9-1990

Surface Reasoning

William C. Purdy
Syracuse University, wcpurdy@ecs.syr.edu

SU-CIS-90-15

Surface Reasoning

William C. Purdy

September 1990

School of Computer and Information Science
Suite 4-116
Center for Science and Technology
Syracuse, New York 13244-4100

(315) 443-2368

Surface Reasoning

William C. Purdy*

September 1990

*School of Computer and Information Science, Syracuse University

Abstract

Surface reasoning is defined to be deduction conducted in the surface language in terms of certain primitive logical relations. The 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 calculus for surface reasoning is presented. Then a model for reasoning in this calculus is developed. The model is similar to but more general than syllogistic. In this model, 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 described. Criteria for strategy selection are proposed.

1 Introduction

To make the notion of surface reasoning sufficiently precise, two subordinate notions are needed. 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. The *primitive logical relations* are taken to be inclusion, exclusion and overlap between classes of entities. *Surface reasoning* is then defined to be deduction conducted in the surface language, in terms of the primitive logical relations.

The position taken in this paper is that the primitive logical relations are the primitive constructs of human reasoning, and moreover, that the surface language is adequate to express and manipulate these constructs. Disparate logics and complex transformations linking them to the surface language are not necessary to explain language understanding and reasoning. As the reasoning process becomes more complex, it passes over into depth reasoning or logical analysis. The property that characterizes surface reasoning, as opposed to depth reasoning, is its immediacy. Of course, given the state of cognitive science, one can only press this position by persuasion. Some of the persuasive arguments that can be marshaled rest on the following observations. Inclusion, exclusion and overlap are primary, perhaps innate, spatial concepts. They are exhibited in children's reasoning at an early age [4]. These relations are the basis of the syllogistic (expressed by the categorical statements A, E and I, respectively). They are also the basis of all intuitive systems such as Venn diagrams.

By definition a calculus for surface reasoning must be a direct representation of written English. However it is desirable that the calculus employ a notation that is briefer

than written English. Syllogistic is such a calculus. But syllogistic is limited in scope to monadic logic. A previous paper [6] introduced a polyadic logic which shares the characteristics of syllogistic. The present paper investigates surface reasoning in the context of this polyadic logic.

The principal results of this investigation are: development of a model of reasoning similar to but more general than syllogistic; and definition of strategies to streamline the reasoning process. Criteria for selection of reasoning strategies are proposed. The criteria are based on syntactic features of the problem statement.

The paper is organized as follows. First an adequate calculus for surface reasoning is presented. Next graphical domains are described in which the reasoning process finds a natural representation. Reasoning is represented as construction of fragments (subposets) of these domains. Finally, to facilitate these constructions, strategies are defined which (i) impose global restrictions or preferences on the logical operations to be used, and (ii) provide local guidance for the direction in which construction is to proceed. A number of examples are presented to illustrate these strategies.

2 A calculus for surface reasoning

The calculus presented in this section is $\mathcal{L}_N$, first described in [6]. The presentation given here follows [6] except that the axiomatization given in [6] is replaced by theorem schemas and derived rules of inference. This calculus is sound since it is based on theorems derived from a sound axiomatization. This calculus is also complete since the previous axiomatization is complete and since it can be derived from the theorem schemas and derived rules of inference to be presented.

The objectives of the design of $\mathcal{L}_N$ are (i) structural similarity to English, in the sense that well-translatable grammars [2] relating the two languages can be defined; and (ii) syntax and semantics that bring the monotonicity principle into prominence. For more on these objectives, see [6].

The language of $\mathcal{L}_N$ does not have variables, individual constants or an identity relation. The capabilities provided by these entities in conventional logic are provided by singular predicates and predicate functors in $\mathcal{L}_N$. In this regard, $\mathcal{L}_N$ is similar to natural languages. Also like natural languages, $\mathcal{L}_N$ is implicitly many-sorted.

It can be shown [5] that the pure predicate calculus without identity ($\mathcal{PP}$) is equivalent to a proper subset of $\mathcal{L}_N$, which in turn is equivalent to a proper subset of the pure predicate calculus with identity ($\mathcal{PPI}$). The deficit in expressiveness relative to $\mathcal{PPI}$ is not significant, since singular predicates provide the essential capability of the identity relation. However, the motivation for $\mathcal{L}_N$ is not to duplicate the expressiveness of $\mathcal{PPI}$, but rather to provide a perspicuous medium in which to study

those aspects of logic that play an important role in natural language understanding and human reasoning in general.

2.1 **Syntax** The alphabet of $\mathcal{L}_N$ consists of the following.

1. Predicate symbols $\mathcal{P} = \mathcal{S} \cup (\bigcup_{j \in \omega} \mathcal{R}_j)$ where $\mathcal{R}_j = \{R_i^j : i \in \omega\}$, $\mathcal{S} = \{S_i : i \in \omega\}$, and $\mathcal{S}$ and the $\mathcal{R}_j$ are mutually disjoint.

2. Selection operators $\{\langle k_1, \ldots, k_n \rangle : n \in (\omega - \{0\}), k_i \in (\omega - \{0\}), 1 \leq i \leq n\}$.

3. Boolean operators $\cap$ and $^{-}$.

4. 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. Each $S_i \in \mathcal{S}$ is a unary expression.

2. For all $n \in \omega$, each $R_i^n \in \mathcal{R}_n$ is a 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 a n-ary expression then $\overline{(X)}$ is a n-ary expression.

