# SOME INTERPOLATION THEOREMS FOR FIRST-ORDER FORMULAS IN WHICH ALL DISJUNCTIONS ARE BINARY (1)

M. R. KROM

The *atomic formulas* of a first-order predicate calculus are the predicate letters with attached individual symbols, and the *signed atomic formulas* are the atomic formulas and the negations of atomic formulas. A first-order formula is in *prenex conjunctive normal form* (cf. page 1 of [1]) in case it consists of a string of quantifiers (called a prefix) followed by a quantifier free part (called a matrix) which is a conjunction of disjunctions of signed atomic formulas. We establish properties of the formulas of any first-order predicate calculus without identity and without function symbols which are in prenex conjunctive normal form and in which the disjunctions are binary (have just two terms). The condition that the disjunctions be binary is not so limiting that it restricts us to trivial classes of formulas; in fact there is an apparently still unsolved decision problem for these classes of formulas( cf. the discussion following Theorem 3 of [5]). However, as we will show, there are no results analogous to ours for classes of formulas defined by restricting to ternary disjunctions.

Most of the properties that we establish are expressed in the form of interpolation theorems. We begin with one such theorem for formulas in a statement calculus which completely characterizes when a formula is logically equivalent to a conjunction of *binary* disjunctions of signed statement letters. We have three interpolation theorems for formulas in first-order languages and one of them is a modification of Craig's Theorem (Theorem 5 page 267 of [2]) which yields a modified form of Beth's Theorem on definability (Theorem 5.2.1 page 118 of [6]).

(1) The research reported here was supported by National Science Foundation Grant number GP-6358.

For any two formulas $\Gamma$, $\Delta$ of a statement calculus or of a first-order predicate calculus we use $\Gamma \vdash \Delta$ to indicate that $\Delta$ is a logical consequence of $\Gamma$. We use $\vee$, $\wedge$, and $\neg$, respectively, for the logical connectives, disjunction, conjunction, and negation.

## § 1. Binary Disjunctions in Statement Calculus

We will use $\alpha$, $\beta$, $\alpha_1$, $\beta_1$, $\alpha_2$, $\beta_2$ ... to denote arbitrary statement letters and negated statement letters of a statement calculus. For any $\alpha$ we let $\alpha'$ be the negation of $\alpha$ if $\alpha$ is a statement letter and the statement letter occurring in $\alpha$ if $\alpha$ is a negated statement letter. A *binary disjunction* is an expression of the form $\alpha \vee \beta$. For any $\alpha$, $\beta$ we let $\alpha \vee \beta$ be either one of the two logically equivalent expressions $\alpha \vee \beta$ and $\beta \vee \alpha$. A *chain* from $\alpha$ to $\beta$ is a set of the form $\{\alpha_1 \vee \alpha_2, \alpha'_2 \vee \alpha_3, \alpha'_3 \vee \alpha_4, \ldots \alpha'_{n-1} \vee \alpha_n\}$ with $\alpha_1 = \alpha$ and $\alpha_n = \beta$. The following lemma is a restatement of Corollary 2.2 of [4].

*Lemma 1.* For any set S of binary disjunctions, S is inconsistent if and only if there is a statement letter $\alpha$ and there are two chains formed with members of S such that one chain is from $\alpha$ to $\alpha$ and one chain is from $\alpha'$ to $\alpha'$.

We will briefly indicate a proof of this lemma and refer the reader to [4] for complete details. Any set of binary disjunctions forming a chain from $\alpha$ to $\beta$ obviously has the disjunction $\alpha \vee \beta$ as a logical consequence. Thus any set S from which chains can be formed, one from $\alpha$ to $\alpha$ and one from $\alpha'$ to $\alpha'$, has both $\alpha \vee \alpha$ and $\alpha' \vee \alpha'$ as logical consequences and so it is inconsistent. Conversely, if S is a set of binary disjunctions for which there are no two chains as indicated in the lemma then we can show by induction that S can be extended to a set S′ of binary disjunctions such that for any predicate letter $\alpha$ occurring in members of S′ there is, with members of S′, a chain from $\alpha$ to $\alpha$ or a chain from $\alpha'$ to $\alpha'$ but not both. This determines an assignment of truth values which shows that S is consistent.

