf{some}TT|$ and $|\overline{\mathbf{some}TT}|$, respectively. Further, if $|X|$ and $|Y|$

are equivalence classes of $\mathcal{L}_N/\approx$, then the meet and join of $|X|$ and $|Y|$ are $|X \cap Y|$ and $|X \cup Y|$, respectively, and the complement of $|X|$ is $|\overline{X}|$.

The following properties of $\mathbf{L}_\Gamma$ are easy to prove:

1. $|\mathbf{some}TT| = |\mathbf{all}TT| = |T| = |\langle n\rangle T|$ where $\langle n\rangle T := \langle n\rangle S_0 \subseteq \langle n\rangle S_0$
2. $|\overline{\mathbf{some}TT}| = |\mathbf{some}T\overline{T}| = |\overline{T}| = |\overline{\langle n\rangle T}|$
3. $|(\mathbf{some}T)^m X^n| \sqsubseteq |\overline{\mathbf{some}TT}|$ iff $|(\mathbf{some}T)^{m+1} X^n| \sqsubseteq |\overline{\mathbf{some}TT}|$ for $0 \leq m < n$
4. $|X^n| \not\sqsubseteq |\overline{\mathbf{some}TT}|$ iff $\Gamma \vdash (\mathbf{some}T)^n X^n$
5. $|\mathbf{some}TT| \sqsubseteq |X^n|$ iff $\Gamma \vdash (\mathbf{all}T)^n X^n$.

Let $\mathcal{L}_n \subseteq \mathcal{L}_N$ be the set of n-ary expressions. Then $\mathbf{L}_{\Gamma,n} = \langle \mathcal{L}_n/\approx, \sqsubseteq\rangle$ is a sublattice of $\mathbf{L}_\Gamma$ for each $n \in \omega$. From Properties 1 and 2, $\mathbf{L}_{\Gamma,n}$ has the same greatest and least elements as $\mathbf{L}_\Gamma$.

Define *rank* $r: \mathcal{L}_N \to \omega$ as follows (cf. Rantala [13]).

1. $r(P) = 0$ for $P \in \mathcal{P}$
2. $r(\langle k_1, \ldots, k_m\rangle P) = 0$ for $P \in \mathcal{P}$
3. $r(\overline{X}) = r(X)$
4. $r(X \cap Y) = max(r(X), r(Y))$
5. $r(\mathbf{some}XY) = max(r(X), r(Y)) + 1$.

If Γ is a set of expressions, then $r(\Gamma) := sup\{r(X) : X \in \Gamma\}$. Now let $\mathcal{L}^{(d)} \subseteq \mathcal{L}_N$ be the set of expressions of rank $\leq d$. It can easily be seen that $\mathbf{L}_\Gamma^{(d)} = \langle \mathcal{L}^{(d)}/\approx, \sqsubseteq\rangle$ is a sublattice of $\mathbf{L}_\Gamma$ for each $d \in \omega$. In general, $\mathbf{L}_{\Gamma,n}^{(d)} = \langle \mathcal{L}_n^{(d)}/\approx, \sqsubseteq\rangle$ is a sublattice of $\mathbf{L}_\Gamma$ for each $n \in \omega$ and $d \in \omega$.

Reasoning can be considered a search of $\mathbf{L}_\Gamma$. The discussion to follow will emphasize refutation, but the same principles hold for direct proof. If a standard form Γ is inconsistent, then $\mathbf{L}_\Gamma$ has only one element. Conversely, inconsistency of Γ can be established by proving that in $\mathbf{L}_\Gamma$, $|\mathbf{some}TT| = |\overline{\mathbf{some}TT}|$. This would follow for example if $\mathbf{some}TX \in \Gamma_+$ and for some $Y: X \sqsubseteq (Y \cap \overline{Y})$. The search for such a Y is the essence of reasoning by refutation. In general, it is not decidable whether such a Y exists (since predicate logic is undecidable). Whether such a Y exists in the restricted lattice $\mathbf{L}_\Gamma^{(d)}$ is decidable. But even in this restricted domain the problem is NP-hard (since SAT can be reduced to it). Therefore, some constraints must be imposed on the search. In the following sections, two types of constraint will be discussed: (i) constraints that require, or at least give preference to, certain theorems and inference rules to be used in the search; and (ii) constraints that give preference to certain search paths.

