1-1993

Binary Resolution in Surface Reasoning

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

SU-CIS-93-22

Binary Resolution in Surface Reasoning

William C. Purdy

January 1993

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

Abstract

Intuition suggests the hypothesis that everyday human reasoning is conducted in the written or spoken natural language, rather than in some disparate representation into which the surface language is translated. An examination of human reasoning reveals patterns of inference that parallel binary resolution. But any standard implementation of resolution requires Skolemization. Skolemization would seem an unlikely component of human reasoning. This appears to contradict the hypothesis that human reasoning takes place at the surface. To reconcile these observations, this paper develops a new rule of inference, which operates on surface expressions directly. This rule is shown to produce results which exactly parallel those produced by Skolemization and resolution. It extends the notion of 'surface reasoning' that was defined in previous papers. Several examples are given to illustrate its use in surface reasoning.

1 **Introduction** A common pattern of inference found in human reasoning is illustrated by the following.

All politicians either love every person or seek power.
No ultra-conservative loves a person who is poor.

All politicians who are ultra-conservative seek power.

The conclusion can be produced by binary resolution. It seems unlikely however that resolution is actually the inference mechanism. If it were, it would require translation of the premises to a subsurface form (viz., clausal form), and translation of the conclusion back to surface English. Rather, intuition suggests that the inference takes place at the surface, that is, directly in terms of the spoken English.

This type of reasoning was examined in Purdy [3] and [4], and given the name *surface reasoning.* A formal language, $\mathcal{L}_N$, was defined, which is similar in structure to surface English. Rules of inference were derived that apply to $\mathcal{L}_N$ (or to surface English) directly. The particular rule involved in the above inference is the *Cancellation Rule.*

Another, more complex but still common, pattern of inference is the following.

Every student likes some sport.
Some boy dislikes all team sports.

Some boy dislikes all team sports and either is not a student or likes some sport which is not a team sport.

Again binary resolution can yield the conclusion. But if resolution were actually the mechanism employed, it would require Skolemization of the premises as well as translation to clausal form, and the reverse for the conclusion. Indeed, all standard systems of formal reasoning involve Skolemization or, equivalently, iterated existential instantiation. Yet, involvement of procedures such as prenexing, Skolemization, and translation to clausal form seems unlikely in human reasoning. Thus intuition again suggests that this inference takes place at the surface.

This hypothesis is of interest, not only relative to human reasoning, but relative to automated reasoning systems as well. For if these patterns of inference are instances of surface reasoning, then it must be possible to reason without any form of Skolemization.

A rule by which this kind of reasoning can be accomplished at the surface level will be called a *Generalized Cancellation Rule*, since in the simplest case it will reduce to the Cancellation Rule.

The objective of this paper is to derive and study such a rule.

Since $\mathcal{L}_N$, its semantics, and its axiomatization have been presented in [3] and [4], only a brief summary is given here. This is followed by definition of Skolemization and resolution in $\mathcal{L}_N$ and development of a Generalized Cancellation Rule. It is shown that the Generalized Cancellation Rule directly parallels Skolemization and resolution. This direct parallel is termed *Skolem equivalence*. Let the premises and the result of applying the Generalized Cancellation Rule to them be the set S_1 of sentences. Let the Skolemized premises and their binary resolvent be the set S_2 of sentences. Then S_1 and S_2 are Skolem equivalent if Skolemizing S_1 yields S_2, and 'deskolemizing' S_2 yields S_1.

Application of the Generalized Cancellation Rule to surface reasoning is discussed. Two frequently encountered cases, which are similar to syllogistic reasoning, are defined. Several examples are used to illustrate surface reasoning with the Generalized Cancellation Rule and its syllogistic cases.

2 **Summary of $\mathcal{L}_N$** This section provides a brief review of $\mathcal{L}_N$. For details, see [3] and [4]. In the following, $\omega_+ := \omega - \{0\}$.

The vocabulary of $\mathcal{L}_N$ consists of the following.

1. Ordinary predicate symbols $\mathcal{R} = \bigcup_{j \in \omega_+} \mathcal{R}_j$, where $\mathcal{R}_j = \{R_i^j : i \in \omega\}$.

2. Singular predicate symbols $\mathcal{S} = \bigcup_{j \in \omega_+} \mathcal{S}_j$, where $\mathcal{S}_j = \{S_i^j : i \in \omega\}$.

3. Selection operators $\{\langle k_1, \ldots, k_n \rangle : n \in \omega_+, k_i \in \omega_+, 1 \leq i \leq n\}$.

4. Boolean operators $\cap$ and $^-$.

5. Parentheses (and).

Let $\mathcal{P} := \mathcal{R} \cup \mathcal{S}$, $\mathcal{P}_j := \mathcal{R}_j \cup \mathcal{S}_j$, and P_i^j be either R_i^j or S_i^j.

Expressions of $\mathcal{L}_N$ and their arities are simultaneously defined as follows.

1. If $P_i^n \in \mathcal{P}_n$ then P_i^n is a n-ary expression.

2. If $P_i^m \in \mathcal{P}_m$ then $\langle k_1, \ldots, k_m \rangle P_i^m$ is a n-ary expression where $n = max(k_i)_{1 \leq i \leq m}$.

3. If X is a n-ary expression then $\overline{(X)}$ is a n-ary expression.

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

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

A nullary expression is also called a *sentence.* An expression of the form YZ is called an *image.*

Metavariables will be used as follows: $k, l, m, n, \ldots$ range over ω, R^n ranges over $\mathcal{R}_n$; S^n ranges over $\mathcal{S}_n$; P^n ranges over $\mathcal{P}_n$; $X^n, Y^n, Z^n, \ldots$ range over n-ary expressions; $X, Y, Z, \ldots$ range over all expressions of $\mathcal{L}_N$; and S ranges over singular expressions (see below). Applying subscripts and primes to these symbols does not change their ranges.

Superscripts and parentheses will be dropped whenever this introduces no confusion.

$\mathcal{L}_N$ is interpreted in a first-order structure in the usual way. The relation between $\mathcal{L}_N$ and the language of predicate calculus can be given as follows. Let function τ' be defined as follows:

1. $\tau'(X^1Y^n, \sigma) = \exists x(\tau'(X^1, x\sigma) \wedge \tau'(Y^n, x\sigma))$ where x does not occur in σ

2. $\tau'(\overline{X}, \sigma) = \neg\tau'(X, \sigma)$

3. $\tau'(X \cap Y, \sigma) = \tau'(X, \sigma) \wedge \tau'(Y, \sigma)$

4. $\tau'((X), \sigma) = (\tau'(X, \sigma))$

5. $\tau'(\langle l_1, \ldots, l_n \rangle R^n, x_{k_1} \cdots x_{k_m}) = \begin{cases} R^n x_{k_{l_1}} \cdots x_{k_{l_n}} \text{ providing } \{l_1, \ldots, l_n\} \subseteq \{1, \ldots, m\} \\ \text{undefined otherwise} \end{cases}$

6. $\tau'(\langle l \rangle S, x_{k_1} \cdots x_{k_m}) = \begin{cases} Sx_{k_l} \wedge \neg(\exists x(Sx \wedge \neg(x = x_{k_l}))) \\ \quad \text{where } x \text{ does not occur in } \sigma \\ \quad \text{providing } l \in \{1, \ldots, m\} \\ \text{undefined otherwise} \end{cases}$

7. $\tau'(P, \sigma) = \tau'(\langle 1, \ldots, n \rangle P, \sigma)$ where $P \in \mathcal{P}$ is of arity n.

Then the *translation* of a sentence $X \in \mathcal{L}_N$ is defined to be $\tau'(X, \epsilon)$, where ϵ is the empty string.

$\mathcal{L}_N$ is extended by abbreviation as usual.

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. $T := (S^1_0 \subseteq S^1_0)$

5. $\wedge X^1 Y := \overline{X^1 \overline{Y}}$

6. $X_n X_{n-1} \cdots X_1 Y := (X_n(X_{n-1} \cdots (X_1 Y) \cdots)$

7. $X^1 Y_n^2 \circ Y_{n-1}^2 \circ \cdots \circ Y_1^2 := (\cdots (X^1 Y_n^2) Y_{n-1}^2) \cdots Y_1^2)$

8. $\overbrace{P^n} := \langle n, \ldots, 1 \rangle P^n$

Singular expressions of $\mathcal{L}_N$ play a role similar to that of functions in predicate logic. They are defined as follows. Each $S_i^1 \in \mathcal{S}_1$ is a singular expression. If $S_1, \ldots, S_n$ are singular expressions, then for each $S_i^{n+1} \in \mathcal{S}_{n+1}$, $\overbrace{S_1 \cdots S_n\, S_i^{n+1}}$ is a singular expression.

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.

Several identities and derived rules of inference will be needed in the sequel. They are stated here without proof. For proofs, see [3] and [4].

Identities:

$XY \equiv T(X \cap Y) \qquad \wedge XY \equiv \wedge T(\overline{X} \cup Y)$

Tautology Rule:

Let X^n be obtained from a Boolean tautology by uniform substitution of expressions of $\mathcal{L}_N$ for sentential variables, $\cap$ for $\wedge$ and $^-$ for $\neg$. Then infer $(\wedge T)^n X^n$.