We say that a binary disjunction is a *singleton* in case its two disjuncts are identical.

*Lemma 2.* For any inconsistent set S of binary disjunctions, the two chains required by Lemma 1 may be formed so that neither one of them has more than one singleton occurring in it.

Proof. Let c be a chain from $\alpha$ to $\alpha$ in which at least two singletons occur. We may assume that $\alpha \vee \alpha$ is not a singleton that occurs in c otherwise we could replace c with the chain $\{\alpha \vee \alpha\}$. Let $\beta_1 \vee \beta_1$ be an endmost singleton occurring in c, that is, $\beta_1 \vee \beta_1$ is a singleton occurring in c such that there is a subchain c′ of c from $\alpha$ to $\beta'_1$ in which no singleton occurs. Then c′ together with $\beta_1 \vee \beta_1$ forms a chain c* from $\alpha$ to $\beta_1$ in which one singleton occurs, and c′ together with c* forms a chain from $\alpha$ to $\alpha$ in which no more than one singleton occurs. Hence c can be replaced by a chain in which no more than one singleton occurs.

*Theorem 1.* If $\Gamma$ is a conjunction of binary disjunctions of signed statement letters and $\Delta$ is a disjunction of signed statement letters such that $\Gamma \vdash \Delta$ then there is a subdisjunction $\bar{\Delta}$ of $\Delta$ of no more than two terms such that $\Gamma \vdash \bar{\Delta}$ and $\bar{\Delta} \vdash \Delta$.

Proof. Let $\Gamma$ and $\Delta$ be formulas as described in the hypothesis of the lemma. Let $S_1$ be the set of binary disjunctions that occur in $\Gamma$ and let $S_2$ be the set of singletons of the form $\beta' \vee \beta'$ for each disjunct $\beta$ of $\Delta$. Since $S_2$ is logically equivalent to the negation of $\Delta$, $S_1 \cup S_2$ is inconsistent. By Lemma 1, there is a statement letter, say $\alpha$, such that there are two chains, one from $\alpha$ to $\alpha$ and one from $\alpha'$ to $\alpha'$, formed with the elements of $S_1 \cup S_2$. By Lemma 2, no more than two elements of $S_2$ need occur in these chains. It follows that there is a subset $\bar{S}_2$ of $S_2$ with no more than two elements such that $S_1 \cup \bar{S}_2$ is inconsistent. Let $\bar{\Delta}$ be a subdisjunction of $\Delta$ of no more than two disjuncts, a disjunct $\beta$ for each singleton $\beta' \vee \beta'$ of $\bar{S}_2$. Then $\Gamma \vdash \bar{\Delta}$ because $\Gamma$ is inconsistent with $\neg \bar{\Delta}$ and $\bar{\Delta} \vdash \Delta$ because $\bar{\Delta}$ is a subdisjunction of $\Delta$.

There is no result corresponding to Theorem 1 which applies to ternary disjunctions instead of binary disjunctions. In fact, for any positive integer n there is a conjunction $\Gamma$ of ternary disjunctions of signed statement letters and a disjunction $\Delta$ of signed statement letters with n terms such that $\Gamma \vdash \Delta$ and such that no proper subdisjunction of $\Delta$ is a consequence of $\Gamma$. For such an example let $\Gamma = (P_1 \vee P_2 \vee Q_1) \wedge (\neg Q_1 \vee P_3 \vee Q_2) \wedge (\neg Q_2 \vee P_4 \vee Q_3) \wedge \ldots \wedge (\neg Q_{n-3} \vee P_{n-2} \vee Q_{n-2}) \wedge (\neg Q_{n-2} \vee P_{n-1} \vee P_n)$ and $\Delta = (P_1 \vee P_2 \vee \ldots \vee P_n)$.