By Property 3 it is sufficient to restrict the search to $\mathbf{L}_{\Gamma,1}$, since $\mathbf{L}_{\Gamma,0}$ is contradictory iff $\mathbf{L}_{\Gamma,1}$ is also. The relations between elements of $\mathbf{L}_{\Gamma,1}$ are inclusion, exclusion, and overlap, and thus search of $\mathbf{L}_{\Gamma,1}$ closely resembles syllogistic reasoning. From Properties 4 and 5, it follows that a standard form directly yields elements of $\mathbf{L}_{\Gamma,1}$.

Let $r(\Gamma) = d$. $\mathbf{L}_{\Gamma,1}^{(d)}$ is finite and therefore atomistic. The atoms of $\mathbf{L}_{\Gamma,1}^{(d)}$ correspond to the *attributive constituents at depth d* of Hintikka's distributive normal forms ([6] and [13]). Thus, the atoms denote all the classes of individuals that can exist in the world entailed by Γ.

For these reasons, construction of a contradictory subposet of $\mathbf{L}_{\Gamma,1}^{(d)}$ is proposed as a model of indirect surface reasoning. Similarly, construction of a subposet of $\mathbf{L}_{\Gamma,1}^{(d)}$ which exhibits the conclusion $X \not\leq \overline{\mathbf{some}TT}$ is proposed as a model of direct surface reasoning.

7 Global strategies This section presents strategies for simplifying proofs by imposing global restrictions and preferences on the reasoning process. The strategies are illustrated by examples. Criteria for strategy selection are proposed.

Let Γ be a standard form which is to be shown inconsistent. Γ might represent the whole or part of a logic problem, or it might represent a natural language discourse with the denial of some conclusion from that discourse. The illustrations of reasoning will be presented graphically as subposets of $\mathbf{L}_{\Gamma,1}^{(d)}$. Expressions of $\mathcal{L}_N$ will represent their equivalence classes.

7.1 Breadth-first strategy Meaning inclusion or entailment as it relates to natural language understanding is often taken to be identical with logical entailment, leading to the paradox of logical omniscience. Hintikka [6] suggests a way to avoid this:

> Whatever the meaning of a sentence is or may be, it seems to me that the (literal) meaning of a (grammatically correct) sentence has to be something that anyone who knows the language in question can effectively find out. . . . [Therefore] trivial implication seems to me a much better explication of the idea of meaning inclusion than logical implication.

Suppose Γ is as described above and $r(\Gamma) = d$. Trivial implication of the conclusion by the premises is indicated by the trivial inconsistency of Γ. Γ is trivially inconsistent if a search restricted to $\mathbf{L}_{\Gamma,1}^{(d)}$ can produce a contradiction.

Generalizing this explication of meaning inclusion yields the following breadth-first strategy. Initially the search is restricted to $\mathbf{L}_{\Gamma,1}^{(d)}$. If this fails to produce a contradiction, the search is extended to $\mathbf{L}_{\Gamma,1}^{(d+1)}$. If this fails as well, the search is extended to $\mathbf{L}_{\Gamma,1}^{(d+2)}$, and so on, until a contradiction is found or some limit on resource use is reached. Of course, as the reasoning process moves to $\mathbf{L}_{\Gamma,1}^{(d+i)}$ for increasing i, it passes from surface reasoning to depth reasoning.

This strategy can be used in conjunction with any other strategy. If a limit is not imposed, it is a complete strategy.

7.2 Cancellation strategy The Cancellation Rule (CANC), used in conjunction with the First Monotonicity Theorem (MON), is very effective for a certain class of problems. A well-known nontrivial example is *Schubert's Steamroller* (see [9] and Stickel [17] for details). However, CANC and MON alone are not complete. Therefore, the *cancellation strategy* limits itself to giving *preference* to the use of CANC and MON (cf. the *unit preference strategy* (Wos [19])).

A simple illustration of the cancellation strategy is provided by the following example.

> If Ben owns a donkey, then he feeds it. Every donkey that Harriet rides is owned by Ben. Susie is a donkey and Harriet rides Susie. Therefore, Ben feeds Susie.

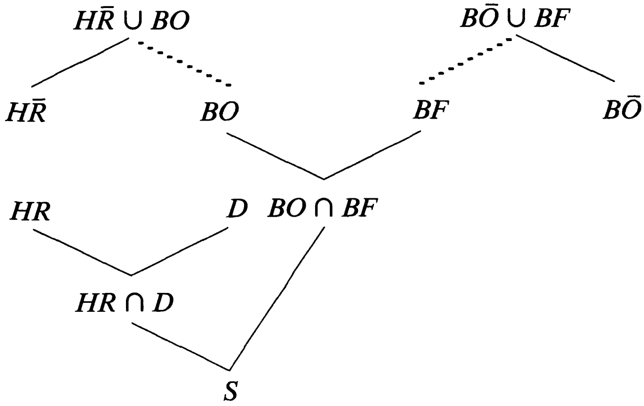

Figure 1. Example of cancellation strategy.