Monotonicity Rule:

Let Y^m occur in sentence 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 W infer W'.

Distributivity Rule:

From $\wedge T(X_n \cup \cdots \cup \wedge T(X_1 \cup Y^n) \cdots) \cap \wedge T(X_n \cup \cdots \cup \wedge T(X_1 \cup V^n) \cdots)$ infer $\wedge T(X_n \cup \cdots \cup \wedge T(X_1 \cup (Y^n \cap V^n) \cdots)$.

Thinning Rule:

From ${}_\wedge T(X_n \cup \cdots \cup {}_\wedge T(X_i \cup \cdots \cup {}_\wedge T(X_1 \cup Y^n) \cdots)$ infer ${}_\wedge T(X_n \cup \cdots \cup {}_\wedge T(X_i \cup Z \cup \cdots \cup {}_\wedge T(X_1 \cup Y^n) \cdots)$.

Singularity Rule:

From ${}_\wedge T(X_n \cup \cdots \cup {}_\wedge T(\overline{\langle 1, k_1, \ldots, k_m \rangle S^{m+1}} \cup {}_\wedge T(X_{i+1} \cup \cdots \cup {}_\wedge T(X_1 \cup Y^n) \cdots)$ infer ${}_\wedge T(X_n \cup \cdots \cup T\langle 1, k_1, \ldots, k_m \rangle S^{m+1} \cap {}_\wedge T(X_{i+1} \cup \cdots \cup {}_\wedge T(X_1 \cup Y^n) \cdots)$.

The Cancellation Rule, although not used, is frequently mentioned. It is stated below. It uses the notion of governance, which is defined in the next section.

Cancellation Rule:

Let Y^m occur in sentence W, governed by $-X_k \cdots -X_1$. Let ${}_\wedge X_k \cdots {}_\wedge X_1 \overline{Y^m}$. Let W' be obtained from W by deleting that occurrence of Y^m and all occurrences of governors $-X_i$ that no longer govern a subexpression. Let TX_i for every governor $-X_i$ that was deleted. Then from W infer W'.

3 Governance This section develops a relation called *governs*. Following common practice, in this and subsequent sections any conjunction (disjunction) will be considered to be the *set* of its conjuncts (disjuncts). More precisely, a conjunction will be considered to be an equivalence class that identifies the conjunctions $(X \cap V)$, $(V \cap X)$, $(X \cap X \cap V)$, and $(X \cap V \cap (Y \cup \overline{Y}))$, and dually for disjunctions.

Let W be an expression in which Y occurs as a subexpression. If there is more than one occurrence of Y in W, the particular one of these occurrences intended will be indicated (e.g., by an integer specifying the position of that occurrence of Y in the structural description of W). Thereafter any reference 'Y' will be a reference to that particular occurrence. W is *normalized with respect to* (a particular occurrence of) Y if (i) every subexpression of W containing Y is of one of the forms $T(X \cap V)$ or ${\wedge}T(X \cup V)$, where V contains Y as a subexpression; (ii) V is either Y itself, or an image, or the complement of one of these; and (iii) neither $T(X \cap V)$ nor ${\wedge}T(X \cup V)$ contains a nullary proper subexpression. W can always be normalized with respect to Y by repeated use of the following identities.

$$
\begin{array}{ll}
XZ \equiv T(X \cap Z) & (\overline{X \cap Z}) \equiv (\overline{X} \cup \overline{Z}) \\
T(X \cup Z) \equiv (TX \cup TZ) & \overline{\overline{X}} \equiv X \\
T(X \cap (Z \cup V)) \equiv (T(X \cap Z) \cup T(X \cap V)) & \overline{XZ} \equiv {\wedge}X\overline{Z} \\
T(X^0 \cap Z) \equiv (X^0 \cap TZ) &
\end{array}
$$

Other forms of these identities can be obtained for the abbreviations defined above. For example, from $T(X \cup Z) \equiv (TX \cup TZ)$ one can obtain ${\wedge}T(X \cap Z) \equiv ({\wedge}TX \cap {\wedge}TZ)$. As an illustration of normalization, the expression ${\wedge}X_3X_2{\wedge}(X_1Y)(V \cup Z)$, normalized with respect to Y, is ${\wedge}T(\overline{X_3} \cup T(X_2 \cap {\wedge}T(V \cup Z \cup {\wedge}T(\overline{X_1} \cup \overline{Y}))))$. The normalized form of W with respect to Y is denoted $\mathbf{Nm}[Y](W)$. Since identities only are used to obtain it, $\mathbf{Nm}[Y](W)$ is equivalent to W.

A subexpression Y^m (or $\overline{Y^m}$) of W is *governed by* $g_n \cdots g_1$ if $\mathbf{Nm}[Y](W)$ has the form

$$\cdots {\wedge}T(X_i \cup V) \cdots) \text{ and } g_i = -\overline{X_i}, \text{ or}$$

$$\cdots T(X_i \cap V) \cdots) \text{ and } g_i = +X_i$$

for $1 \leq i \leq n$. Each g_i is referred to as a *governor of* Y^m, and the string $g_n \cdots g_1$ as the *governance* of Y^m. In case W is Y^m (or $\overline{Y^m}$) itself, Y^m (or $\overline{Y^m}$) is said to be governed by the empty string, ϵ. The subexpression ${}_\wedge T(X_i \cup V)$ or $T(X_i \cap V)$ will be referred to as the subexpression at *level i*. For example, continuing the illustration above, $\overline{Y}$ is governed in W by $-X_3 + X_2 - (\overline{V \cup Z}) - X_1$. Three simple observations about governance are the following. First, n and m are in general unrelated. If W is nullary, then $m \leq n$. But if not, it can be that $n < m$ as well. For example, in the illustration, Y is of arity 2, but it is governed by a string of length 4, since $V \cup Z$ is of at least arity 3. Second, in case X_i is a singular expression, S, the identity ${}_\wedge T(\overline{S} \cup V) \equiv T(S \cap V)$ implies that g_i can be taken as either $-S$ or $+S$. Third, from the governance and the nullary or sentential level expression, the normalized form of W can be reconstructed. (It may be that W is a conjunction or disjunction, in which case it is assumed that each sentence is dealt with separately.)

The notion of the normalization of an expression W with respect to a subexpression Y was defined to simplify the definition of the governance of Y in W. However, it is not necessary to normalize W to obtain the governance of Y. This is now explained.

Let $\alpha = g_n \cdots g_1$. Define $(+X)^R := -X$, $(-X)^R := +X$, and $\alpha^R := g_n^R \cdots g_1^R$. Let Y occur in the arbitrary expression V, governed by α. Then from the definition given above, Table 1 is constructed.

The algorithm implied by Table 1 can be demonstrated using the expression ${}_\wedge X_3 X_2 {}_\wedge (X_1 Y)(V \cup Z)$.

Y is governed in $X_1 Y$ by $+X_1$ (row 2 of the table)

$\overline{Y}$ is governed in ${}_\wedge(X_1 Y)(V \cup Z)$ by $-(\overline{V \cup Z}) - X_1$ (row 5)

$\overline{Y}$ is governed in $X_2 {}_\wedge(X_1 Y)(V \cup Z)$ by $+X_2 - (\overline{V \cup Z}) - X_1$ (row 2)

$\overline{Y}$ is governed in ${}_\wedge X_3 X_2 {}_\wedge(X_1 Y)(V \cup Z)$ by $-X_3 + X_2 - (\overline{V \cup Z}) - X_1$ (row 3)

If V occurs in $\overline{V}$	then $\overline{Y}$	is governed in $\overline{V}$	by α^R
XV	Y	XV	$+X\alpha$
$\wedge XV$	Y	$\wedge XV$	$-X\alpha$
VX	Y	VX	$+X\alpha$
$\wedge VX$	$\overline{Y}$	$\wedge VX$	$-\overline{X}\alpha^R$
$X(Z \cap V)$	Y	$X(Z \cap V)$	$+(X \cap Z)\alpha$
$X(Z^0 \cap V)$	Y	$Z^0 \cap XV$	$+X\alpha$
$X(Z \cap V^0)$	Y	$XZ \cap V^0$	α
$\wedge X(Z \cup V)$	Y	$\wedge X(Z \cup V)$	$-(X \cap \overline{Z})\alpha$
$\wedge X(Z^0 \cup V)$	Y	$Z^0 \cup \wedge XV$	$-X\alpha$
$\wedge X(Z \cup V^0)$	Y	$\wedge XZ \cup V^0$	α
$X(Z \cup V)$	Y	$XZ \cup XV$	$+X\alpha$
$X(Z^0 \cup V)$	Y	$(TX \cap Z^0) \cup XV$	$+X\alpha$
$X(Z \cup V^0)$	Y	$XZ \cup (TX \cap V^0)$	α
$\wedge X(Z \cap V)$	Y	$\wedge XZ \cap \wedge XV$	$-X\alpha$
$\wedge X(Z^0 \cap V)$	Y	$Z^0 \cap \wedge XV$	$-X\alpha$
$\wedge X(Z \cap V^0)$	Y	$\wedge XZ \cap V^0$	α
$(Z \cap V)X$	Y	$(Z \cap V)X$	$+(X \cap Z)\alpha$
$(Z^0 \cap V)X$	Y	$Z^0 \cap VX$	$+X\alpha$
$(Z \cap V^0)X$	Y	$ZX \cap V^0$	α
$\wedge(Z \cup V)X$	$\overline{Y}$	$\wedge ZX \cap \wedge VX$	$-\overline{X}\alpha^R$
$\wedge(Z^0 \cup V)X$	$\overline{Y}$	$Z^0 \cap \wedge VX$	$-\overline{X}\alpha^R$
$\wedge(Z \cup V^0)X$	$\overline{Y}$	$\wedge ZX \cap V^0$	α^R
$(Z \cup V)X$	Y	$ZX \cup VX$	$+X\alpha$
$(Z^0 \cup V)X$	Y	$(TX \cap Z^0) \cup VX$	$+X\alpha$
$(Z \cup V^0)X$	Y	$ZX \cup (TX \cap V^0)$	α
$\wedge(Z \cap V)X$	$\overline{Y}$	$\wedge(Z \cap V)X$	$-(\overline{X} \cap Z)\alpha^R$
$\wedge(Z^0 \cap V)X$	$\overline{Y}$	$\overline{Z^0} \cup \wedge VX$	$-\overline{X}\alpha^R$
$\wedge(Z \cap V^0)X$	$\overline{Y}$	$\wedge ZX \cup \overline{V^0}$	α^R

