命题逻辑公式中的冗余子句及冗余文字

翟翠红 秦克云
(西南交通大学数学学院 成都 610031)

摘 要 主要研究命题逻辑公式中的冗余子句和冗余文字。针对子句集中必需的、有用的、无用的子句,分别给出了一些等价描述方法,进而讨论子句集的无冗余等价子集。另外,得到了子句集中冗余文字的判别方法,借助可满足性给出了冗余子句的一种等价条件。上述结果为命题逻辑公式的化简奠定了一些理论基础。

关键词 冗余子句,冗余文字,无冗余等价子集,可满足性
中图法分类号 O141.1 **文献标识码** A

Redundancy Clause and Redundancy Literal of Propositional Logic

ZHAI Cui-hong QIN Ke-yun
(College of Mathematics, Southwest Jiaotong University, Chengdu 610031, China)

Abstract We mainly studied the redundancy of clause and literal in the propositional logic formula. Some equivalent descriptions of necessary, useful and useless clause for a particular set of clauses were given respectively, and the irredundant equivalent subsets of any set of clauses were discussed. In addition, an approach of identification of redundant literal was presented. At last, an equivalent condition of the redundancy clause was obtained by using the concept of satisfiability. These results serve as the theoretical foundation of a new method for simplifying the propositional logic formula.

Keywords Redundancy clause, Redundancy literal, Irredundant equivalent subset, Satisfiability

一个知识库是冗余的,如果它包含有在不减少其所携信息的情况下可以删除的部分。冗余以及与其密切相关的化简已经成为具有重要现实意义的问题。首先,化简在某些情况下有一些计算方面的优势(例如,它可以降低指数);此外,代表同样信息的简化了的公式更容易理解,因为冗余部分可能会掩盖或混淆所要表达的信息。

逻辑公式的可满足性问题是理论计算机科学和人工智能中的著名问题。世界各国学者在这方面做了大量的研究工作,特别是命题逻辑的可满足性问题(即 SAT)[1,2]。其中,逻辑公式的化简是重要的研究方向[3-7]。命题逻辑(又称命题演算、布尔逻辑)是最简单的一种形式逻辑系统,命题是其研究对象。每个命题可能为真,也可能为假。我们用 1,0 分别表示命题的真、假值。本文主要研究命题逻辑中公式的冗余子句和冗余文字。针对子句集中的必需的、有用的、无用的子句,分别给出一些等价描述,进而讨论子句集的无冗余等价子集。另外,得到了子句集中冗余文字的判别方法。

1 冗余子句和无冗余子句集

本文讨论命题逻辑中的公式。称原子公式及其否定为文字(literal),若干个文字的析取称为子句(clause)。合取范式(conjunctive normal form,CNF)是如下形式的公式:

$$(x_{11}\vee\cdots\vee x_{1n_1})\wedge\cdots\wedge(x_{m1}\vee\cdots\vee x_{mn_m})$$

这里的 x_{ij} 是文字。

设 Π 和 Π' 为两个子句集。如果对任一赋值 ν,都有 $\nu(\wedge\Pi)=\nu(\wedge\Pi')$,则称 Π 与 Π' 等价,记作 $\Pi\equiv\Pi'$;如果对任一赋值 ν,当 $\nu(\Pi)=1$ 时,有 $\nu(\Pi')=1$,则称 Π 可以语义导出 Π',记作 $\Pi\models\Pi'$[8]。

约定 文中出现的子句集如不加特别说明,则默认为是 CNF 形式的命题逻辑公式,并假设子句集中的子句为非重言式。

定义 1[3] 设 Π 是子句集,$\gamma\in\Pi$,称 γ 在 Π 中是冗余的,如果 $\Pi-\{\gamma\}\models\gamma$。

设 $S=\{C_1,\cdots,C_m,D\}$ 是命题逻辑中的子句集。显然,D 是 S 中的冗余子句,当且仅当:

$$C_1\wedge\cdots\wedge C_m\wedge D\equiv C_1\wedge\cdots\wedge C_m$$

从上面的定义中可以看出,一个子句是冗余的,暗示此子句可以从子句集中删除[4,5],不会影响子句集所要表示的信息。同理,一个子句集是冗余的可以定义为它和它的一个真子集等价。

定义 2[3] 子句集 Π 是冗余的,当且仅当存在 $\Pi'\subset\Pi$,使 $\Pi'\equiv\Pi$。

在命题逻辑中,此定义和以下说法是等价的:

到稿日期:2012-07-20 返修日期:2012-10-13 本文受国家自然科学基金项目(61175055)资助。
翟翠红(1987-),女,硕士生,主要研究方向为粗糙集理论与方法,E-mail:865031165@163.com;秦克云(1962-),男,教授,博士生导师,主要研究方向为多值逻辑、粗糙集理论与方法。

(1)存在 $\Pi' \subset \Pi$,使 $\Pi' \mid = \Pi$;

(2)Π 中含有冗余子句。

定义 3[3] 称子句集 Π' 是 Π 的无冗余等价子集(irredundant equivalent subset, I. E. S.),如果:

(1)$\Pi' \subseteq \Pi$;

(2)$\Pi' \equiv \Pi$;

(3)Π' 是非冗余的。

任何一个子句集至少含有一个无冗余等价子集。

例 1 令 $\Pi = \{a \vee \bar{b}, \bar{a} \vee b, a \vee c, b \vee c\}$。$\Pi_1 = \Pi - \{a \vee c\}$ 是 Π 的一个无冗余等价子集。事实上,若赋值 ν 使得 $\nu(\Pi_1) = 1$,当 $\nu(a) = 1$ 时,显然 $\nu(a \vee c) = 1$;当 $\nu(a) = 0$ 时,由 $\nu(a \vee \bar{b}) = 1$ 可得 $\nu(b) = 0$,又由 $\nu(b \vee c) = 1$ 可得 $\nu(c) = 1$,从而 $\nu(a \vee c) = 1$。于是 $\Pi' \equiv \Pi$。若赋值 ν 满足 $\nu(b) = \nu(b) = \nu(a) = 0$,则 $\nu(a \vee \bar{b}) = 1, \nu(\bar{a} \vee b) = 1, \nu(b \vee c) = 0$。故 $b \vee c$ 不是 Π_1 中的冗余子句。同理,可以验证 $a \vee \bar{b}, \bar{a} \vee b$ 都不是 Π_1 中的冗余子句,即 Π_1 是非冗余子句集。

同理,$\Pi_2 = \Pi \backslash \{b \vee c\}$ 也是 Π 的一个无冗余等价子集。

如果令 $\Pi_n = \bigcup_{i=1,\cdots,n} \Pi[\{a/a_i, b/b_i, c/c_i\}]$,从 Π_n 中删除冗余子句时有 n 个相互独立的选择,并且 $\forall i \in \{1, 2, \cdots, n\}$,我们可以删除 $a_i \vee c_i$、$b_i \vee c_i$ 中的一个,这就说明 Π_n 有 2^n 个无冗余等价子集。

一个公式可能含有多个 I. E. S.,根据这种情况可以把公式中的子句分为如下 3 种类型。

定义 4[3] 设 Π 是子句集,$\gamma \in \Pi$。

(1)称 γ 在 Π 中是必需的(necessary),如果对于 Π 的任一无冗余等价子集 Π',有 $\gamma \in \Pi'$。

(2)称 γ 在 Π 中是有用的(useful),如果存在 Π 的一个无冗余等价子集 Π',使 $\gamma \in \Pi'$。

(3)称 γ 在 Π 中是无用的(useless),如果对于 Π 的任一无冗余等价子集 Π',有 $\gamma \notin \Pi'$。

从定义中注意到,必需的子句也是有用的,而有用的和无用的是两个相对的概念。就知识库来说,必需的子句是非冗余的,无用的子句是"强冗余"的,即它不但可以被删除,而且总可以删除。文献[3]中给出了子句是必需的等价描述。

定理 1[3] 设 Π 是子句集,$\gamma \in \Pi$。γ 在 Π 中是必需的当且仅当 $\Pi - \{\gamma\} \mid \neq \gamma$。

下面给出有用和无用子句的等价描述。

定理 2 设 Π 是子句集,$\gamma \in \Pi$。γ 在 Π 中是有用的当且仅当存在 Π 的一个无冗余等价子集 Π',使 $\Pi' - \{\gamma\} \mid \neq \gamma$。

证明:(充分性)设 Π' 是 Π 的无冗余等价子集,且 $\Pi' - \{\gamma\} \mid \neq \gamma$。于是存在一个赋值 ν,使 $\nu(\Pi' - \{\gamma\}) = 1, \nu(\gamma) = 0$。若 $\gamma \notin \Pi'$,则有 $\nu(\Pi' - \{\gamma\}) = \nu(\Pi') = 1$,但由 $\gamma \in \Pi, \nu(\gamma) = 0$ 知 $\nu(\Pi) = 0$,这与 $\Pi' \equiv \Pi$ 矛盾,故 $\gamma \in \Pi'$,即 γ 是有用的。