In standard form, $\Gamma = \{\mathbf{all}D(B\bar{O} \cup BF), \mathbf{all}D(H\bar{R} \cup BO), SD, SHR\}$. The relevant subposet of $\mathbf{L}^{(d)}_{\Gamma,1}$ is shown in Figure 1. Inferences based on cancellation appear as dotted arcs. The conclusion follows from $S \sqsubseteq BF$. Notice that $S \sqsubseteq D$ immediately implies (by MON) that $S \sqsubseteq (B\bar{O} \cup BF)$ and $S \sqsubseteq (H\bar{R} \cup BO)$. These inferences correspond to unification in conventional logic. Subsequent cancellations correspond to unit resolution. This example also illustrates direct reasoning.

7.3 Instantiation strategy If $\langle k_1, \ldots, k_m \rangle$ and $\langle l_1, \ldots, l_m \rangle$ are distinct selection operators, then $\langle k_1, \ldots, k_m \rangle R^m$ and $\langle l_1, \ldots, l_m \rangle R^m$ will be called *variants* of each other. A set Γ of sentences in which no predicate symbol occurs with two or more distinct selection operators will be said to be *without variants.*

If the problem statement is not without variants, use of EIC and Axiom C may be necessary to establish the connection between sentences involving distinct variants of the same predicate. EIC is the contrapositive form of Rule EI and is sound for refutation proofs. If such predicates are already governed by singular predicates, then only Axiom C need be used. Early use of EIC and Axiom C under these conditions will be called the *instantiation strategy.*

The next example, taken from Quine [12], illustrates the instantiation strategy.

> All natives of Ajo have a cephalic index in excess of 96. All women who have a cephalic index in excess of 96 have Pima blood. Therefore, anyone whose mother is a native of Ajo has Pima blood. (The following tacit assumptions are also made. Every mother is a woman. Everyone whose mother has Pima blood also has Pima blood.)

The premises and denial of the conclusion are given by the standard form: $\Gamma = \{\mathbf{all}AC, \mathbf{all}(W \cap C)P, \mathbf{some}(\mathbf{some}AM)\bar{P}, \mathbf{all}(\mathbf{something}\breve{M})W, \mathbf{all}(\mathbf{some}PM)P\}$. Use of EIC is necessary to relate the sentences involving M and $\breve{M}$. The construction is shown in Figure 2. Heavy arcs represent premises; dotted arcs show the