5. If X is a m-ary expression and Y is a l-ary expression then $(X \cap Y)$ is a n-ary expression where $n = max(l, m)$.

6. If X is a unary expression and Y is a $(n+1)$-ary expression then (XY) is a n-ary expression.

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

2.2 **Semantics** 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 $S_i \in \mathcal{S}$, $\mathcal{F}(S_i) = \{\langle d \rangle\}$ for some (not necessarily unique) $d \in \mathcal{D}$, and
2. 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 $\mathcal{I} \models_\alpha X$) 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 $\langle d_{k_1}, \ldots, d_{k_m} \rangle \models P$
3. $X = \overline{Y}$ and $\mathcal{I} \not\models_\alpha Y$
4. $X = Y \cap Z$ and $\mathcal{I} \models_\alpha Y$ and $\mathcal{I} \models_\alpha Z$
5. $X = Y^1 Z^{n+1}$ and for some $d \in \mathcal{D}$, $\langle d \rangle \models Y^1$ and $\langle d \rangle \models Z^{n+1}$

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

X is true in $\mathcal{I}$ (written $\mathcal{I} \models X$) iff $\mathcal{I} \models_\alpha X$ for every $\alpha \in \mathcal{D}^\omega$. *$X$ is valid* (written $\models X$) 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 $\mathcal{I}$.

2.3 **Abbreviations** The following abbreviations are introduced to improve readability.

1. $X \cup Y := \overline{(\overline{X} \cap \overline{Y})}$
2. $X \subseteq Y := \overline{X \cap \overline{Y}}$
3. $X \equiv Y := (X \subseteq Y) \cap (Y \subseteq X)$
4. $T := (S_0 \subseteq S_0)$
5. $X_n X_{n-1} \cdots X_1 Y := (X_n(X_{n-1} \cdots (X_1 Y) \cdots)$
6. $X^1 Y_n^2 \circ Y_{n-1}^2 \circ \cdots \circ Y_1^2 := (\cdots (X^1 Y_n^2) Y_{n-1}^2) \cdots Y_1^2)$
7. $\wedge X^1 Y := \overline{X^1 \overline{Y}}$
8. $\breve{R}^n := \langle n, \ldots, 1 \rangle R^n$

It is easy to see that:

1. $\mathcal{I} \models_\alpha X \cup Y$ iff ($\mathcal{I} \models_\alpha X$ or $\mathcal{I} \models_\alpha Y$)

2. $\mathcal{I} \models_\alpha X \subseteq Y$ iff ($\mathcal{I} \models_\alpha X$ implies $\mathcal{I} \models_\alpha Y$)

3. $\mathcal{I} \models_\alpha X \equiv Y$ iff ($\mathcal{I} \models_\alpha X$ iff $\mathcal{I} \models_\alpha Y$)

4. $\mathcal{I} \models_\alpha T$ for every $\mathcal{I}$ and α

5. $\mathcal{I} \models_\alpha X^1 Y_n^2 \circ \cdots \circ Y_1^2$ iff for some $d \in \mathcal{D}$, $\langle d \rangle \models X^1$ and $\langle d \rangle \models Y_n^2 \circ \cdots \circ Y_1^2$ where $\circ$ denotes composition of relations in $\mathcal{I}$

6. $\mathcal{I} \models_\alpha \wedge X^1 Y$ iff for all $d \in \mathcal{D}$, $\langle d \rangle \models X^1$ implies $\langle d \rangle \models Y$

2.4 **Theorem schemas and derived rules of inference** Some definitions are needed first. 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.

An occurrence of a subexpression Y^m, where $m \geq 1$, is *governed by X in W* if W is XY^m, $X\overline{Y^m}$, or $X(Y^m \cap Z^l)$, 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.

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

The *universal closure* of a n-ary expression X is the nullary expression $(\wedge T)^n X$. The theorem schemas are the universal closures of the following.

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

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

S1. SS

S2G. $\phi[S_{i_{n_1}} \cdots S_{i_1} X_1^{n_1}, \ldots, S_{i_{n_k}} \cdots S_{i_1} X_k^{n_k} / p_1, \ldots, p_k] \equiv S_{i_n} \cdots S_{i_1} \phi[X_1^{n_1}, \ldots, X_n^{n_k} / p_1, \ldots, p_k]$, where ϕ is obtained from a Boolean wff in sentential variables $p_1, \ldots, p_k$ by uniform substitution of $\cap$ for $\wedge$ and $^-$ for $\neg$, $X_1^{n_1}, \ldots, X_k^{n_k} \in \mathcal{L}_N$, $S_{i_1}, \ldots, S_{i_n} \in \mathcal{S}$, and $n = max(n_1, \ldots, n_k)$

IMAG. $X^1 Y^{n+1} \equiv T(X^1 \cap Y^{n+1})$

DIST. $(W \cap \wedge X_k \cdots \wedge X_1 Z^l) \subseteq W'$, where Y^m occurs disjunctively in W, governed by $X_k \cdots X_1$, $k \leq l \leq m$, and W' is obtained from W by replacing that occurrence of Y^m with $(Z^l \cap Y^m)$

The rules of inference are the following.

EI. From $\overline{(Z^0 \cap SX^1 \cap S_{i_n} \cdots S_{i_1} SY^{n+1})}$, where S does not occur in X^1, Y^{n+1}, or Z^0, and is distinct from $S_{i_1}, \ldots, S_{i_n}$, infer $\overline{(Z^0 \cap S_{i_n} \cdots S_{i_1} X^1 Y^{n+1})}$

MON. Let Y^m occur in W with positive (respectively, negative) polarity. Let $(\wedge T)^m (Y^m \subseteq Z^l)$ (respectively, $(\wedge T)^m (Z^l \subseteq Y^m)$), 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 TX for every governing subexpression X with an occurrence of negative polarity that was eliminated in (iii). Then from $(\wedge T)^h W$ infer $(\wedge T)^{h'} W'$, where h and h' are the arities of W and W', respectively.