(必要性)设 γ 在 Π 中是有用的,则存在 Π 的一个无冗余等价子集 Π',使 $\gamma \in \Pi'$,又 Π' 是无冗余的,Π' 的无冗余等价子集只有 $\Pi', \gamma \in \Pi'$。故 γ 在 Π' 中是必需的,由定理 1 知 $\Pi' - \{\gamma\} \mid \neq \gamma$。

定理 3 设 Π 是子句集,$\gamma \in \Pi$。γ 在 Π 中是无用的当且仅当 Π 的无冗余等价子集恰为 $\Pi - \{\gamma\}$ 的无冗余等价子集。

证明:(充分性)设 Π' 是 Π 的无冗余等价子集,则 Π' 亦是 $\Pi - \{\gamma\}$ 的无冗余等价子集,故 $\Pi' \subseteq \Pi - \{\gamma\}$,所以 $\gamma \notin \Pi'$。故 γ 在 Π 中是无用的。

(必要性)设 γ 是无用的。若 Π' 是 Π 的无冗余等价子集,则 $\gamma \notin \Pi'$,于是 $\Pi' \subseteq \Pi - \{\gamma\}$,$\wedge \Pi' \geqslant \wedge(\Pi - \{\gamma\}) \geqslant \wedge \Pi$,又 $\wedge \Pi' = \wedge \Pi$,故 $\wedge \Pi' = \wedge(\Pi - \{\gamma\})$,又 Π' 是无冗余的,故 Π' 是 $\Pi - \{\gamma\}$ 的无冗余等价子集。另一方面,设 Π' 是 $\Pi - \{\gamma\}$ 的无冗余等价子集,则 Π' 是无冗余的,$\Pi' \subseteq \Pi - \{\gamma\} \subseteq \Pi$,且由定理 1 可得 $\Pi - \{\gamma\} \mid = \gamma$,于是有:

$$\wedge \Pi' = \wedge(\Pi - \{\gamma\}) = \wedge \Pi$$

即 Π' 是 Π 的无冗余等价子集。

借助定理 2 与定理 3,可以判断一个子句集中某子句是否为有用子句或无用子句。

定理 4[3] 设 Π 是子句集,Π 有唯一的无冗余等价子集,当且仅当 $\Pi_N \equiv \Pi$,其中 $\Pi_N = \{\gamma \in \Pi \mid \gamma$ 在 Π 中是必需的$\}$。

利用定理 4 可以判断任一子句集的无冗余等价子集是否唯一。

2 冗余文字

本节讨论子句中的冗余文字。

定义 5 设 $S = \{C_1, \cdots, C_m, D\}$ 是命题逻辑中子句集,$D = x \vee D_1$,其中 x 是一文字,D_1 是一子句,如果

$$D \wedge C_1 \wedge \cdots \wedge C_m = D_1 \wedge C_1 \wedge \cdots \wedge C_m$$

则称 x 是 D 中关于 S 的冗余文字。

定理 5 设 S, x, D 如定义 5,如果 D_1 是 $S' = \{C_1, \cdots, C_m, D_1\}$ 中的冗余子句,则 x 是 D 中关于 S 的冗余文字。

证明:因 D_1 是 S' 中的冗余子句,故 $D_1 \wedge C_1 \wedge \cdots \wedge C_m = C_1 \wedge \cdots \wedge C_m$,从而

$$\begin{aligned} & D \wedge C_1 \wedge \cdots \wedge C_m \\ &= (D_1 \vee x) \wedge C_1 \wedge \cdots \wedge C_m \\ &= (D_1 \wedge C_1 \wedge \cdots \wedge C_m) \vee (x \wedge C_1 \wedge \cdots \wedge C_m) \\ &= (C_1 \wedge \cdots \wedge C_m) \vee (x \wedge C_1 \wedge \cdots \wedge C_m) \\ &= C_1 \wedge \cdots \wedge C_m = D_1 \wedge C_1 \wedge \cdots \wedge C_m \end{aligned}$$

即 x 是 D 中关于 S 的冗余文字。

定理 6 设 S, x, D 如定义 5,x 是 D 中关于 S 的冗余文字当且仅当 D_1 是子句集 $S' = \{D_1, x, C_1, \cdots, C_m\}$ 中的冗余子句。

证明:(充分性)因 D_1 是 S' 中冗余子句,故 $D_1 \wedge x \wedge C_1 \wedge \cdots \wedge C_m = x \wedge C_1 \wedge \cdots \wedge C_m$。