Table 1: Rules for governance assuming that Y is governed in V by α.

4 Resolution

In this section, some standard results of resolution theory are translated to $\mathcal{L}_N$. These results are used in the next section as a point of departure for generalizing the Cancellation Rule.

Let W be a sentence, normalized with respect to subexpression Y. Then $\mathbf{Sk}[Y](W)$, the *Skolemization of W with respect to* (that occurrence of) Y, is defined as follows. (It is assumed that Y occurs in V.)

$\mathbf{Sk}[Y](_\wedge T(X \cup V)) := {}_\wedge T(X \cup \mathbf{Sk}[Y](V))$

$\mathbf{Sk}[Y](T(X_i \cap V)) := {}_\wedge T(\overline{f} \cup (X_i \cap V \cap \mathbf{Sk}[Y](V))$, where f denotes a fresh singular predicate (a *Skolem constant*) of arity one greater than the number of negative governors of level greater than i, with an appropriate selection operator

$\mathbf{Sk}[Y](Y) := Y$

$\mathbf{Sk}[Y](\overline{Y}) := \overline{Y}$

It is evident that normalization of $\mathbf{Sk}[Y](W)$ will yield expressions with the property that every proper subexpression containing Y is of the form $_\wedge T(X \cup V)$, where V contains that occurrence of Y. Skolemization as defined here differs in two inessential ways from the usual definition. First, Skolemization here is only partial; specifically, only those Skolem constants are introduced that are necessary to achieve the property just stated. Second, the results at each step in the Skolemization procedure are preserved by conjoining X_i, V, and $\mathbf{Sk}[Y](V)$ (see the second part of the definition of $\mathbf{Sk}[Y]$). It is a standard result (e.g., Andrews [1], Corollary 3302) that W is satisfiable iff $\mathbf{Sk}[Y](W)$ is satisfiable. For example, the Skolemization of

$$W = {}_\wedge T(X_3 \cup T(X_2 \cap {}_\wedge T(V \cup Z \cup T(X_1 \cap \overline{Y}))))$$

is

$$W' = {}_\wedge T(X_3 \cup {}_\wedge T(\overline{f^2} \cup (X_2 \cap {}_\wedge T(V \cup Z \cup T(X_1 \cap \overline{Y})) \cap {}_\wedge T(V \cup Z \cup {}_\wedge T(\overline{g^3} \cup (X_1 \cap \overline{Y}))))))$$

g^3 denotes $\langle 1,2,4\rangle S_i^3$ and f^2 denotes $\langle 1,2\rangle S_j^2$, where S_i^3 and S_j^2 are fresh singular predicate symbols. When normalized again, W' yields the conjunction of the following.

$$W'\mathbf{3}' = {}_{\wedge}T(X_3 \cup {}_{\wedge}T(\overline{f^2} \cup X_2))$$

$$W'\mathbf{3} = {}_{\wedge}T(X_3 \cup {}_{\wedge}T(\overline{f^2} \cup {}_{\wedge}T(V \cup Z \cup T(X_1 \cap \overline{Y}))))$$

$$W'\mathbf{1}' = {}_{\wedge}T(X_3 \cup {}_{\wedge}T(\overline{f^2} \cup {}_{\wedge}T(V \cup Z \cup {}_{\wedge}T(\overline{g^3} \cup X_1))))$$

$$W'\mathbf{1} = {}_{\wedge}T(X_3 \cup {}_{\wedge}T(\overline{f^2} \cup {}_{\wedge}T(V \cup Z \cup {}_{\wedge}T(\overline{g^3} \cup \overline{Y}))))$$

This example illustrates a notational convention that will be followed in the sequel. The sentences that result from first Skolemizing and then normalizing an expression will be called the *constituents of Skolemization with respect to* Y. A constituent will be denoted by suffixing in bold type the level of the positive governor that generated that constituent. As shown in the example, the suffix will be primed or unprimed according as that constituent contains the positive governing subexpression or the subexpression that it governs. Thus if Y is governed in W by the positive subexpressions $X_{i_k}, \ldots, X_{i_1}$, then the results of Skolemizing and normalizing will be denoted $W\boldsymbol{i}'_k, W\boldsymbol{i}_k, \ldots, W\boldsymbol{i}'_1, W\boldsymbol{i}_1$. $W\boldsymbol{i}_1$ will be called the *principal constituent.*

Let

$$W_1 = {}_{\wedge}T(X_m \cup {}_{\wedge}T(X_{m-1} \cup \cdots \cup {}_{\wedge}T(X_1 \cup Y) \cdots)$$
$$W_2 = {}_{\wedge}T(Z_l \cup {}_{\wedge}T(Z_{l-1} \cup \cdots \cup {}_{\wedge}T(Z_1 \cup \overline{Y}) \cdots)$$

be principal constituents of Skolemization with respect to Y. Further, let $n = max(l, m)$ and

$$U_i = \begin{cases} X_i & \text{if } l < i \leq n \\ Z_i & \text{if } m < i \leq n \\ X_i \cup Z_i & \text{otherwise} \end{cases}$$

If for $1 \leq i \leq n$, U_i contains at most one Skolem constant, then the *binary resolvent* of W_1, W_2 with respect to Y is defined

$$\mathbf{Res}[Y](W_1, W_2) := {}_{\wedge}T(U_n \cup {}_{\wedge}T(U_{n-1} \cup \cdots \cup {}_{\wedge}TU_1) \cdots)$$

Y is the *expression resolved upon.* It is convenient to extend the definition to subexpressions so that if

$$W_1 = {}_\wedge T(X_m \cup \cdots \cup {}_\wedge T(X_i \cup V_{1,i}) \cdots)$$

$$W_2 = {}_\wedge T(Z_l \cup \cdots \cup {}_\wedge T(Z_i \cup V_{2,i}) \cdots)$$

then

$$\mathbf{Res}[Y](W_1, W_2) = {}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_i \cup \mathbf{Res}[Y](V_{1,i}, V_{2,i})) \cdots)$$

It is a standard result that $(W_1 \cap W_2) \subseteq \mathbf{Res}[Y](W_1, W_2)$. This is proved as follows. Thinning W_1, W_2 to conform to $\mathbf{Res}[Y](W_1, W_2)$ yields

$$W_1' = {}_\wedge T(U_n \cup {}_\wedge T(U_{n-1} \cup \cdots \cup {}_\wedge T(U_1 \cup Y) \cdots)$$

$$W_2' = {}_\wedge T(U_n \cup {}_\wedge T(U_{n-1} \cup \cdots \cup {}_\wedge T(U_1 \cup \overline{Y}) \cdots)$$

By the Distributive Rule,

$$W_1' \cap W_2' = {}_\wedge T(U_n \cup {}_\wedge T(U_{n-1} \cup \cdots \cup {}_\wedge T(U_1 \cup (Y \cap \overline{Y})) \cdots)$$

This use of thinning and the Distributive Rule will be used frequently in subsequent sections.

5 **Generalizing the Cancellation Rule** This section introduces a generalized cancellation rule, which is shown to be equivalent to binary resolution. But this rule differs from binary resolution in that its arguments need not be Skolemized. At first it will be assumed that the arguments are normalized. Later it will be seen that this too is unnecessary.