Note that MON could as well be a theorem schema (similar to DIST) with modus ponens providing for detachment. But MON embodies the monotonicity principle and, to make this principle central in the calculus, is given the status of an inference rule. Because of the importance of this principle, it is appropriate to elaborate on the use of MON.

As an inference rule, MON provides a generalized detachment capability. Also, subject to the conditions stated above, instances of the schema $(\wedge T)^h(W \subseteq W')$ are immediately deducible by applying MON to instances $(\wedge T)^h(W \subseteq W)$ of BTG. The use of BTG in this deduction will be implicit in subsequent discussion.

It should be noted that an equivalent form of IMAG is $\wedge X^1 Y^{n+1} \equiv \wedge T(X^1 \subseteq Y^{n+1})$. This form of IMAG will also be used implicitly in connection with the application of MON.

From previous definitions, it follows that if the expression $\wedge YX$ occurs with positive (negative) polarity, then the occurrence of Y has negative (positive) polarity while the occurrence of X has positive (negative) polarity; if the expression $Y \subseteq X$ occurs with positive (negative) polarity, then the occurrence of Y has negative (positive) polarity while the occurrence of X has positive (negative) polarity; if the expression $Y \cup X$ occurs with positive (negative) polarity, then the occurrence of Y and the occurrence

of X both have positive (negative) polarity; and if the expression $Y \equiv X$ occurs with either positive or negative polarity, then the occurrence of Y and the occurrence of X both have positive *and* negative polarity. With these provisions, rule MON applies to expressions containing occurrences of defined operators. In this connection, singular predicates require special mention. Since $\wedge SX := \overline{S\overline{X}} \equiv S\overline{\overline{X}} \equiv SX$, any occurrence of a singular predicate can be taken to have *either* positive or negative polarity.

In reasoning by refutation, it is convenient to replace EI with its contrapositive form:

EIC. From $Z^0 \cap S_{i_n} \cdots S_{i_1} X^1 Y^{n+1}$, infer $Z^0 \cap SX^1 \cap S_{i_n} \cdots S_{i_1} SY^{n+1}$, where S does not occur in X^1, Y^{n+1}, or Z^0, and is distinct from $S_{i_1}, \ldots, S_{i_n}$

Let $\Gamma \subseteq \mathcal{L}_N$ be a set of sentences. A *deduction of* W from Γ is a finite sequence $W_1, W_2, \ldots, W_n = W$ of sentences in which each W_q is either (i) a member of Γ, (ii) an instance of a theorem schema, or (iii) the result of EI or MON applied to elements of the subsequence preceding W_q. The notation $\Gamma \vdash W$ will be used to abbreviate the assertion that there exists a deduction of the sentence W from Γ; $\Gamma \nvdash W$ will abbreviate its denial. Similarly, $\Gamma \vdash_{EIC} W$ will abbreviate the assertion that there exists a deduction of the sentence W from Γ using the rules of inference MON and EIC. It is easy to see that $\Gamma \vdash \overline{TT}$ iff $\Gamma \vdash_{EIC} \overline{TT}$.

3 A model of surface reasoning

Reasoning is viewed as theorem proving, using either direct or indirect proof methods. The objective of this section is to develop a model for reasoning in $\mathcal{L}_N$ 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. It will be assumed that $\mathcal{R}_0 = \emptyset$, i.e., that there are no nullary predicate symbols.

Any sentence of $\mathcal{L}_N$ can be *purified* [7], that is, put in a form in which all quantifiers have minimum scope. The procedure is well-known, using IMAG, DeMorgan's laws (instances of BTG), and the following lemmas, which follow directly from DIST and MON.

LEMMA 1 (schema) $X_n \cdots X_1(Y^n \cap Z^0) \equiv (X_n \cdots X_1 Y^n \cap Z^0)$. $\square$

LEMMA 2 (schema) $\wedge X_n \cdots \wedge X_1(Y^l \cap Z^m) \equiv (\wedge X_l \cdots \wedge X_1 Y^l \cap \wedge X_m \cdots \wedge X_1 Z^m)$, *where* $n = max(l, m)$. $\square$