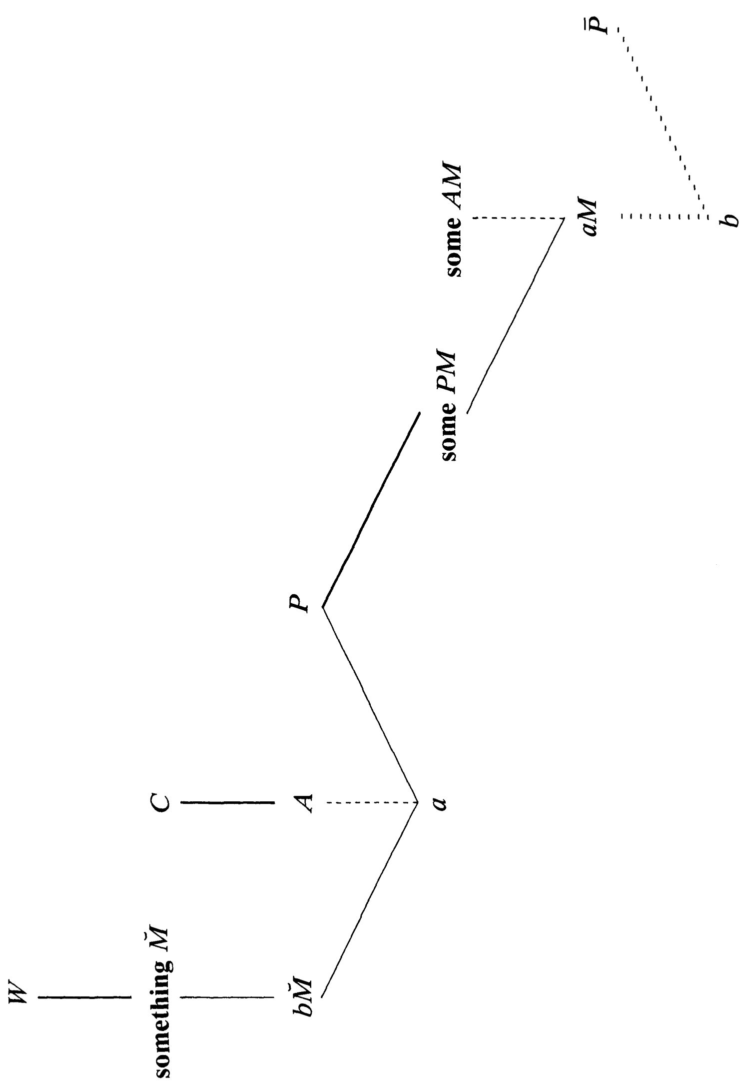

Figure 2. Example of existential instantiation.

two uses of EIC. Except for one use of Axiom C, the lighter arcs represent inferences involving MON. Contradiction is evidenced by $b \sqsubseteq (P \cap \bar{P})$.

Some problem statements involve variants, but the variants can be eliminated. In other problems, the variants are noneliminable, as illustrated by the following example.

Any transitive symmetric binary relation is reflexive.

In standard form for refutation: $\Gamma = \{(\mathbf{allthing})^2(R \circ R \subseteq R), (\mathbf{allthing})^2(R \equiv \breve{R}), \mathbf{something}\overline{\langle 1,1 \rangle R}\}$.

In Quine's example, the variants can be eliminated to yield the equivalent standard form: $\Gamma' = \{\mathbf{all}AC, \mathbf{all}(W \cap C)P, \mathbf{some}(\mathbf{some}AM)\bar{P}, \mathbf{allthing}\ \mathbf{all}\overline{W}\overline{M}, \mathbf{all}(\mathbf{some}PM)P\}$. The construction based on Γ' is given in Figure 3. Heavy and lighter arcs have the same significance. Notice that this construction is no smaller

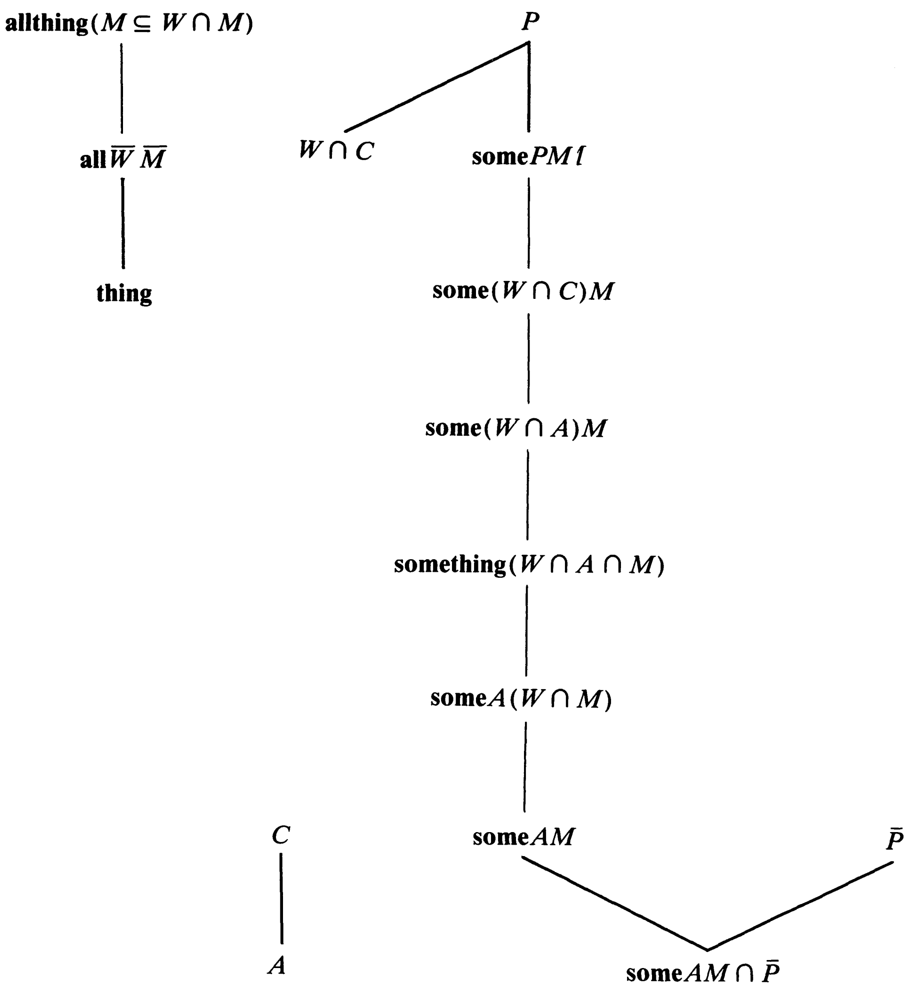

Figure 3. Previous example with variants eliminated.

and the variety of inferences no less than the previous one. Indeed by some criteria it may be judged more complex.