The rule is motivated as follows. Let W_1 and W_2 be normalized with respect to Y. Let Y occur in W_1 and W_2 with opposite polarities. Let the rightmost occurrence of a positive governor in either W_1 or W_2 be at level i. Suppose it occurs in W_1. Let

$$\mathbf{Sk}[Y](W_1) = {}_\wedge T(X_m \cup \cdots \cup {}_\wedge T(X_{i-1} \cup {}_\wedge T(\overline{f} \cup (X_i \cap V_{1,i}))\cdots)$$

Then

$$W_1\boldsymbol{i}' = {}_\wedge T(X_m \cup \cdots \cup {}_\wedge T(X_{i-1} \cup {}_\wedge T(\overline{f} \cup X_i)\cdots)$$

$$W_1\boldsymbol{i} = {}_\wedge T(X_m \cup \cdots \cup {}_\wedge T(X_{i-1} \cup {}_\wedge T(\overline{f} \cup V_{1,i})\cdots)$$

Let the principal constituent of the Skolemization of W_2 with respect to Y be

$$W_2\boldsymbol{j} = {}_\wedge T(Z_l \cup \cdots \cup {}_\wedge T(Z_{i-1} \cup {}_\wedge T(Z_i \cup V_{2,i})\cdots)$$

Let

$$\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j}) = {}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{f} \cup Z_i \cup \mathbf{Res}[Y](V_{1,i}, V_{2,i}))\cdots)$$

Now thinning $W_1\boldsymbol{i}'$ and $W_1\boldsymbol{i}$ to conform to $\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j})$ through level $i-1$ yields

$$_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{f} \cup X_i)\cdots)$$

$$_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{f} \cup V_{1,i})\cdots)$$

By the Distributive Rule it follows that

$$_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{f} \cup (X_i \cap V_{1,i} \cap (Z_i \cup \mathbf{Res}[Y](V_{1,i}, V_{2,i}))\cdots)$$

Because f is singular, this is equivalent to

$$_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup T(f \cap X_i \cap V_{1,i} \cap (Z_i \cup \mathbf{Res}[Y](V_{1,i}, V_{2,i}))\cdots)$$

Hence by the Monotonicity Rule it follows that

$$\wedge T(U_n \cup \cdots \cup \wedge T(U_{i-1} \cup T(X_i \cap V_{1,i} \cap (Z_i \cup \mathbf{Res}[Y](V_{1,i}, V_{2,i}))) \cdots)$$

This latter expression conveys all that which was deduced relative to the Skolem constant f. That is, f as a (parameterized) name, links $W_1\boldsymbol{i}'$, $W_1\boldsymbol{i}$ and $\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j})$, and so asserts the existence of something satisfying these three expressions. The last result asserts the existence of something satisfying the same expressions, but without explicitly naming it. Indeed, if the last result is Skolemized and then normalized, it yields $W_1\boldsymbol{i}'$ and $W_1\boldsymbol{i}$ (thinned), and $\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j})$, up to the choice of Skolem constants.

The same can be repeated with the new rightmost Skolem constant, and can be continued until there are no Skolem constants left. The final result is an expression that asserts precisely that which is asserted by all the constituents of Skolemization of W_1 and W_2 along with $\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j})$. The process of thinning each of the constituents of Skolemization of W_1 and W_2 to conform to $\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j})$ and then combining them with $\mathbf{Res}[Y](W_1\boldsymbol{i}, W_2\boldsymbol{j})$ using the Distributive Rule will be referred to as *collecting and eliminating Skolem constants.*

This process suggests Rule $\mathbf{R}[Y]$, the *Generalized Cancellation Rule*, defined in Table 2. Indeed the following theorem claims that Rule $\mathbf{R}[Y]$, applied directly to W_1, W_2, produces a result which is in a certain sense equivalent to binary resolution.

THEOREM 1 *Let W_1, W_2 be normalized with respect to Y, and let Y occur in W_1, W_2 with opposite polarities. Let $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$ and $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$ be the constituents of Skolemization of W_1, W_2 respectively with respect to Y. Then W_1, W_2 and $\mathbf{R}[Y](W_1, W_2)$ are* Skolem equivalent *to $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$, $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$, and $\mathbf{Res}[Y](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$. Specifically,*
(i) collecting and eliminating Skolem constants in $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$, $W_2\boldsymbol{j}'_l$, $W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$, and $\mathbf{Res}[Y](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$ yields W_1, W_2 and $\mathbf{R}[Y](W_1, W_2)$, and

$$\mathbf{R}[Y](T(X_i \cap V_{1,i}), W_2) = T(X_i \cap \mathbf{R}[Y](V_{1,i}, W_2)) \text{ if level of } W_2 < i$$

$$\mathbf{R}[Y](_\wedge T(X_i \cup V_{1,i}), W_2) = {}_\wedge T(X_i \cup \mathbf{R}[Y](V_{1,i}, W_2)) \text{ if level of } W_2 < i$$

$$\mathbf{R}[Y](W_1, T(Z_i \cap V_{2,i})) = T(Z_i \cap \mathbf{R}[Y](W_1, V_{2,i})) \text{ if level of } W_1 < i$$

$$\mathbf{R}[Y](W_1, {}_\wedge T(Z_i \cup V_{2,i})) = {}_\wedge T(Z_i \cup \mathbf{R}[Y](W_1, V_{2,i})) \text{ if level of } W_1 < i$$

$$\mathbf{R}[Y](T(X_i \cap V_{1,i}), {}_\wedge T(Z_i \cup V_{2,i})) = T(X_i \cap V_{1,i} \cap (Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

$$\mathbf{R}[Y](_\wedge T(X_i \cup V_{1,i}), T(Z_i \cap V_{2,i})) = T(Z_i \cap V_{2,i} \cap (X_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

$$\mathbf{R}[Y](_\wedge T(X_i \cup V_{1,i}), {}_\wedge T(Z_i \cup V_{2,i})) = {}_\wedge T(X_i \cup Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

$$\mathbf{R}[Y](T(X_1 \cap Y), {}_\wedge T(Z_1 \cup \overline{Y})) = T(X_1 \cap Y \cap Z_1)$$

$$\mathbf{R}[Y](_\wedge T(X_1 \cup Y), T(Z_1 \cap \overline{Y})) = T(Z_1 \cap \overline{Y} \cap X_i)$$

$$\mathbf{R}[Y](_\wedge T(X_1 \cup Y), {}_\wedge T(Z_1 \cup \overline{Y})) = {}_\wedge T(X_1 \cup Z_1)$$

Table 2: Definition of Rule $\mathbf{R}[Y]$

(ii) Skolemizing W_1, W_2 and $\mathbf{R}[Y](W_1, W_2)$ yields, up to choice of Skolem constants, $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$, $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$, and $\mathbf{Res}[Y](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$.

proof: Proof is by induction on p, the total number of positive governors in W_1, W_2 of Y and $\overline{Y}$.

The basis case ($p = 0$) is trivial. But it is of interest to note that this case is precisely the Cancellation Rule (see [3] and [4]) which Rule $\mathbf{R}[Y]$ generalizes.

For the induction step, $p > 0$. Let

$$\mathbf{Res}[Y](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1) = {\wedge}T(U_n \cup \cdots \cup {\wedge}T(U_{i-1} \cup {\wedge}T(\overline{f} \cup Z_i \cup \mathbf{Res}[Y](V_{1,i}, V_{2,i})) \cdots)$$

where f is the leftmost occurrence of a Skolem constant. f came from Skolemization of either W_1 or W_2. Since the same argument applies in either case, suppose that f came from W_1. That is, suppose that $i = i_k$ and

$$W_1 = {\wedge}T(X_m \cup \cdots \cup T(X_i \cap V_{1,i}) \cdots)$$

so that

$$W_1\boldsymbol{i}' = {\wedge}T(X_m \cup \cdots \cup {\wedge}T(\overline{f} \cup X_i) \cdots)$$

$$W_1\boldsymbol{i} = {\wedge}T(X_m \cup \cdots \cup {\wedge}T(\overline{f} \cup V_{1,i}) \cdots)$$

(i) The induction hypothesis applies to $W_1\boldsymbol{i}$ and W_2 since the total number of positive governors is $p - 1$. Then by the induction hypothesis, collecting and eliminating Skolem constants in $W_1\boldsymbol{i}'_{k-1}, W_1\boldsymbol{i}_{k-1}, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$, $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$, and $\mathbf{Res}[Y](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$ yields $W_1\boldsymbol{i}$, W_2, and

$$\mathbf{R}[Y](W_1\boldsymbol{i}, W_2) = {\wedge}T(U_n \cup \cdots \cup {\wedge}T(U_{i-1} \cup {\wedge}T(\overline{f} \cup Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})) \cdots)$$

Now collecting and eliminating the Skolem constant f in $W_1\boldsymbol{i}'$, $W_1\boldsymbol{i}$, and $\mathbf{R}[Y](W_1\boldsymbol{i}, W_2)$ yields W_1 and

$${\wedge}T(U_n \cup \cdots \cup {\wedge}T(U_{i-1} \cup T(X_i \cap V_{1,i} \cap (Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})) \cdots)$$

which is equal to $\mathbf{R}[Y](W_1, W_2)$.