After purification, the prime subexpressions all have the form $T(Z_1 \cap \cdots \cap Z_g)$ or $\wedge T(V_1 \cup \cdots \cup V_h)$. Putting the result of purification in disjunctive normal form yields a disjunction of expressions of the form $TX_1 \cap \cdots \cap TX_k \cap \wedge TY_1 \cap \cdots \cap \wedge TY_l$, where the X_i are conjunctions of prime subexpressions and the Y_j are disjunctions of prime subexpressions. A set $\Gamma = \{TX_1, \ldots, TX_k, \wedge TY_1, \ldots, \wedge TY_l\}$ of sentences comprising such a disjunct, or a set of sentences equivalent to these under S2G and IMAG, 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 $\wedge SX$ or $\wedge T(S \subseteq X)$. 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 $\Gamma_+ = \{TX_1, \ldots, TX_k\}$ will be called the *positive part*, and the subset $\Gamma_- = \{\wedge TY_1, \ldots, \wedge TY_l\}$ 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 $\Gamma \subseteq \mathcal{L}_N$ be a consistent set of sentences. The relation $\sqsubseteq_\Gamma$, or simply $\sqsubseteq$ when no confusion can result, is defined: $X^l \sqsubseteq Y^m :\Leftrightarrow \Gamma \vdash (\wedge T)^n(X^l \subseteq Y^m)$, where $n = max(l, m)$. It is easy to see that $\sqsubseteq$ is a quasi-order on $\mathcal{L}_N$. Moreover, if $\approx$ is defined $X \approx Y :\Leftrightarrow (X \sqsubseteq Y) \cap (Y \sqsubseteq X)$, then $\sqsubseteq$ is a partial order on $\mathcal{L}_N/\approx$. The poset $\mathbf{L}_\Gamma = \langle \mathcal{L}_N/\approx, \sqsubseteq \rangle$ is the *Lindenbaum algebra of* Γ. It can be shown (e.g., [1]) that $\mathbf{L}_\Gamma$ is a Boolean lattice with greatest and least elements $|TT|$ and $|\overline{TT}|$, 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. $|TT| = |\wedge TT| = |T| = |\langle n\rangle T|$ where $\langle n\rangle T := \langle n\rangle S_0 \subseteq \langle n\rangle S_0$

2. $|\overline{TT}| = |T\overline{T}| = |\overline{T}| = |\overline{\langle n\rangle T}|$

3. $|T^m X^n| \sqsubseteq |\overline{TT}|$ iff $|T^{m+1}X^n| \sqsubseteq |\overline{TT}|$ for $0 \leq m < n$

4. $|X^n| \not\sqsubseteq |\overline{TT}|$ iff $\Gamma \vdash T^n X^n$

5. $|TT| \sqsubseteq |X^n|$ iff $\Gamma \vdash (\wedge T)^n X^n$

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. [8]).

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. $r(XY) = r(X) + r(Y) + 1$

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 $\mathbf{L}_\Gamma$, $|TT| = |\overline{TT}|$. This would follow for example if $TX \in \Gamma_+$ 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 [3, 8]. 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 $\mathbf{L}^{(d)}_{\Gamma,1}$ which exhibits the conclusion $X \not\sqsubseteq \overline{TT}$ is proposed as a model of direct surface reasoning.

4 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. To bring the presentation closer to surface English, the following "syntactic sugar" is added to $\mathcal{L}_N$.

$$\textbf{thing} := T$$

$$\textbf{some}X^1Y := X^1Y$$

$$\textbf{all}X^1Y := {}_\wedge X^1Y$$

$$\textbf{no}X^1Y := \overline{\textbf{some}X^1Y}$$

The illustrations of reasoning will be presented graphically as subposets of $\mathbf{L}^{(d)}_{\Gamma,1}$. Expressions of $\mathcal{L}_N$ will represent their equivalence classes. In the graphical presentations, to make polarity syntactically (and visually) explicit, **no** will not be used.

4.1 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 [3] 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.

This insight has more than philosophical importance. 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.

4.2 **Cancellation strategy** As a direct consequence of theorem schemas DIST and BTG and inference rule MON, the following rule can be derived.

CANC. Let Y^m occur disjunctively in W, governed by $X_k \cdots X_1$. Let W' be obtained from W by deleting that occurrence of Y^m and all occurrences of $\wedge X_i$ that no

longer govern a subexpression. Let TX_i for every $\wedge X_i$ that was deleted. Then infer $(\wedge T)^h((W \cap \wedge X_k \cdots \wedge X_1 \overline{Y^m}) \subseteq W')$, where h is the arity of W.

This rule, used in conjunction with rule MON, is very effective for a certain class of problems. A well-known example is *Schubert's Steamroller* (see [6, 10] for details). However, CANC and MON do not constitute a complete strategy. Therefore, the *cancellation strategy* limits itself to giving *preference* to the use of this rule along with rule MON (cf. the *unit preference strategy* [11]).

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.

In standard form, $\Gamma = \{\mathbf{all}D(B\overline{O} \cup BF), \mathbf{all}D(H\overline{R} \cup BO), SD, SHR\}$. 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 $S \sqsubseteq (B\overline{O} \cup BF)$ and $S \sqsubseteq (H\overline{R} \cup BO)$. These inferences correspond to unification in conventional logic. Subsequent cancellations correspond to unit resolution. This example also illustrates direct reasoning.

4.3 Reasoning without CONV and EIC

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.*

Some problem statements do not involve variants. Others involve variants but the variants can be eliminated. An example of noneliminable variants is:

Any transitive symmetric binary relation is reflexive.

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

In those cases where the problem statement is without variants, it would seem that CONV could have no essential role in a proof. Moreover, if variants do not occur in the proof, it would seem also that EIC would not be required. This gives rise to the following conjecture.

Let $\Gamma \vdash_{EIC} W$ where $\Gamma \cup \{W\}$ is without variants. Then there exists a deduction of W from Γ which involves neither CONV nor EIC.

This matter will be considered further in a subsequent paper. It is remarked in passing that if the conjecture were valid, the language $\mathcal{L}_N$ without selection operators would be completely axiomatized by BTG, S1, S2G, IMAG, DIST, and MON.

This conjecture motivates the following strategy, which will be called the *instantiation strategy.* If the problem statement is not without variants, early use of EIC and CONV is mandated wherever necessary to establish the connection between sentences involving distinct variants of the same predicate. If such predicates are already governed by singular predicates, then only CONV need be used. Conversely, if the

problem statement is without variants, then the use of EIC and CONV is prohibited. The completeness of this strategy is dependent on the validity of the conjecture.