*Corollary 1.* For any formula $\Gamma$ of a statement calculus the following two conditions are equivalent:

(1) $\Gamma$ is logically equivalent to a conjunction of binary disjunctions of signed statement letters.

(2) For every disjunction $\Delta$ of signed statement letters such that $\Gamma \vdash \Delta$,there exists a subdisjunction $\bar{\Delta}$ of $\Delta$ of no more than two terms such that $\Gamma \vdash \bar{\Delta}$.

Proof. By Theorem 1 it is sufficient to show that (2) implies (1).

Let $\Gamma$ be a formula satisfying (2) and let $\Omega$ be a formula in conjunctive normal form that is equivalent to $\Gamma$. Then each conjunct of $\Omega$ is a consequence of $\Gamma$ and by (2) there is a formula $\bar{\Omega}$ obtained from $\Omega$ by deleting all but two disjuncts from each conjunct of $\Omega$ and such that $\Gamma \vdash \bar{\Omega}$. Then $\bar{\Omega} \vdash \Omega$ by construction, so $\Gamma$ is logically equivalent to $\bar{\Omega}$.

## § 2. Binary Disjunctions in First-order Predicate Calculus.

Formulas referred to in this section will be assumed to be formulas of first-order languages without identity and without function symbols. We will briefly describe the proof theory of Linear Reasoning which we will use (vid. [2]). For any formulas $\Sigma$ and $\Omega$, an L-deduction of $\Omega$ from $\Sigma$ is an ordered (n+1)-tuple $\langle \Sigma_0, \Sigma_1, \ldots, \Sigma_n \rangle$ where $\Sigma_0 = \Sigma$ and $\Sigma_n = \Omega$, together with a specification of how, for $m<n$, $\Sigma_{m+1}$ results from $\Sigma_m$ by an application of an L-rule. The reader is referred to pages 252 and 253 of [2] for the definitions of the eleven L-rules. An L-deduc-

tion is symmetric if and only if the order in which the different kinds of L-rules are applied satisfies conditions (iii) through (vi) on page 257 of [2]. In addition, for convenience, we require that exactly one application of the L-rule *matrix change* occur in any symmetric L-deduction. Theorem 2 of [2] says that for any prenex formulas $\Gamma$ and $\Lambda$ such that $\Gamma \vdash \Lambda$ there is a symmetric L-deduction of $\Lambda$ from $\Gamma$.

We first show to what extent Theorem 1 above generalizes to first-order languages.

*Theorem 2.* Let $\Gamma$ and $\Lambda$ be prenex formulas such that $\Gamma \vdash \Lambda$. Assume that the matrix of $\Gamma$ is a conjunction of binary disjunctions of signed atomic formulas, that the matrix of $\Lambda$ is a conjunction of arbitrary disjunctions of signed atomic formulas, and that the prefix of $\Lambda$ contains no existential quantifiers. Then there is a formula $\bar{\Lambda}$ obtained from $\Lambda$ by deleting all but at most two disjuncts from each conjunct of the matrix of $\Lambda$ and such that $\Gamma \vdash \bar{\Lambda}$ and $\bar{\Lambda} \vdash \Lambda$.

Proof. Let $\Gamma$ and $\Lambda$ be formulas satisfying the hypotheses of the theorem and let $\Lambda^*$ be a formula obtained from $\Lambda$ by replacing the universally quantified variables with distinct individual constant symbols that do not occur in $\Gamma$ or $\Lambda$ and by deleting the universal quantifiers. Then $\Gamma \vdash \Lambda^*$, so there exists a symmetric L-deduction $\mathfrak{D}$ of $\Lambda^*$ from $\Gamma$. By properties of L-rules that may be applied after the application of matrix change in a symmetric L-deduction and since $\Lambda^*$ is quantifier free, we may assume that $\Lambda^*$ is the matrix of the formula resulting from the application of matrix change in $\mathfrak{D}$. Let $\Sigma$ be the formula of $\mathfrak{D}$ to which matrix change is applied. By properties of the L-rules that may be applied before the application of matrix change in a symmetric L-deduction, it follows that the matrix of $\Sigma$ is a conjunction of binary disjunctions of signed atomic formulas. By the definition of the L-rule matrix change, $\Lambda^*$ is a logical consequence of the matrix of $\Sigma$ and by Theorem 1 above there is a formula $\bar{\Lambda}^*$ obtained from $\Lambda^*$ by deleting all but at most two disjuncts from each conjunct of $\Lambda^*$ such that $\bar{\Lambda}^*$ is a logical consequence of the matrix of $\Sigma$. Thus $\Gamma \vdash \Sigma$ and $\Sigma \vdash \bar{\Lambda}^*$, so