(ii) Let Skolemization of W_1 yield $W_1\boldsymbol{i}'$ and $W_1\boldsymbol{i}$ as above. Skolemizing $\mathbf{R}[Y](W_1, W_2)$ with respect to $(Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i}))$ yields ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{g} \cup X_i) \cdots)$, ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{g} \cup V_{1,i}) \cdots)$, and ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{g} \cup Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})) \cdots)$. Now h is defined equal to g over the domain described by these three expressions, and equal to f elsewhere. Specifically, ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(g \equiv h)) \cdots)$, and ${}_\wedge T(\overline{U_n} \cup \cdots \cup {}_\wedge T(X_m \cup {}_\wedge T(X_{m-1} \cup \cdots \cup {}_\wedge T(X_{i-1} \cup {}_\wedge T(f \equiv h)) \cdots)$, $\ldots$, ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(X_m \cup \overline{Z_m} \cup {}_\wedge T(X_{m-1} \cup \cdots \cup {}_\wedge T(X_{i-1} \cup {}_\wedge T(f \equiv h)) \cdots)$, ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(X_m \cup {}_\wedge T(X_{m-1} \cup \overline{Z_{m-1}} \cup \cdots \cup {}_\wedge T(X_{i-1} \cup {}_\wedge T(f \equiv h)) \cdots)$, $\ldots$, ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(X_m \cup {}_\wedge T(X_{m-1} \cup \cdots \cup {}_\wedge T(X_{i-1} \cup \overline{Z_{i-1}} \cup {}_\wedge T(f \equiv h)) \cdots)$. Then h is the desired Skolem constant, that is, it satisfies ${}_\wedge T(X_m \cup \cdots \cup {}_\wedge T(\overline{h} \cup X_i) \cdots)$, ${}_\wedge T(X_m \cup \cdots \cup {}_\wedge T(\overline{h} \cup V_{1,i}) \cdots)$, and ${}_\wedge T(U_n \cup \cdots \cup {}_\wedge T(U_{i-1} \cup {}_\wedge T(\overline{h} \cup Z_i \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})) \cdots)$. Finally, the induction hypothesis is applied to complete the proof.

6 **Selection operators** Up to this point it was not necessary to consider selection operators in connection with Rule $\mathbf{R}[Y]$. However if the expression resolved upon is a predicate symbol P, and the occurrences of the predicate symbol are variants of each other, then special consideration must be given to the selection operators in applying Rule $\mathbf{R}[P]$. This section derives a necessary and sufficient condition for Rule $\mathbf{R}[P]$ to apply when variants of P are present. Moreover, it is shown that this same condition is necessary and sufficient for resolution of the sentences to be possible. When the condition is not met, the result is an 'occur-check'.

The role of selection operators in $\mathcal{L}_N$ is the same as the role of bound variables in predicate calculus. If $\langle k_1, \ldots, k_m \rangle P^m$ occurs in sentence W governed by $g_q \cdots g_1$, where $max(k_i)_{1 \leq i \leq m} = n \leq q$, then arguments $1, 2, \ldots, m$ of P^m are *filled by* $g_{k_1}, g_{k_2}, \ldots, g_{k_m}$, respectively. A different sequence can be imposed by use of a different selection operator. By definition, $\langle 1, 2, \ldots, m \rangle P^m$ is equivalent to P^m. So if P^m occurs in a sentence W governed by $g_q \cdots g_1$, arguments $1, 2, \ldots, m$ of P^m are filled by $g_1, g_2, \ldots, g_m$, respectively.

If an integer between 1 and $max(k_i)_{1 \leq i \leq m}$ is absent from a selection operator, that selection operator is termed *vacuous*. If an integer occurs more than once in a selection operator, that selection operator is *reflexive*. A selection operator that is neither vacuous nor reflexive is a *permutation*.

Let $\langle k_1, \ldots, k_m \rangle P^m$ occur in sentence W_1 governed by $g_q \cdots g_1$, where $max(k_i)_{1 \leq i \leq m} = n \leq q$. Let $\overline{\langle l_1, \ldots, l_m \rangle P^m}$ occur in sentence W_2 governed by $h_r \cdots h_1$, where $max(l_i)_{1 \leq i \leq m} = p \leq r$. If W_1 and W_2 can be transformed to W_1' and W_2', respectively, containing occurrences of $\langle k_1', \ldots, k_m' \rangle P^m$ and $\overline{\langle k_1', \ldots, k_m' \rangle P^m}$, respectively, such that $W_1 \subseteq W_1'$ and $W_2 \subseteq W_2'$, then Rule $\mathbf{R}[\langle k_1', \ldots, k_m' \rangle P^m]$ can be applied. A necessary and sufficient condition for the existence of such transformations is considered next.

6.1 **The case of unary governors** To gain an understanding of the problem, it is useful to initially make the simplifying assumption that all governing subexpressions are unary. This assumption is frequently satisfied in natural language, and so also is of special interest for surface reasoning, which will be considered in the following sections. Under this assumption, if P^m occurs in a sentence W governed by $g_q \cdots g_1$, it follows that $q = m$; and if $\langle k_1, \ldots, k_m \rangle P^m$, occurs in sentence W governed by $g_q \cdots g_1$, where $max(k_i)_{1 \le i \le m} = n$, then $q = n$.

A governance $g_n \cdots g_1$ induces a *dependence ordering* $\prec$ on the g_i defined as follows.

$$g_i \prec g_j :\Leftrightarrow (i < j) \wedge (g_i \text{ is } -) \wedge (g_j \text{ is } +)$$

This corresponds in predicate calculus to a universal quantifier controlling an existential quantifier.

The following indicates the significance of the dependence ordering. Suppose $\langle k_1, \ldots, k_m \rangle P^m$ occurs in sentence W governed by $g_m \cdots g_1$, where $\langle k_1, \ldots, k_m \rangle$ is a permutation. Then P^m is governed in W by $g_{k_m} \cdots g_{k_1}$ providing the dependence ordering induced by $g_{k_m} \cdots g_{k_1}$ is identical to that induced by $g_m \cdots g_1$. Otherwise, the governance of P^m in W is undefined.

Let $\langle k_1, \ldots, k_m \rangle P^m$ occur in sentence W_1 governed by $g_n \cdots g_1$, where $max(k_i)_{1 \le i \le m} = n$, and let $\overline{\langle l_1, \ldots, l_m \rangle P^m}$ occur in sentence W_2 governed by $h_p \cdots h_1$, where $max(l_i)_{1 \le i \le m} = p$. Let a graph $\mathcal{G}'$ be defined with nodes $g_1, \ldots, g_n, h_1, \ldots, h_p$ and arcs as follows. If $g_i \prec g_j$ then (g_j, g_i) is an arc. If $h_i \prec h_j$ then (h_j, h_i) is an arc. There are no other arcs. Let $\sim$ be the least equivalence relation containing $\{(g_{k_i}, h_{l_i}) : 1 \le i \le m\}$. Now define $\mathcal{G}$, the *dependence graph* of W_1, W_2 with respect to P^m, to be $\mathcal{G}'$ reduced by $\sim$.

The nodes of $\mathcal{G}$ are equivalence classes, denoted $[g]$, consisting of corresponding governors of P^m in W_1 and W_2. The arcs of $\mathcal{G}$ give the combined dependence ordering. A node containing only negative governors will be called a *negative* node; one containing a positive governor, a *positive* node. A positive node containing only one governor will be called *strictly positive.* A positive node containing more than one positive governor will be called *inconsistent.* A graph containing no inconsistent node is *consistent.*

Suppose $\mathcal{G}$ contains no cycles (i.e., directed closed paths). In this case, there exists a linear ordering $\mathcal{C}$ that extends $\mathcal{G}$. Let $[g_i] \prec_{\mathcal{G}} [g_j]$ iff $[g_j]$ covers $[g_i]$ in $\mathcal{G}$. Let

$[g_i] \prec_{\mathcal{C}} [g_j]$ iff $[g_i]$ precedes $[g_j]$ in $\mathcal{C}$. In general, there are many linear extensions. An important property of these linear extensions is the following.

$$cost(\mathcal{C},\mathcal{G}) := card(\{(+X_j,-X_i) : [-X_i] \prec_{\mathcal{C}} [+X_j] \wedge [-X_i] \not\prec_{\mathcal{G}} [+X_j]\})$$

A smaller cost means that fewer new dependences have been introduced by the linear extension. An *optimal* linear extension is one with minimum cost. The situation is similar to that encountered in scheduling 'jobs' which are constrained by a partial order (see Rival [6]). The complexity of constructing an optimal linear extension will not be considered here, but it is obviously an interesting question.