The next example, taken from Quine [7], 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: $\Gamma = \{\mathbf{all}AC, \mathbf{all}(W \cap C)P, \mathbf{some}(\mathbf{some}AM)\overline{P}, \mathbf{all}(\mathbf{something}\breve{M})W, \mathbf{all}(\mathbf{some}PM)P\}$. If the variant of M is not eliminated, it is necessary to use EIC to relate the sentences involving M and $\breve{M}$. The construction is shown in Figure 2. Heavy arcs represent premises; dotted arcs show the two uses of EIC. Except for one use of CONV, the lighter arcs represent inferences involving MON. Contradiction is evidenced by $b \sqsubseteq (P \cap \overline{P})$.

An equivalent standard form without variants is: $\Gamma' = \{\mathbf{all}AC, \mathbf{all}(W \cap C)P, \mathbf{some}(\mathbf{some}AM)\overline{P}, \mathbf{allthing}\ \mathbf{all}\overline{W}\ \overline{M}, \mathbf{all}(\mathbf{some}PM)P\}$. The construction based on Γ' is given in Figure 3. Heavy and lighter arcs have the same significance. It must be pointed out that this construction is no smaller and the variety of inferences no less than the previous one. The advantage afforded by this strategy is that only a subset of the possible inferences need be considered.

4.4 **Reasoning with MON only** Even when variants are eliminated, the construction for the previous example remains complex. The complexity is due to the sentence **allthing all**$\overline{W}\ \overline{M}$, i.e., "Of all things all non-women are non-mothers." An equivalent form, (**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. Sentences whose Boolean subexpressions involve only prime subexpressions of the same arity will be called *homogeneous.* A set of homogeneous sentences will also be called homogeneous. Where only homogeneous sentences are involved, MON assumes the following simpler form.

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

Let Γ be a standard form. If a subposet of $\mathbf{L}_\Gamma$ is constructed using MON, but not EI (or EIC) or any instance of a theorem schema except $(\wedge T)^h(W \subseteq W)$ and $\wedge XY \equiv \wedge T(X \subseteq Y)$, then the construction will be said to *involve MON 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.

In many cases, a problem statement can be rephrased to be homogeneous and without variants. The following is Sommers' [9] 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: $\Gamma = \{\mathbf{all}AC, \mathbf{all}(W \cap C)P, \mathbf{some}(\mathbf{some}AB)\overline{P}, \mathbf{all}(\mathbf{some}P\breve{D})P, \mathbf{all}(\mathbf{some}AB)\mathbf{some}(W \cap A)\breve{D}\}$. The last tacit assumption, $\mathbf{all}AC \subseteq \mathbf{all}(W \cap A)(W \cap C)$, is redundant since it is a valid sentence. The construction of a contradictory subposet is shown in Figure 4. It involves MONH only.

4.5 **Strategy selection** Four global strategies have been defined: breadth-first, cancellation, instantiation, and monotone. They are not completely compatible. What is wanted is a classification of standard forms by their syntactic properties that correlates with the optimal strategy. Whether such a classification exists is an open question. In its place the following approximate classification is proposed.

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 without variants, the instantiation strategy prohibits inferences using CONV and EIC. Conversely, if the standard form is not without variants, early use of EIC and CONV 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.

This classification is based partly on the conjecture given above, and partly on a survey of problems, some of which were presented above. A more precise classification is the objective of a subsequent investigation.