$\Gamma \vdash \bar{\Delta}^*$. Let $\bar{\Delta}$ be the formula obtained from $\bar{\Delta}^*$ by replacing individual constant symbols, that were introduced to obtain $\Delta^*$ as indicated above, with their corresponding variables and by attaching the prefix of $\Delta$. Since $\Gamma \vdash \bar{\Delta}^*$ and the constant symbols replaced in $\bar{\Delta}^*$ to form $\bar{\Delta}$ do not occur in $\Gamma$, it follows that $\Gamma \vdash \bar{\Delta}$. Also $\bar{\Delta} \vdash \Delta$ by its construction, so $\bar{\Delta}$ has the properties required in the theorem.

We give an example which shows that the requirement that the prefix of $\Delta$ contains no existential quantifiers is essential in the above theorem. Let $\Gamma = \exists a \exists b((Fa \vee Gb) \wedge (Ga \vee Lb) \wedge (Ma \vee Hb) \wedge (Na \vee Mb))$ and $\Delta = \exists a((Fa \vee Ga \vee Ha) \wedge (La \vee Ma \vee Na))$. Then $\Gamma \vdash \Delta$ but for any formula $\bar{\Delta}$ obtained from $\Delta$ by deleting one disjunct from each conjunct of the matrix of $\Delta$, not $(\Gamma \vdash \bar{\Delta})$.

The requirement that the prefix of $\Delta$ contains no existential quantifiers can be deleted from this theorem, however, if one adds the requirement that the conjunction forming the matrix of $\Delta$ has exactly one term. This result can also be established with Linear Reasoning and Theorem 1 above using the fact that the formula resulting from matrix change in any symmetric L-deduction of such a formula $\Delta$ would have a matrix which is a disjunction of signed atomic formulas.

*Theorem 3*. Let $\Gamma$ and $\Delta$ be prenex formulas such that $\Gamma \vdash \Delta$. Assume that the matrix of $\Gamma$ is a conjunction of binary disjunctions of signed atomic formulas and that $\Gamma$ and $\Delta$ have a least one predicate letter in common. Then there exists a prenex formula $\Sigma$ such that the matrix of $\Sigma$ is a conjunction of binary disjunctions, such that all predicate letters in $\Sigma$ also occur both in $\Gamma$ and in $\Delta$, and such that $\Gamma \vdash \Sigma$ and $\Sigma \vdash \Delta$.

Proof. Observe that the theorem is a statement of a well known interpolation theorem (cf. Theorem 5 of [2]) with an added feature corresponding to an assumption that disjunctions are binary in $\Gamma$. Let $\Gamma$ and $\Delta$ be as described in the hypotheses of this theorem. Then by the known interpolation theorem cited, there is a formula $\Omega$ such that all predicate letters occurring in $\Omega$ also occur both in $\Gamma$ and in $\Delta$ and such that $\Gamma \vdash \Omega$ and $\Omega \vdash \Delta$.