Let $W_1\boldsymbol{i} = {}_\wedge T(X_n \cup \cdots \cup {}_\wedge T(X_1 \cup \overline{\langle k_1,\ldots,k_m\rangle P^m})\cdots)$ and $W_2\boldsymbol{j} = {}_\wedge T(Z_p \cup \cdots \cup {}_\wedge T(Z_1 \cup \langle l_1,\ldots,l_m\rangle P^m)\cdots)$ be the principal Skolem constituents of W_1 and W_2, respectively. Suppose that $\mathcal{G}$ is consistent and acyclic. In this case, there is an optimal linear extension $\mathcal{C} = C_q \prec C_{q-1} \prec \cdots \prec C_1$, where $1 \le q < n+p$. Let $\mathcal{G}_{Sk}$ and $\mathcal{C}_{Sk}$ be the corresponding dependence graph and linear extension obtained by replacing each positive governor $+X$ by its corresponding Skolem governor $-\overline{f}$. Define $W_1' := {}_\wedge T(U_q \cup \cdots \cup {}_\wedge T(U_1 \cup \overline{\langle k_1',\ldots,k_m'\rangle P^m})\cdots)$ and $W_2' := {}_\wedge T(V_q \cup \cdots \cup {}_\wedge T(V_1 \cup \langle k_1',\ldots,k_m'\rangle P^m)\cdots)$, where $U_i = \bigcup\{\overline{X} : -X \in C_i\}$, $V_i = \bigcup\{\overline{Z} : -Z \in C_i\}$, and $\langle k_1',\ldots,k_m'\rangle$ is derived from $\mathcal{C}_{Sk}$ as follows: if $-X_{k_i} \in C_j$ (equivalently $-Z_{l_i} \in C_j$) then $k_i' = j$. Notice that either U_i or V_i may be empty. Notice also that the Skolem constants may be given modified selection operators which reflect their positions in W_1' and W_2'. Now the resolvent of $W_1\boldsymbol{i}$ and $W_2\boldsymbol{j}$ with respect to P^m is defined to be $\mathbf{Res}[\langle k_1',\ldots,k_m'\rangle P^m](W_1',W_2')$. An optimal linear extension corresponds to a most general unifier.

It is important to note that if $\mathcal{G}$ consists of disjoint acyclic subgraphs, then the resolvent of $W_1\boldsymbol{i}$ and $W_2\boldsymbol{j}$ with respect to P^m can be written as a disjunction of sentences, one determined by each subgraph.

Since an arc of $\mathcal{G}$ represents a (control) dependence, the corresponding arc of $\mathcal{G}_{Sk}$ represents a *functional* dependence between a Skolem constant and one of its arguments. It is clear that if $\mathcal{G}_{Sk}$ contains a cycle, some Skolem constant is functionally dependent on itself. This is exactly the condition that produces an occur-check during unification. Thus W_1, W_2 have a resolvent iff $\mathcal{G}$ is consistent and acyclic.

It is possible to extend Rule $\mathbf{R}[P^m]$ as binary resolution has just been extended. Let W_1 and W_2 be as before and let $\mathcal{G}$, the dependence graph of W_1 and W_2 with respect to P^m, be consistent and acyclic and consist of the disjoint components $\mathcal{G}_1, \dots, \mathcal{G}_s$ $(s \geq 1)$. Let $\mathcal{C}_1, \dots, \mathcal{C}_s$ be the corresponding optimal linear extensions. Then

$$\mathbf{R}[P^m](W_1, W_2) := \mathbf{R}[P^m](\mathcal{C}_1) \cup \dots \cup \mathbf{R}[P^m](\mathcal{C}_s), \text{ where}$$

$$\mathbf{R}[P^m](C\mathcal{C}) := \begin{cases} \wedge X(\mathbf{R}[P^m](\mathcal{C})) & \text{if } C \text{ is } - \\ & \text{and } X = \bigcap\{X_j : -X_j \in C\} \\ X(\mathbf{R}[P^m](\mathcal{C})) & \text{if } C \text{ is strictly } + \\ & \text{and } +X \in C \\ X(V \cap (Z \cup \mathbf{R}[P^m](\mathcal{C}))) & \text{if } C \text{ is } + \\ & \text{and } +X \in C,\ Z = \bigcup\{\overline{Z_j} : -Z_j \in C\}, \\ & V \text{ is the scope of } X \text{ in } W_1 \text{ or } W_2 \end{cases}$$

$$\mathbf{R}[P^m](\emptyset) = \overline{T}$$

The scope of a positive governor in W is defined as follows. Let V be the subexpression of W containing the distinguished occurrence of P^m. Then V is the scope of X if XV is a subexpression of W or VX is a subexpression of W, and V is the scope of $X \cap U$ if $X(U \cap V)$ is a subexpression of W or $(U \cap V)X$ is a subexpression of W.

THEOREM 2 *Let $\langle k_1, \dots, k_m \rangle P^m$ occur in sentence W_1 governed by unary $g_n \cdots g_1$, where $max(k_i)_{1 \leq i \leq m} = n$. Let $\overline{\langle l_1, \dots, l_m \rangle P^m}$ occur in sentence W_2 governed by unary $h_p \cdots h_1$, where $max(l_i)_{1 \leq i \leq m} = p$. Let $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \dots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$ and $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \dots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$ be the constituents of Skolemization of W_1, W_2 respectively with respect to Y. Let $\mathcal{G}$ be the dependence graph of W_1, W_2. Then $\mathbf{R}[P^m](W_1, W_2)$ and $\mathbf{Res}[P^m](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$ exist iff $\mathcal{G}$ is consistent and acyclic. Moreover, W_1, W_2, and $\mathbf{R}[P^m](W_1, W_2)$ are Skolem equivalent to $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \dots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$, $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \dots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$, and $\mathbf{Res}[P^m](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$*

proof: The proof follows that of Theorem 1.

6.2 **The general case** The general case can be treated as an extension of the unary case. Let $\langle k_1,\ldots,k_m\rangle P^m$ occur in sentence W_1 with unrestricted governors $g_q\cdots g_1$, where $max(k_i)_{1\le i\le m}=n\le q$. Similarly, let $\overline{\langle l_1,\ldots,l_m\rangle P^m}$ occur in sentence W_2 with unrestricted governors $h_r\cdots h_1$, where $max(l_i)_{1\le i\le m}=p\le r$. Temporarily ignoring all but the unary governors, form the dependence graph $\mathcal{G}$ and without distinguishing disjoint subgraphs, form an optimal linear extension $\mathcal{C}$ as before.

Suppose that the k-ary expression $(k>1)$ $\langle c_1,\ldots,c_h\rangle R^h$ is a governor of P^m in W_1. Suppose further that arguments $1,2,\ldots,h$ of R^h are filled (partly or completely) by unary governors occurring in nodes $A_{v_1},A_{v_2},\ldots,A_{v_h}$, respectively, of $\mathcal{G}$, where $min(v_j)_{1\le j\le h}=u$. In this event, $\langle v_1+1-u,v_2+1-u,\ldots,v_h+1-u\rangle R^h$ is placed in node A_u.

Suppose that an arbitrary k-ary expression $(k>1)$ X^k is a governor of P^m in W_1, and that arguments $1,2,\ldots,h$ of X^k are filled (partly or completely) by unary governors occurring in nodes $A_{v_1},A_{v_2},\ldots,A_{v_h}$, respectively, of $\mathcal{G}$, where $min(v_j)_{1\le j\le h}=u$. To formalize this situation, generalize selection operators to apply to arbitrary expressions, and provide the following rules for the elimination of these generalized selection operators.

1. $\langle k_1,\ldots,k_m\rangle\langle l_1,\ldots,l_n\rangle X^n := \langle k_{l_1},\ldots,k_{l_n}\rangle X^n$, where $m=max(l_i)_{1\le i\le n}$
2. $\langle k_1,\ldots,k_m\rangle\overline{X^m} := \overline{\langle k_1,\ldots,k_m\rangle X^m}$
3. $\langle k_1,\ldots,k_m\rangle(X^n\cap Y^l) := (\langle k_1,\ldots,k_n\rangle X^n\cap\langle k_1,\ldots,k_l\rangle Y^l)$, where $m=max(l,n)$
4. $\langle k_1,\ldots,k_m\rangle(X^1Y^{m+1}) := X^1(\langle 1,k_1+1,\ldots,k_m+1\rangle Y^{m+1})$

With these generalized selection operators, X^k can be treated in the same manner as $\langle c_1,\ldots,c_h\rangle R^h$. Again the resulting expression is placed in node A_u.

The same is done for each governor of arity greater than 1 in both W_1 and W_2. Now $\mathcal{C}$ is decomposed into disjoint subchains just so long as distinct subchains do not separate the arguments of any governor. Finally, Rule $\mathbf{R}[P^m](W_1, W_2)$ in a slightly more general form can be applied to the subchains of $\mathcal{C}$.