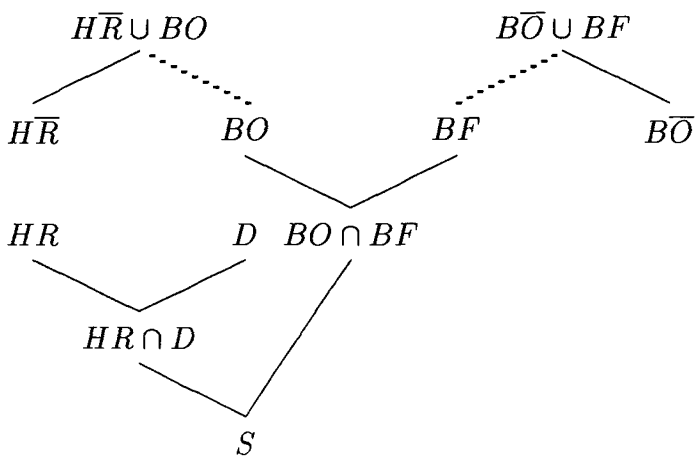

Figure 1: Example of cancellation strategy

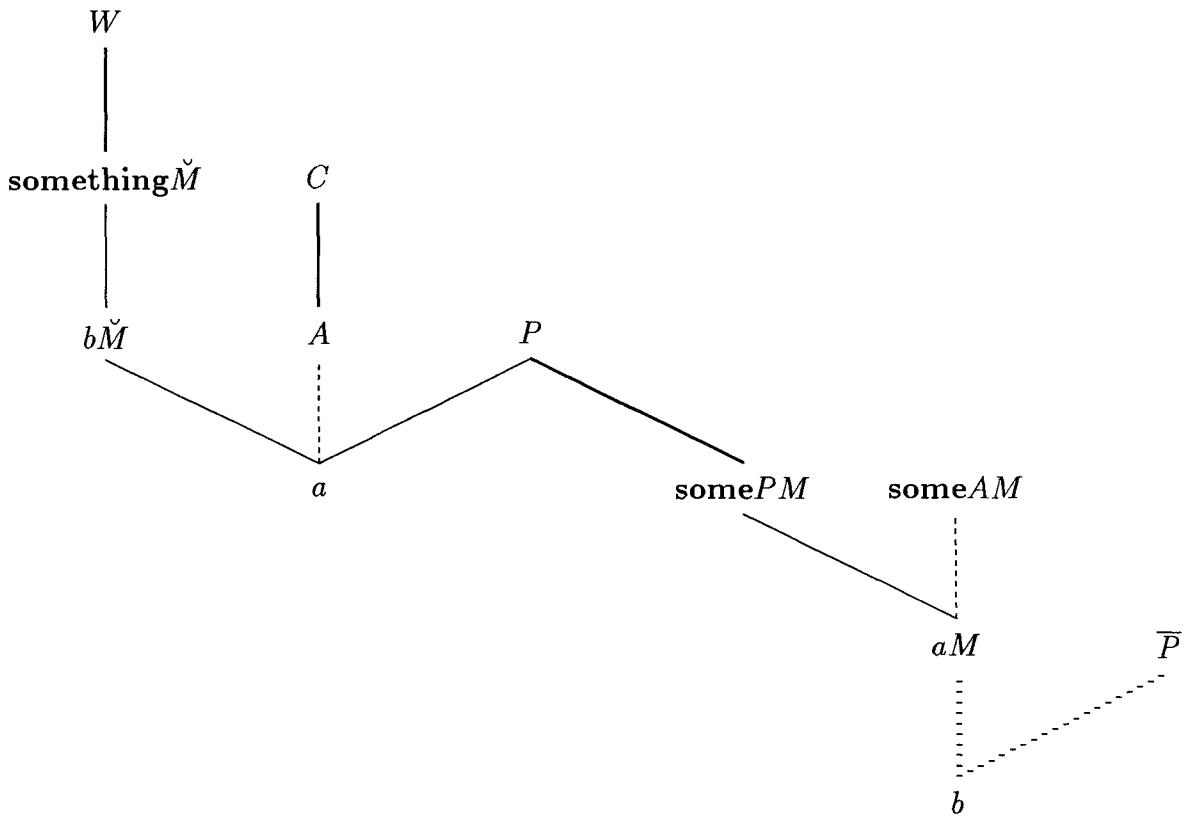

Figure 2: Example of existential instantiation

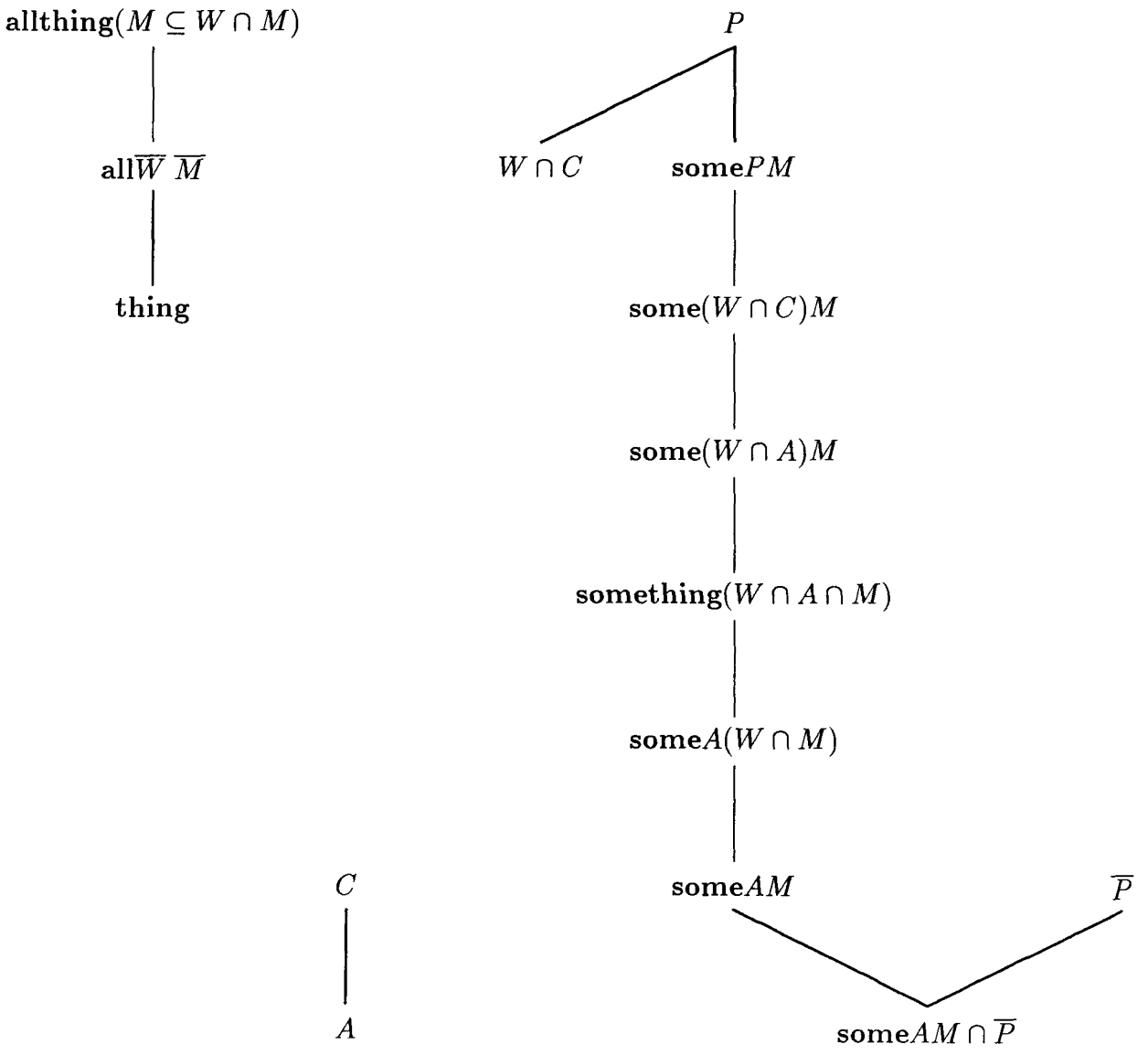

Figure 3: Previous example with variants eliminated

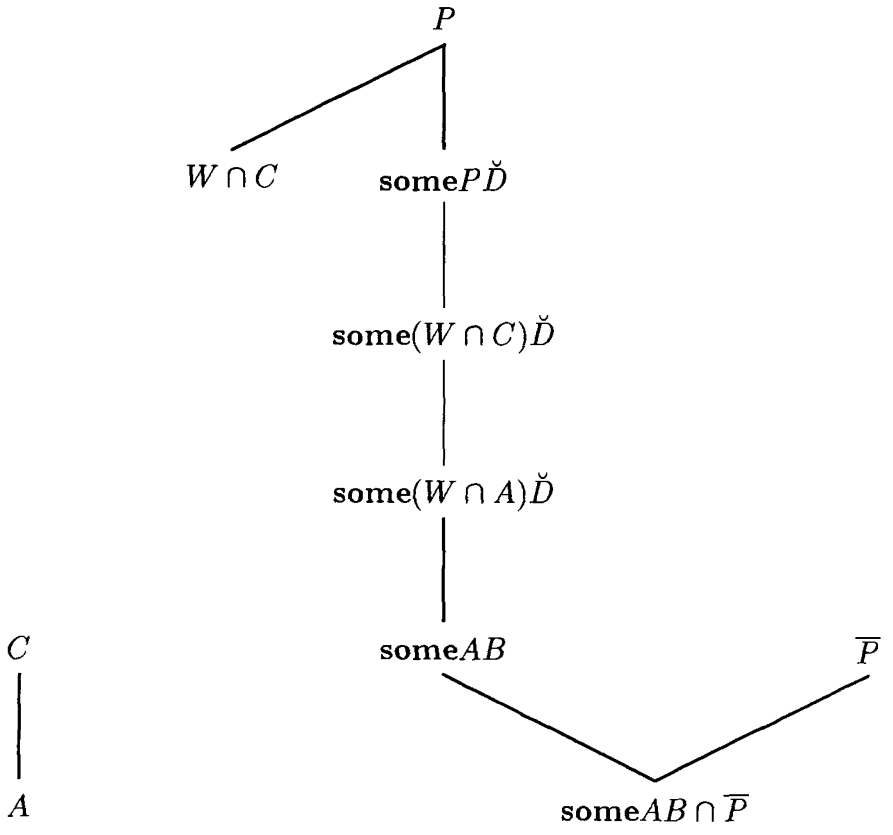

Figure 4: Previous example reformulated

5 A local strategy

The strategies of the previous section reduce the search space to the point that, for simple problems such as those considered in this paper, an exhaustive search is feasible. Nonetheless further economies are possible and in more complex problems necessary. This section considers the use of pattern matching in the subposet to guide the search for a contradiction.

The following example [7] introduces the strategy.

> The guard searched all who entered the building except those who were members of the firm. Some of Fiorecchio's men entered the building unaccompanied by anyone else. The guard searched none of Fiorecchio's men. Therefore, some of Fiorecchio's men were members of the firm.

A standard form for the problem for proof by refutation is: $\Gamma = \{\mathbf{all}(\mathbf{no}M\breve{A})S, \mathbf{some}F(\mathbf{no}\overline{F}\breve{A}), \mathbf{no}FS, \mathbf{no}FM\}$. Figure 5 shows a partial construction. The elements of Γ are represented by heavy