Thus there is a symmetric L-deduction $\mathfrak{C}$ of $\Omega$ from $\Gamma$. Let $\vartheta$ and $\Phi$, respectively, be the formulas just before and just after matrix change in $\mathfrak{C}$. By properties of L-rules which occur after matrix change in a symmetric L-deduction, the predicate letters that occur in $\Phi$ are exactly those that occur in $\Omega$. Let $\pi$ be a formula obtained from $\Phi$ by writing its matrix in conjunctive normal form. The matrix of $\pi$ is a logical consequence of the matrix of $\vartheta$ (which is also a conjunction of binary disjunctions of signed atomic formulas) and we may apply Theorem 1 above to the matrix of $\vartheta$ together with each conjunct of the matrix of $\pi$. Thus there is a formula $\Sigma$ obtained from $\pi$ by deleting all but at most two disjuncts from each conjunct of the matrix of $\pi$ and such that $\vartheta \vdash \Sigma$. Since, by construction, $\Sigma \vdash \pi$, it follows that $\Sigma$ is a formula as required in the theorem.

We observe that Theorem 3 can be used to establish a correspondingly modified version of a theorem known as Beth's Theorem on Definability. In particular, in Theorem 5.2.1 page 118 of [6], if we add an assumption that the sentence defining the relation *implicitly* is in prenex conjunctive form in which all disjunctions are binary then we may conclude also that the sentence defining the relation *explicitly* may be assumed to be in prenex conjunctive normal form in which all disjunctions are binary.

Theorem 3 has a consequence concerning the existence of reduction classes for satisfiability of formulas in pure first-order languages. Let $L_1$ and $L_2$ be any two first-order languages and let $\mathfrak{R}$ be a set of formulas of $L_1$. Then $\mathfrak{R}$ is called a *reduction class* for satisfiability of formulas of $L_2$ in case there is an effective procedure for finding, for any formula X of $L_2$, an associated formula $\varrho(X)$ in $\mathfrak{R}$ such that X is satisfiable (has a model) if and only if $\varrho(X)$ is satisfiable (cf. page 32 of [7]). Now suppose that all predicate letters that occur in $L_2$ also occur in $L_1$. Then for any structure (relational system) $\mathfrak{M}$ of the similarity type of $L_1$, the *$L_2$-reduct* of $\mathfrak{M}$ is the structure obtained from $\mathfrak{M}$ by deleting all predicates for which there is no corresponding predicate letter in $L_2$. Also, if $\mathfrak{N}$ is an $L_2$-reduct of a structure $\mathfrak{M}$ of the similarity type of $L_1$, then we say that $\mathfrak{M}$ is an *$L_1$-expansion* of $\mathfrak{N}$. We will say that $\mathfrak{R}$ is a *strong re-*

*duction class* for satisfiability of formulas of $L_2$ in case $\mathfrak{R}$ is a reduction class such that for any X of $L_2$ and any model $\mathfrak{M}$ of $\varrho(X)$ the $L_2$-reduct of $\mathfrak{M}$ is a model of X and for any model $\mathfrak{N}$ of X, $\mathfrak{N}$ has an $L_1$-expansion which is a model of $\varrho(X)$. We observe that the Skolem normal form for satisfiability (cf. Satz VII page 49 of [7]) determines a strong reduction class. If $L_2$ is any pure first-order language and $L_1$ is an extension of $L_2$ including infinitely many new predicate letters of all ranks and $\mathfrak{R}$ is the class of the prenex formulas of $L_1$ with the Skolem normal form prefix, then $\mathfrak{R}$ is a strong reduction class for satisfiability of formulas of $L_2$.

*Theorem 4.* For any pure first-order languages $L_1$ and $L_2$ such that $L_2$ has at least three predicate letters and such that $L_1$ is an extension of $L_2$, the class of the prenex formulas of $L_1$ whose matrices are conjunctions of binary disjunctions of signed atomic formulas is not a strong reduction class for satisfiability of the formulas of $L_2$.