$$\mathbf{R}[P^m](W_1, W_2) := \mathbf{R}[P^m](\mathcal{C}_1) \cup \cdots \cup \mathbf{R}[P^m](\mathcal{C}_s), \text{ where}$$

$$\mathbf{R}[P^m](C\mathcal{C}) := \begin{cases} {}_{\wedge}T(X \cup \mathbf{R}[P^m](\mathcal{C})) & \text{if } C \text{ is } - \\ & \text{and } X = \bigcup\{\overline{X_j} : -X_j \in C\} \\ T(X \cap \mathbf{R}[P^m](\mathcal{C})) & \text{if } C \text{ is strictly } + \\ & \text{and } +X \in C \\ T(X \cap V \cap (Z \cup \mathbf{R}[P^m](\mathcal{C}))) & \text{if } C \text{ is } + \\ & \text{and } +X \in C,\ Z = \bigcup\{\overline{Z_j} : -Z_j \in C\}, \\ & V \text{ is the scope of } X \text{ in } W_1 \text{ or } W_2 \end{cases}$$

$$\mathbf{R}[P^m](\emptyset) = \overline{T}$$

THEOREM 3 *Let* $\langle k_1, \ldots, k_m\rangle P^m$ *occur in sentence* W_1 *governed by unary* $g_n \cdots g_1$, *where* $max(k_i)_{1\leq i\leq m} = n$. *Let* $\overline{\langle l_1, \ldots, l_m\rangle P^m}$ *occur in sentence* W_2 *governed by unary* $h_p \cdots h_1$, *where* $max(l_i)_{1\leq i\leq m} = p$. *Let* $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$ *and* $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$ *be the constituents of Skolemization of* W_1, W_2 *respectively with respect to* Y. *Let* $\mathcal{G}$ *be the dependence graph of* W_1, W_2. *Then* $\mathbf{R}[P^m](W_1, W_2)$ *and* $\mathbf{Res}[P^m](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$ *exist iff* $\mathcal{G}$ *is consistent and acyclic. Moreover,* W_1, W_2, *and* $\mathbf{R}[P^m](W_1, W_2)$ *are Skolem equivalent to* $W_1\boldsymbol{i}'_k, W_1\boldsymbol{i}_k, \ldots, W_1\boldsymbol{i}'_1, W_1\boldsymbol{i}_1$, $W_2\boldsymbol{j}'_l, W_2\boldsymbol{j}_l, \ldots, W_2\boldsymbol{j}'_1, W_2\boldsymbol{j}_1$, *and* $\mathbf{Res}[P^m](W_1\boldsymbol{i}_1, W_2\boldsymbol{j}_1)$

proof: The proof follows that of Theorem 1.

7 Syllogistic cases of Rule $\mathbf{R}[Y]$ Certain cases of Rule $\mathbf{R}[Y]$ are of special interest for surface reasoning. In these cases, application of Rule $\mathbf{R}[Y]$ is simplified so that it resembles simple syllogistic reasoning. Its arguments need not be normalized.

7.1 Rule AII Let Y occur in sentences W_1 and W_2 with opposite polarities. Let Y ($\overline{Y}$) be governed in W_1 by $-X_m \cdots - X_1$. Let $\overline{Y}$ (Y) be governed in W_2 by $+Z_l \cdots + Z_1$, where $m \leq l$. To motivate this rule assume temporarily that W_1 and W_2 are normalized with respect to Y. Under these assumptions, Rule $\mathbf{R}[Y]$ specializes to

$$\mathbf{R}[Y](\wedge T(\overline{X_i} \cup V_{1,i}), T(Z_i \cap V_{2,i})) = T(Z_i \cap V_{2,i} \cap (\overline{X_i} \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

Hence

$$\mathbf{R}[Y](\wedge T(\overline{X_i} \cup V_{1,i}), T(Z_i \cap V_{2,i})) \equiv T(Z_i \cap V_{2,i} \cap \overline{X_i}) \cup T(Z_i \cap V_{2,i} \cap \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

Also under the current assumptions, by the Monotonicity Rule,

$$\mathbf{R}[Y](V_{1,i}, V_{2,i}) \subseteq T(Z_{i+1} \cap V_{2,i+1}) = V_{2,i}$$

Hence, again by the Monotonicity Rule and the identity $(X \subseteq Y) \equiv ((X \cap Y) \equiv X)$,

$$\mathbf{R}[Y](\wedge T(\overline{X_i} \cup V_{1,i}), T(Z_i \cap V_{2,i})) \equiv T(Z_i \cap V_{2,i} \cap \overline{X_i}) \cup T(Z_i \cap \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

Thus under the current assumptions, Rule $\mathbf{R}[Y]$ becomes

$$\mathbf{R}[Y](W_1, W_2) = W_2^{(m)} \cup \cdots \cup W_2^{(1)}$$

where for $1 \leq i \leq m$, $W_2^{(i)}$ is obtained from W_2 by replacing $V_{2,i}$ with $V_{2,i} \cap \overline{X_i}$. But this result is not dependent upon W_1 and W_2 being normalized. So Rule AII can be given as follows.

> Rule AII: Let Y occur in sentences W_1 and W_2 with opposite polarities. Let Y ($\overline{Y}$) be governed in W_1 by $-X_m \cdots - X_1$. Let $\overline{Y}$ (Y) be governed in W_2 by $+Z_l \cdots + Z_1$, where $m \leq l$. Then infer $W_2^{(m)} \cup \cdots \cup W_2^{(1)}$, where $W_2^{(i)}$ is obtained from W_2 by replacing the subexpression V in the scope of Z_i with $V \cap \overline{X_i}$.

As a first example of Rule AII, consider the syllogism *Baroco*: all X are Y, and some Z are not Y, so some Z are not X. That is, $W_1 = \wedge XY$ and $W_2 = Z\overline{Y}$. Thus Y is governed in W_1 by $-X$ and $\overline{Y}$ is governed in W_2 by $+Z$. Therefore Rule AII yields $Z(\overline{Y} \cap \overline{X})$.

As a second more general example, consider $W_1 = \wedge X_3 \wedge X_2(U^2 \cup \wedge X_1 Y)$ and $W_2 = (Z_2 Z_1 \overline{Y})Z_3$. Here Y is governed in W_1 by $-X_3 - (X_2 \cap \overline{U^2}) - X_1$ and $\overline{Y}$ is governed in W_2 by $+Z_3 + Z_2 + Z_1$. Therefore by Rule AII one infers

$$(Z_2 Z_1 \overline{Y} \cap \overline{X_3})Z_3 \cup (Z_2(Z_1 \overline{Y} \cap \overline{(X_2 \cap \overline{U^2})}))Z_3 \cup (Z_2 Z_1(\overline{Y} \cap \overline{X_1}))Z_3$$

7.2 **Rule AAA** Let Y occur in sentences W_1 and W_2 with opposite polarities. Let Y be governed in W_1 by $-X_m \cdots - X_1$. Let $\overline{Y}$ be governed in W_2 by $-Z_l \cdots - Z_1$. Again assume temporarily that W_1 and W_2 are normalized with respect to Y. Under these assumptions, Rule $\mathbf{R}[Y]$ becomes identical to $\mathbf{Res}[Y]$:

$$\mathbf{R}[Y](\wedge T(\overline{X_i} \cup V_{1,i}), \wedge T(\overline{Z_i} \cup V_{2,i})) = \wedge T(\overline{X_i} \cup \overline{Z_i} \cup \mathbf{R}[Y](V_{1,i}, V_{2,i})))$$

This is put in a more perspicuous form as Rule AAA.

> Rule AAA: Let Y occur in sentences W_1 and W_2 with opposite polarities. Let $Y \cup V_1$ be governed in W_1 by $-X_m \cdots - X_1$. Let $\overline{Y} \cup V_2$ be governed in W_2 by $-Z_l \cdots - Z_1$. Let $n = max(l, m)$ and
>
> $$U_i = \begin{cases} X_i & \text{if } l < i \leq n \\ Z_i & \text{if } m < i \leq n \\ X_i \cup Z_i & \text{otherwise} \end{cases}$$
>
> Then infer $\wedge T(\overline{U_n} \cup \cdots \cup \wedge T(\overline{U_1} \cup V_1 \cup V_2) \cdots)$.

If $V_2 = \overline{TT}$, then Rule AAA is just the Cancellation Rule. If $V_1 = V_2 = \overline{TT}$, then Rule AAA is still further simplified. In this case the inference may be stated: $\wedge T(\overline{U_n} \cup \cdots \cup \wedge T(\overline{X_1} \cup \overline{Z_1}) \cdots)$.

When the governors are unary, the inferred sentence(s) retain the form of the premises. Thus in general one infers $\wedge U_k \cdots \wedge U_1(V_1 \cup V_2) \cup \overline{TU_n} \cup \cdots \cup \overline{TU_{k+1}}$, where $k \leq n$ is the arity of $V_1 \cup V_2$.

For illustration of Rule AAA, again consider first a syllogism (*Camestres*): all X are Y, and all Z are not Y, so all Z are not X. That is, $W_1 = {\wedge}XY$ and $W_2 = {\wedge}Z\overline{Y}$. Thus Y is governed in W_1 by $-X$ and $\overline{Y}$ is governed in W_2 by $-Z$. Therefore by Rule AAA one infers $\overline{X}$ is governed in the resolvent by $-Z$. That is, one infers ${\wedge}Z\overline{X}$. A more general example is Schubert's Steamroller (see [3]).

8 **Surface reasoning** The impetus for investigating a generalization of the Cancellation Rule was to better understand surface reasoning, that is, reasoning conducted in (surface level) natural language. An underlying assumption is that surface reasoning characterizes much of human reasoning.

Previous papers ([3], [4]) emphasized the Monotonicity Rule and its corollaries. The Cancellation Rule is one of these corollaries. It is similar to unit resolution. However, it is applied to expressions at the surface level rather than clausal form expressions. Hence the Cancellation Rule has a direct and intuitive rendition in surface English.

Similarly, the Generalized Cancellation Rule developed in this paper is equivalent to binary resolution, but applies directly to surface expressions. Skolemization, prenexing, and conjunctive normal form play no part. Like the Cancellation Rule, this generalization, particularly in its syllogistic forms, has a direct and intuitive rendition in surface English. This will be demonstrated by several examples.