In those cases where the problem statement is without variants, it would seem that Axiom C could have no essential role in a proof. Moreover, if conversion is unnecessary, it would seem also that EIC is not required. However, this is not the case. A counterexample is given below. It is based on the observation that in any inference, polarity of corresponding subexpressions is unchanged except when Axiom S is used. This can be seen by examining each axiom and inference rule. For each axiom, corresponding subexpressions in the two expressions under the main connective have the same polarity, with the sole exception of Axiom S. For each inference rule, corresponding subexpressions in the antecedent and consequent have the same polarity.

Now consider the example: if $\mathbf{some}B\mathbf{all}AR$ and $\mathbf{some}AC$ then $\mathbf{some}B\mathbf{some}CR$. The premises and the conclusion are without variants. However, the subexpression A in the premise $\mathbf{some}B\mathbf{all}AR$ corresponds to the subexpression C in the conclusion $\mathbf{some}B\mathbf{some}CR$, and they have different polarities. By the above observation, this change in polarity can be accomplished only by application of Axiom S. This, in turn, can be accomplished only by use of Rule EI.

7.4 Reasoning with MON only Even when variants are eliminated, the construction for the previous example remains complex. The complexity is due to the sentence $\mathbf{allthing}\ \mathbf{all}\overline{W}\overline{M}$, i.e., "Of all things all non-women are non-mothers." An equivalent form, $(\mathbf{allthing})^2(M \subseteq W)$, i.e., "All who stand in the mother relation are women," is less awkward. In the latter form one recognizes a property that is not typical in natural language, viz., an inclusion relation between expressions of differing arities. Inclusions that involve only prime subexpressions of the same arity will be called *homogeneous*. A set of sentences containing only homogeneous inclusions will also be called homogeneous. Where only homogeneous inclusions are involved, MON assumes the following simpler form.

MONH Let Y^m occur in W with positive (respectively, negative) polarity. Let $(\mathbf{all}T)^m(Y^m \subseteq Z^m)$ (respectively, $(\mathbf{all}T)^m(Z^m \subseteq Y^m)$). Let W' be obtained from W by substituting Z^m for that occurrence of Y^m. Then $(\mathbf{all}T)^h(W \subseteq W')$, where h is the arity of W.

In many cases, a problem statement can be rephrased to be homogeneous and without variants. The following is Sommers' [14] version of Quine's problem, which does just this.

> All natives of Ajo have a cephalic index in excess of 96. All women who have a cephalic index greater than 96 have Pima blood. Therefore, anyone Ajoan on both sides has Pima blood. (Tacit assumptions are as follows. All descended from someone with Pima blood have Pima blood. Anyone who is Ajoan on both sides is a descendent of some woman Ajoan. All cases of [the first statement] are cases of every woman Ajoan being a woman with a cephalic index greater than 96.)

In standard form this problem can be given: $\Gamma = \{\mathbf{all}AC, \mathbf{all}(W \cap C)P, \mathbf{some}(\mathbf{some}AB)\bar{P}, \mathbf{all}(\mathbf{some}P\breve{D})P, \mathbf{all}(\mathbf{some}AB)\mathbf{some}(W \cap A)\breve{D}\}$. The last tacit