arcs. The first inference is $\overline{S} \sqsubseteq \mathbf{some}M\breve{A}$, represented by the lighter arc. At this point, the intersection of the two chains, one containing $\mathbf{some}M\breve{A}$, and the other containing $\overline{(\mathbf{some}\overline{F}\breve{A})}$, focuses the construction. Because these two elements lie on intersecting chains and have the potential (syntactically) to produce complementary expressions, the search is directed to extension of either or both of these chains. The two possibilities are shown as dotted arcs. Both result in a contradiction.

The next example [9] illustrates a slightly different situation.

All supporters of Nixon will vote for Reagan. Avery will vote for none but a friend of Harriman. No friend of Khrushchev has Reagan for a friend. Harriman is a friend of Khrushchev. Therefore, Avery will not support Nixon.

A standard form resulting from direct translation of the premises and the denial of the conclusion is: $\Gamma = \{\mathbf{all}(N\breve{S})R\breve{V}, \mathbf{all}(AV)H\breve{F}, \mathbf{no}(K\breve{F})RF, HK\breve{F}, AN\breve{S}\}$. The variants of predicates F and V can be eliminated to obtain the equivalent standard form: $\Gamma' = \{R\,\mathbf{all}(N\breve{S})V, \mathbf{all}(AV)H\breve{F}, R\,\mathbf{no}(K\breve{F})\breve{F}, HK\breve{F}, AN\breve{S}\}$.