Consider the sentences **some man is unkind to every donkey**, and **every farmer is kind to some animal**. The *de re* reading of the first and the *de dicto* reading of the second are rendered in $\mathcal{L}_N$ as $M{\wedge}D\overline{K}$ and ${\wedge}FAK$, respectively. Rule $\mathbf{R}[K]$ yields the result $M({\wedge}D\overline{K} \cap (\overline{F} \cup A(K \cap \overline{D}))$. The direct rendition of the latter in English is **some man is unkind to every donkey and either is not a farmer or is kind to some animal which is not a donkey**. Although this is a rather complex inference, it can be understood and accepted by any competent English speaker.

The simpler sentences **every farmer is unkind to every donkey**, and **some man is kind to some animal**, permit an inference by use of Rule AII. These sentences are rendered in $\mathcal{L}_N$ as ${\wedge}F{\wedge}D\overline{K}$ and MAK, respectively. Rule AII yields the result $M(AK \cap \overline{F}) \cup MA(K \cap \overline{D})$. This result is rendered directly in English as **either some man who is not a farmer is kind to some animal, or some man is kind to some animal which is not a donkey**. Again, it is apparent that this inference can be understood and accepted by any competent English speaker.

Rule AAA can be illustrated with the sentences **every farmer is unkind to every donkey**, and **every gentleman is kind to every animal**, which are rendered in $\mathcal{L}_N$ as $\wedge F \wedge D\overline{K}$ and $\wedge G \wedge AK$, respectively. Rule AAA yields $\wedge F\overline{G} \cup \wedge D\overline{A}$, that is, **either every farmer is not a gentleman or every donkey is not an animal.**

A sequence of inferences constitutes a proof or a line of reasoning. If the proof is indirect, then in addition to Rule $\mathbf{R}[Y]$ and its variations, a criterion for contradiction needs to be given. In $\mathcal{L}_N$, a contradiction is present when a sentence of the form $X \cap \overline{X}$ is deduced. This will require, in addition to the Generalized Cancellation Rule, some simple Boolean identities and definitions.

$$(X \cap Z)Y \equiv X(Z \cap Y) \qquad X(Z \cap Y \cap \overline{Y}) \equiv (TY \cap \overline{TY})$$

$$(X \cap (Z \cup Y)) \equiv ((X \cap Z) \cup (X \cap Y)) \qquad \wedge X\overline{Y} \equiv \overline{XY}$$

The next example, taken from Quine [5], illustrates a sequence of inferences which constitute a proof.

> 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 set of sentences:

$$\{\wedge AC, \wedge(W \cap C)P, (AM)\overline{P}, \wedge(T\check{M})W, \wedge(PM)P\}$$

To illustrate the Generalized Cancellation Rule, a proof by contradiction is given with all details explicit. The sentences will be denoted $X_1, X_2, \ldots, X_5$.

1. $\mathbf{R}[A](X_1, X_3)$:

   (a) $\overline{A}$ is governed in X_1 by $-\overline{C}$

(b) A is governed in X_3 by $+\overline{P}+M$

(c) the linear extension of $\mathcal{G}$ is $[+\overline{P}][+M,-\overline{C}]$

(d) the result is $X_6 = \overline{P}T(M \cap A \cap C)$

2. $\mathbf{R}[M](X_4, X_6)$:

(a) $\overline{M}$ is governed in X_4 by $-T-\overline{W}$

(b) M is governed in X_6 by $+\overline{P}+(A \cap C)$

(c) the linear extension of $\mathcal{G}$ is $[+\overline{P},-T][+(A \cap C),-\overline{W}]$

(d) the result is $\overline{P}(T(M \cap A \cap C) \cap (\overline{T} \cup (A \cap C)(M \cap W)))$, which is equivalent to $X_7 = \overline{P}(T(M \cap A \cap C) \cap (A \cap C)(M \cap W))$

3. $\mathbf{R}[(W \cap C)](X_2, X_7)$:

(a) $\overline{W \cap C}$ is governed in X_2 by $-\overline{P}$

(b) $W \cap C$ is governed in X_7 by $+\overline{P}+(T(M \cap A \cap C) \cap M \cap A)$

(c) the linear extension of $\mathcal{G}$ is $[+\overline{P}][+(T(M \cap A \cap C) \cap M \cap A),-\overline{P}]$

(d) the result is $X_8 = \overline{P}T(T(M \cap A \cap C) \cap M \cap A \cap W \cap C \cap P)$

4. $\mathbf{R}[M](X_5, X_8)$:

(a) $\overline{M}$ is governed in X_5 by $-\overline{P}-P$

(b) M is governed in X_8 by $+\overline{P}+(T(M \cap A \cap C) \cap A \cap W \cap C \cap P)$

(c) the linear extension of $\mathcal{G}$ is $[+\overline{P},-\overline{P}][+(T(M \cap A \cap C) \cap A \cap W \cap C \cap P),-P]$

(d) the result is $X_9 = \overline{P}(T(T(M \cap A \cap C) \cap M \cap A \cap W \cap C \cap P) \cap (P \cup T(T(M \cap A \cap C) \cap A \cap W \cap C \cap P \cap M \cap \overline{P})))$, which is equivalent to $TP \cap \overline{TP}$

Rule $\mathbf{R}[Y]$ can be optimized by observing that a node $[+X,-T]$ can be treated the same as a strictly positive node. This treatment would eliminate the subexpression $T(M \cap A \cap C)$ in the second and subsequent steps.

While the Generalized Cancellation Rule suffices to yield an indirect proof, surface reasoning is generally much simpler. This can be illustrated by using Rule AII to yield a proof.

1. X_1 and X_3 yield $(A(M \cap C))\overline{P}$
2. X_4 and the previous result yield $(A(M \cap W \cap C))\overline{P} \cup (A(M \cap C))(\overline{P} \cap \overline{T})$, i.e., $(A(M \cap W \cap C))\overline{P}$
3. X_2 and the previous result yield $(A(M \cap W \cap C \cap P))\overline{P}$
4. X_5 and the previous result yield $(A(M \cap W \cap C \cap P \cap \overline{P}))\overline{P} \cup (A(M \cap W \cap C \cap P))(\overline{P} \cap P)$, which is equivalent to $TP \cap \overline{TP}$

These proofs are based on use of a single rule of inference. It seems unlikely that surface reasoning is restricted to a single rule of inference. Using a combination of the Monotonicity Rule and Rule AII, an even simpler proof can be obtained.

1. Using the Monotonicity Rule, X_1 and X_3 yield $(CM)\overline{P}$
2. Using Rule AII, X_4 and the previous result yields $(C(M \cap W))\overline{P} \cup (CM)(\overline{P} \cap \overline{T})$, i.e., $((W \cap C)M)\overline{P}$
3. Using the Monotonicity Rule, X_2 and the previous result yield $(PM)\overline{P}$
4. Using the Monotonicity Rule, X_5 and the previous result yields $P\overline{P}$, a contradiction

It is well-known that binary resolution is not complete as a refutation procedure. For example, the sentences $\wedge Ta(P \cup \breve{P}), \wedge Ta(\overline{P} \cup \overline{\breve{P}})$ are inconsistent, but a contradiction cannot be obtained by binary resolution alone. Factoring is required. But while binary resolution in the form of the Generalized Cancellation Rule seems natural to surface reasoning, factoring seems quite unnatural.

It has been shown that binary resolution alone is complete for Horn sentences (Henschen and Wos [2]). A test for the Horn property in $\mathcal{L}_N$ is straightforward. Define $w(X)$ for an expression X as follows.

1. $w(Y \cap Z) = max(w(Y), w(Z))$
2. $w(\overline{Y \cap Z}) = w(\overline{Y}) + w(\overline{Z})$
3. $w(YZ) = w(Y \cap Z)$
4. $w(\overline{YZ}) = w(\overline{Y \cap Z})$
5. $w(P) = 1$
6. $w(\overline{P}) = 0$

A sentence X is Horn iff $w(X) \leq 1$. In some cases, a uniform substitution of $\overline{R_i^n}$ for R_j^n in X makes X Horn.

It is clear that the Horn property is not necessary for binary resolution to be complete. However it appears that a precise characterization of sentences that yield to binary resolution alone is an open problem.

References

[1] Andrews, Peter B. *An Introduction to Mathematical Logic and Type Theory*, Academic Press, Orlando, 1986.

[2] Henschen, L. and L. Wos "Unit Refutations and Horn Sets," *Journal of the Association for Computing Machinery* 21 (1974), 590-605.

[3] Purdy, William C. "A Logic for Natural Language," *Notre Dame Journal of Formal Logic* 32 (1991), 409-425.

[4] Purdy, William C. "Surface Reasoning," *Notre Dame Journal of Formal Logic* 33 (1992), 13-36.

[5] Quine, W. V. *Methods of Logic*, Fourth Edition, Harvard University Press, Cambridge, 1976.

[6] Rival, Ivan "Linear Extensions of Finite Ordered Sets," in Pouzet and Richard (eds.) *Orders: Description and Roles*, Annals of Discrete Mathematics (23), North-Holland, Amsterdam, 1984.