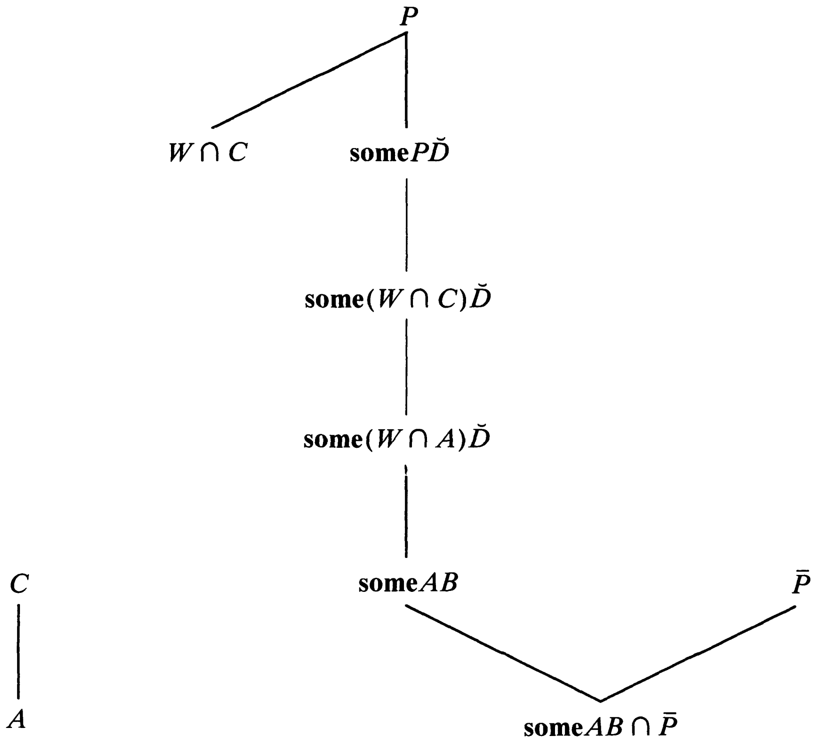

Figure 4. Previous example reformulated.

assumption, **all**$AC \subseteq$ **all**$(W \cap A)(W \cap C)$, is redundant since it is a valid sentence. The construction of a contradictory subposet is shown in Figure 4. It involves MONH only.

These considerations motivate the following simple strategy, called the *monotone strategy.* Whenever the problem statement Γ is homogeneous and without variants, give preference to inferences involving MONH only. This strategy is very effective for a restricted class of problems, notably problems of the kind appropriate for Sommers' Term Calculus. Like the cancellation strategy, it is not complete and so is limited to giving preference to reasoning with MONH only.

There is a connection here with Quine's *fluted* expressions (Noah [8] and Quine [11]). The language of fluted expressions forms a proper sublanguage of the language of $\mathcal{L}_N$ expressions without variants. For example, the problem just considered cannot be stated in the language of fluted expressions. The predicate calculus restricted to fluted expressions is decidable. It is not known whether $\mathcal{L}_N$ without variants is decidable [8].

7.5 Strategy selection Four global strategies have been defined: breadth-first, cancellation, instantiation, and monotone. They constitute only a first step toward a set of global strategies. What is wanted is a classification of standard forms by their syntactic properties that correlates with an optimal strategy.

Whether such a classification exists is an open question. The strategies of this section can be summarized as follows.

1. The breadth-first strategy is indicated for all problems. It seems likely that for natural language understanding d rarely exceeds 3. Therefore a limit of 3 or 4 on d would appear reasonable.
2. If the standard form contains sentences in which some subexpression has disjunctive occurrences of opposite polarities, the cancellation (preference) strategy is indicated.
3. If the standard form is not without variants, use of EIC and Axiom C is indicated to relate sentences containing different variants of the same predicate.
4. If the standard form is both homogeneous and without variants, the monotone (preference) strategy is indicated.

A standard form may decompose into subsets, each belonging to a distinct class. In this case, the subproblems are treated independently.

8 A local strategy In this section, the use of pattern matching to guide construction of a contradictory subposet is considered. A general description of this local strategy is as follows. The subposet construction proceeds under the appropriate global strategies. It begins with a nonnull node (known to be nonnull from the premises). This node may be an intersection of the form $X \cap Y$. In this case, the goal is to find a Z such that $X \sqsubseteq Z$ and $Y \sqsubseteq \bar{Z}$. This is symbolized:

$$\textbf{Goal: unify } (X)\uparrow, (\bar{Y})\downarrow.$$

Alternatively, the nonnull node may not be an intersection, having the form X. In this case, the goal is to find $X \sqsubseteq \bar{X}$, symbolized:

$$\textbf{Goal: unify } (X)\uparrow, (\bar{X})\downarrow.$$

Goals are pursued by recursion on the structure of the expressions involved. Subexpressions are considered in right to left order. For example, from the goal:

$$\textbf{Goal: unify } (\mathbf{some}XZ)\uparrow, (\mathbf{some}YW)\downarrow$$

the first subgoal would be:

$$\textbf{Goal: unify } (Z)\uparrow, (W)\downarrow.$$

The strategy will be illustrated using the following example.

> All who enter the country unaccompanied by a VIP are searched by a customs official. All who enter the country carrying contraband and who are searched by an honest customs official are reported. All smugglers carry contraband. All who carry contraband are dishonest. Yesterday, some smugglers entered the country unaccompanied by anyone else, but were not reported. Therefore, some VIPs or customs officials are dishonest.