Figure 6, based on Γ', shows the premises as heavy arcs and the first inference as the lighter arc: $\mathbf{all}(N\breve{S})V \sqsubseteq AV$. This step connects the chains $R \sqsubseteq \mathbf{all}(N\breve{S})V$ and $AV \sqsubseteq H\breve{F}$ and results in intersecting chains with elements $\overline{\mathbf{some}(K\breve{F})\breve{F}}$ and $H\breve{F}$, respectively. The syntactic patterns $\overline{X\breve{F}}$ and $Y\breve{F}$ lead to a focusing of the search to extensions of these chains. The two possible extensions are shown as dotted arcs. Both yield contradictions.

The local strategy can now be defined as follows. As subposet construction proceeds under the appropriate global strategies, the constituent chains are monitored for the occurrence of a pair of chains having a nonempty intersection and containing expressions of the form XZ and $\overline{YZ}$, respectively. Expressions equivalent to these forms also qualify. The case in which one chain has zero length (i.e., both expressions are on the same chain) is included. Preference is then given to inferences that extend these chains upward.

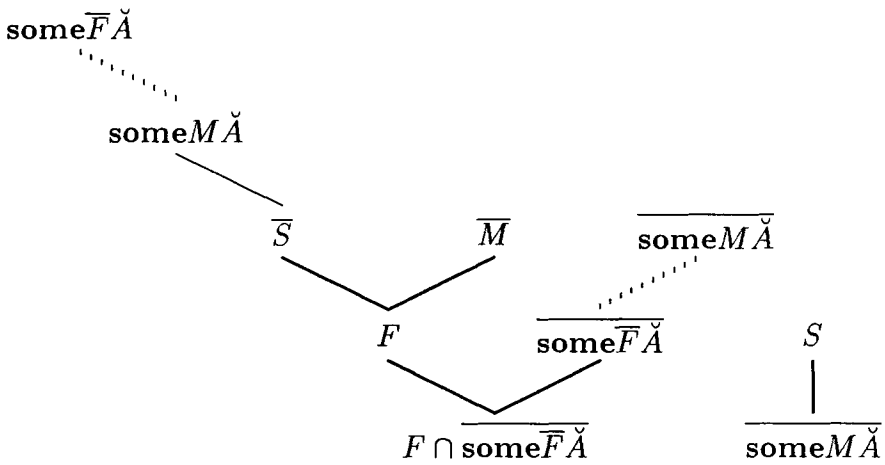

Figure 5: A first example of guiding the search

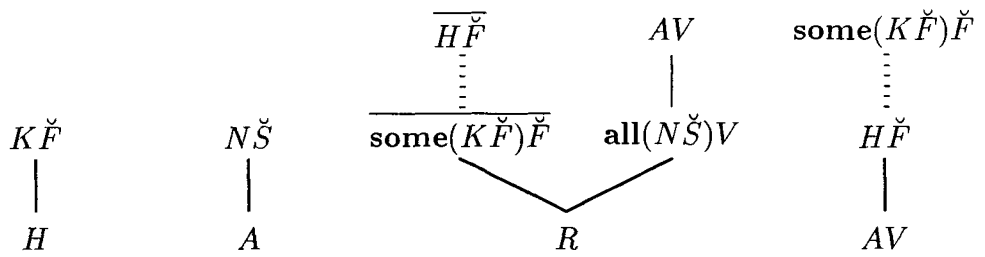

Figure 6: A second example of guiding the search

6 Conclusion

In the theory of reasoning presented in this paper, a problem statement in standard form is modeled as 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 (called *sorts* in conventional logic). This 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. $\mathcal{L}_N$ has been proposed [6] as an appropriate calculus. The model presented in this paper, together with the strategies for efficient construction of model fragments, constitutes an operational definition of an appropriate control algorithm.

It is argued that this reasoning procedure models important aspects of human reasoning. While no definite conclusion is possible on this issue, it is clear that this reasoning procedure can be automated. The means for input and output can be based on the direct intertranslatability of $\mathcal{L}_N$ and English, which facilitates construction of standard forms for problem statements as well as reporting of the reasoning process in English. Such an automated reasoning system would permit more definite conclusions regarding its success in emulating human understanding of natural language.

References

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

[2] Čulík, K "Well-Translatable Grammars and Algol-Like Languages," in Steel, T. B. Jr., ed., *Formal Language Description Languages for Computer Programming*, North Holland, 1966, 76-85.

[3] Hintikka, Jaakko "Surface Information and Depth Information," in Hintikka, J. and Suppes, P., eds., *Information and Inference*, D. Reidel, 1970, 263-297.

[4] Johnson-Laird, P. N. *Mental Models*, Harvard University Press, 1983.

[5] Purdy, William C. *A Logic for Natural Language*, Report CIS-90-02, School of Computer and Information Science, Syracuse University, 1990.

[6] Purdy, William C. "A Logic for Natural Language," to appear in *Notre Dame Journal of Formal Logic.*

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

[8] Rantala, Veikko "Constituents," in Bogdan, Radu J., ed., *Jaakko Hintikka*, D. Reidel Publishing Company, 1987, 43-76.

[9] Sommers, Fred "The Calculus of Terms," in Englebretsen, George, ed., *The New Syllogistic*, Peter Lang, 1987, 11-56.

[10] Stickel, Mark E. "Schubert's Steamroller Problem: Formulations and Solutions," *Journal of Automated Reasoning* 2 (1986), 89-101.

[11] Wos, Lawrence, D. Carson, and G. Robinson "The unit preference strategy in theorem proving," *Proceedings - Fall Joint Computer Conference* 26 (1964), 616-621.