Proof. Let $L_1$ and $L_2$ be pure first-order languages such that $L_1$ is an extension of $L_2$ and suppose that the class $\mathfrak{B}$ of prenex conjunctive formulas of $L_1$ in which all disjunctions are binary is a strong reduction class for $L_2$. Let X be a formula of $L_2$ and let $\varrho(X)$ be the associated formula of $\mathfrak{B}$. By definition of strong reduction, $\varrho(X) \vdash X$. By Theorem 3 above there is a formula $\Sigma$ of $\mathfrak{B}$ such that $\varrho(X) \vdash \Sigma$, $\Sigma \vdash X$ and such that all predicate letters occurring in $\Sigma$ also occur in X. But any model $\mathfrak{N}$ of X has some $L_1$-expansion $\mathfrak{N}'$ which is a model of $\varrho(X)$. Since $\varrho(X) \vdash \Sigma$, $\mathfrak{N}'$ is a model of $\Sigma$ and thus also the $L_2$-reduct $\mathfrak{N}$ of $\mathfrak{N}'$ is a model of $\Sigma$. Thus $X \vdash \Sigma$ and we conclude that for any X of $L_2$ there is a logically equivalent formula which is in prenex conjunctive normal form and in which all disjunctions are binary. But we will show that this is not true for first-order languages $L_2$ with at least three predicate letters. Suppose $L_2$ is a first-order language in which three predicate letters $\alpha$, $\beta$, and $\gamma$ occur. Let $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\gamma}$, respectively, be the formulas obtained from these predicate letters by introducing distinct individual variables into each argument place and then universally quantifying the variables. Let X be $\bar{\alpha} \vee \bar{\beta} \vee \bar{\gamma}$. Suppose that $\Sigma$ is a prenex conjunctive formula in which all disjunctions are binary and such that all predicate letters that occur in $\Sigma$ also occur in X. By considering various one element structures as possible models, we can see that $\Sigma$ can not be logically equivalent to X. In particular, either $\Sigma$ is satisfied by all one element structures or there is a conjunct of $\Sigma$ in which the two disjuncts are not a negated and an unnegated instance of the same predicate letter. But if there is such a conjunct then the weaker formula $\overline{\Sigma}$ obtained by deleting all but that conjunct from the matrix of $\Sigma$ is not a logical consequence of X.

Notice that there are instances of Theorem 4 in which $L_2$ is a statement calculus, i.e. when $L_2$ is a language with just three predicate letters each one being of rank zero. The statement obtained from Theorem 4 by replacing "three predicate letters" with "one binary predicate letter" is also true. That follows from the first part of the above proof and Theorem 1 of [3].

Finally we observe that statements corresponding to Theorems 3 and 4 above with "binary disjunctions" replaced by "ternary disjunctions" are false. In particular there is a simple procedure for strongly reducing first-order formulas to formulas in prenex conjunctive normal form in which disjunctions are ternary. For an illustration consider the replacement of $P_1 \vee P_2 \vee P_3 \vee P_4$ with $(P_1 \vee P_2 \vee Q) \wedge (\neg Q \vee P_3 \vee P_4)$ where $P_1$, $P_2$, $P_3$, $P_4$, and Q are distinct statement letters. For further details of how this reduction can be carried out for arbitrary first-order formulas see the proof of Theorem 1 in [5].

*University of California, Davis* M. R. KROM

## BIBLIOGRAPHY

[1] C. C. CHANG and H. Jerome KEISLER, An Improved Prenex Normal Form, *Jour. Symb. Logic* 27 (1962), 317-326.

[2] W. CRAIG, Linear Reasoning. A New Form of the Herbrand-Gentzen Theorem, *Jour. Symb. Logic* 22 (1957), 250-268.

[3] M. R. KROM, A Property of Sentences that Define Quasi-order, *Notre Dame Journ. of Formal Logic,* forthcoming.

[4] M. R. KROM, The Decision Problem for a Class of First-order Formulas in which all Disjunctions are Binary, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik,* 13 (1967), 15-20.

[5] M. R. KROM, The Decision Problem for Segregated Formulas in Frist-order Logic, *Mathematica Scandinavica,* forthcoming.

[6] Abraham ROBINSON, *Introduction to Model Theory and to the Metamathematics of Algebra,* North-Holland Publishing Company, Amsterdam (1963).

[7] Janos SURANYI, *Reductionstheorie Des Entscheidungsproblems,* Akademiai Kiado, Budapest (1959).