A standard form of the problem for proof by refutation is: $\Gamma = \{\mathbf{all}(E \cap \overline{\mathbf{some}\overline{V}A})\mathbf{some}CS$, $\mathbf{all}(E \cap \overline{\mathbf{some}D\breve{F}} \cap \mathbf{some}(C \cap H)S)R$, $\mathbf{all}B\mathbf{some}D\breve{F}$, $\mathbf{all}(\mathbf{some}D\breve{F})\bar{H}$, $\mathbf{some}(E \cap B \cap \overline{\mathbf{some}\bar{B}A})\bar{R}$, $\mathbf{all}(V \cup C)H\}$. Initially, the given subposet of $\mathbf{L}^{(2)}_{\Gamma,1}$ contains only one node known to be nonnull, viz., $(E \cap B \cap \overline{\mathbf{some}\bar{B}A}) \cap \bar{R}$. This node is the intersection of two chains. These chains will be extended upward, seeking a node on one chain labeled Z, such that a node on the other chain is labeled $\bar{Z}$. This objective is pursued by recursion on the structure of the expressions found on the two chains.

1. **Goal: unify** $(E \cap B \cap \overline{\mathbf{some}\bar{B}A})\uparrow, (R)\downarrow$
   observation: $(E \cap \mathbf{some}D\breve{F} \cap \mathbf{some}(C \cap H)S) \sqsubseteq R$
2. **Goal: unify** $(E \cap B \cap \overline{\mathbf{some}\bar{B}A})\uparrow, (E \cap \mathbf{some}D\breve{F} \cap \mathbf{some}(C \cap H)S)\downarrow$
   2.1. **Goal: unify** $(\overline{\mathbf{some}\bar{B}A})\uparrow, (\mathbf{some}(C \cap H)S)\downarrow$
      2.1.1. **Goal: unify** $(A)\downarrow, (S)\downarrow$
         observation: no inferences possible
      2.1.2. **Goal: unify** $(\overline{\mathbf{some}\bar{B}A})\uparrow, (\mathbf{some}(C \cap H)S)\downarrow$
         observation: $C \sqsubseteq (C \cap H)$; hence, $\mathbf{some}CS \sqsubseteq \mathbf{some}(C \cap H)S$
      2.1.3. **Goal: unify** $(\overline{\mathbf{some}\bar{B}A})\uparrow, (\mathbf{some}CS)\downarrow$
         observation: $(E \cap \overline{\mathbf{some}VA}) \sqsubseteq \mathbf{some}CS$
      2.1.4. **Goal: unify** $(\overline{\mathbf{some}\bar{B}A})\uparrow, (\overline{\mathbf{some}VA})\downarrow$
         observation: (E is absorbed by Goal 2)
         2.1.4.1. **Goal: unify** $(B)\uparrow, (V)\uparrow$
            observation: $V \sqsubseteq H$
         2.1.4.2. **Goal: unify** $(B)\uparrow, (H)\uparrow$
            observation: $\mathbf{some}D\breve{F} \sqsubseteq \bar{H}$; hence, $H \sqsubseteq \overline{\mathbf{some}D\breve{F}}$
         2.1.4.3. **Goal: unify** $(B)\uparrow, (\mathbf{some}D\breve{F})\downarrow$
            observation: $B \sqsubseteq \mathbf{some}D\breve{F}$
   2.2. **Goal: unify** $(B)\uparrow, (\mathbf{some}D\breve{F})\downarrow$
      observation: $B \sqsubseteq \mathbf{some}D\breve{F}$.

The initial goal has been achieved, and so a contradiction has been obtained.

This strategy could employ heuristics to order alternative derived goals. For example, for the goal:

Goal: unify $(X)\uparrow, (Y)\downarrow$

possible heuristics include:

(i) consider the expression with the simpler construction first
(ii) consider the expression with the fewer possible inferences first
(iii) consider inferences $X \sqsubseteq X'$ ($Y' \sqsubseteq Y$) first that make X' and Y (X and Y') most similar in terms of polarity and rank of their subexpressions.

9 Conclusion In the theory of reasoning presented in this paper, a problem statement in standard form defines a lattice of expressions, each denoting a class of individuals. The reasoning process is represented as construction of a fragment of this lattice. Restriction of the reasoning process to unary expressions and construction of a partially ordered subset are salient features of the theory. A number of advantages follow.

First and most important, the reasoning process is similar to syllogistic, dealing with classes and their relation by inclusion, exclusion, and overlap. The monotonicity of natural language quantifiers, which is the basis of syllogistic, is the unifying principle of surface reasoning, embodied in inference rule MON. The simplicity and directness of surface reasoning is a result. Where the problem statement is homogeneous and without variants rule MON alone usually suffices. Reasoning in such cases is virtually identical to syllogistic reasoning.

Second, the local strategy, which guides the search for a contradiction by syntactic pattern matching, is based on an explicit order. Patterns exhibited by expressions of the subposet can be interpreted only in the context of the partial order; while several pairs of expressions may have the syntactic potential to produce a complementary pair, only those that lie on intersecting chains can produce a contradiction.

Third, the partial order provides a subsumption relation on the classes of individuals. (These classes are called *sorts* in conventional logic.) The subsumption relation allows MON to unify expressions without processing variables and in particular without an "occur check". In the cancellation strategy, which corresponds to the unit resolution strategy of conventional logic, unification is provided by MON with resolution performed by CANC.

A reasoning procedure is a calculus together with an algorithm to control deduction in the calculus. $\mathcal{L}_N$ is proposed as an appropriate calculus. The paradigm presented in this paper, together with the strategies for efficient construction of model fragments, constitutes an operational definition of a control algorithm.

This paper only begins the study of strategies in the context of surface reasoning. Much more can be accomplished. And beyond uniform strategies, such as those considered here, methods for incorporating problem domain-specific information into the reasoning process will be crucially important in emulating human reasoning and understanding of natural language.

Acknowledgment The author is indebted to the anonymous referee for a number of suggestions that resulted in significant improvement of this revised version over the original paper.

REFERENCES

[1] Barwise, J. and R. Cooper, "Generalized quantifiers and natural language," *Linguistics and Philosophy*, vol. 4 (1981), pp. 159–219.

[2] Bell, J. L. and A. B. Slomson, *Models and Ultraproducts: An Introduction*, North-Holland, Amsterdam, 1969.

[3] Bell, J. L. and M. Machover, *A Course in Mathematical Logic*, North-Holland, Amsterdam, 1977.

[4] van Benthem, J., "Questions about quantifiers," *The Journal of Symbolic Logic*, vol. 49 (1984), pp. 443–466.

[5] Čulík, K., "Well-translatable grammars and algol-like languages," pp. 76–85 in *Formal Language Description Languages for Computer Programming*, edited by T. B. Steel, Jr., North-Holland, Amsterdam, 1966.

[6] Hintikka, J., "Surface information and depth information," pp. 263–297 in *Information and Inference*, edited by J. Hintikka and P. Suppes, Reidel, Dordrecht, 1970.

[7] Mendelson, E., *Introduction to Mathematical Logic*, Third Edition, Wadsworth and Brooks/Cole, Monterey, 1987.

[8] Noah, A., "Predicate-functors and the limits of decidability in logic," *Notre Dame Journal of Formal Logic*, vol. 21 (1980), pp. 701–707.

[9] Purdy, W. C., "A logic for natural language," *Notre Dame Journal of Formal Logic*, vol. 32 (1991), pp. 409–425.

[10] Quine, W. V., *Mathematical Logic*, Revised Edition, Harvard University Press, Cambridge, 1951.

[11] Quine, W. V., *The Ways of Paradox and Other Essays*, Enlarged Edition, Harvard University Press, Cambridge, 1976.

[12] Quine, W. V., *Methods of Logic*, Fourth Edition, Harvard University Press, Cambridge, 1982.

[13] Rantala, V., "Constituents," pp. 43–76 in *Jaakko Hintikka*, edited by R. J. Bogdan, Reidel, Dordrecht, 1987.

[14] Sommers, F., *The Logic of Natural Language*, Clarendon Press, New York, 1982.

[15] Sommers, F., "The calculus of terms," pp. 11–56 in *The New Syllogistic*, edited by G. Englebretsen, Peter Lang, New York, 1987.

[16] Steedman, M., "Combinators and grammars," pp. 417–442 in *Categorial Grammars and Natural Language Structures*, edited by R. T. Oehrle et al., Reidel, Dordrecht, 1988.

[17] Stickel, M. E., "Schubert's steamroller problem: Formulations and solutions," *Journal of Automated Reasoning*, vol. 2 (1986), pp. 89–101.

[18] Westerståhl, D., "Some results on quantifiers," *Notre Dame Journal of Formal Logic*, vol. 25 (1984), pp. 152–170.

[19] Wos, L., D. Carson and G. Robinson, "The unit preference strategy in theorem proving," *Proceedings—Full Joint Computer Conference*, vol. 26 (1964), pp. 616–621.

School of Computer and Information Science
Syracuse University
Syracuse, New York 13244