于是有:

$$\begin{aligned} & D \wedge C_1 \wedge \cdots \wedge C_m \\ &= (D_1 \vee x) \wedge C_1 \wedge \cdots \wedge C_m \\ &= (D_1 \wedge C_1 \wedge \cdots \wedge C_m) \vee (x \wedge C_1 \wedge \cdots \wedge C_m) \\ &= (D_1 \wedge C_1 \wedge \cdots \wedge C_m) \vee (D_1 \wedge x \wedge C_1 \wedge \cdots \wedge C_m) \\ &= D_1 \wedge C_1 \wedge \cdots \wedge C_m \end{aligned}$$

即 x 是 D 中关于 S 的冗余文字。

(必要性)设 x 是 D 中关于 S 的冗余文字,则

$$\begin{aligned} & (D_1 \wedge C_1 \wedge \cdots \wedge C_m) \vee (x \wedge C_1 \wedge \cdots \wedge C_m) \\ &= (D_1 \vee x) \wedge C_1 \wedge \cdots \wedge C_m \\ &= D_1 \wedge C_1 \wedge \cdots \wedge C_m \end{aligned}$$

故 $x \wedge C_1 \wedge \cdots \wedge C_m \leqslant D_1 \wedge C_1 \wedge \cdots \wedge C_m, x \wedge C_1 \wedge \cdots \wedge C_m \leqslant$

D_1，从而

$$D_1 \wedge x \wedge C_1 \wedge \cdots \wedge C_m = x \wedge C_1 \wedge \cdots \wedge C_m$$

即 D_1 是 S' 中冗余子句。

定理 5 与定理 6 给出了子句集中某子句中的某个文字是冗余文字的等价刻画，可用于命题逻辑公式的化简。

注：设 S, D 如定义 5，若 D 是 S 中的冗余子句，D 中的文字并不一定是 D 中关于 S 的冗余文字。例如，令 $S=\{x_1, x_1 \vee x_2\}$，显然 $D=x_1 \vee x_2$ 是子句集 S 中的冗余子句，当 $\nu(x_1)=1, \nu(x_2)=0$ 时，有 $x_1 \wedge x_2 \neq x_1 \wedge (x_1 \vee x_2)$，所以 x_1 不是 D 中关于 S 的冗余文字。

3 公式的可满足性判定

徐扬在文献[6]中针对析取范式讨论了命题逻辑公式中文字与小项的可消性，给出了文字及小项可消去的等价条件。在此基础上，本节讨论子句冗余性与可满足性的关系。

对于子句 A, B，以下用 $\underline{A}$ 表示 A 中出现的文字构成的集合，且 $A-B=\bigvee_{x \in \underline{A}-\underline{B}} x$。

定理 7 设 $S=\{C_1, \cdots, C_m, D\}$ 是命题逻辑中子句集，且 D 中不含互补文字。D 是 S 中的冗余子句当且仅当子句集 $S'=\{C_1-D, C_2-D, \cdots, C_m-D\}$ 不可满足。

证明：(充分性)若 D 不是 S 中的冗余子句，则

$$C_1 \wedge C_2 \wedge \cdots \wedge C_m \neq C_1 \wedge \cdots \wedge C_m \wedge D$$

于是存在赋值 v 使得 $v(C_1 \wedge C_2 \wedge \cdots \wedge C_m)=1$ 且 $v(C_1 \wedge \cdots \wedge C_m \wedge D)=0$，从而有 $v(D)=0$。故对于任意文字 $y \in \underline{D}$，有 $v(y)=0$。对于任意的 $i, 1 \leqslant i \leqslant m$，有 $v(C_i)=1$，从而存在 $x_i \in \underline{C_i}$ 使 $v(x_i)=1$。显然 $x_i \notin \underline{D}$，即 $x_i \in \underline{C_i}-\underline{D}$，故有 $v(C_i-D)=1$。由 i 的任意性可得 S' 是可满足的，矛盾。

(必要性)假设 S' 是可满足的。由于

$$(C_1-D) \wedge (C_2-D) \wedge \cdots \wedge (C_m-D) = \bigvee_{(x_1, \cdots, x_m) \in \underline{C_1}-\underline{D} \times \cdots \times \underline{C_m}-\underline{D}} x_1 \wedge \cdots \wedge x_m$$

故存在赋值 v 及 $(x_1, \cdots, x_m) \in \underline{C_1}-\underline{D} \times \cdots \times \underline{C_m}-\underline{D}$ 使得 $v(x_1)=v(x_2)=\cdots=v(x_m)=1$。

由于 $\{x_1, \cdots, x_m\} \cap \underline{D}=\emptyset$，且 D 中不含互补文字，可设对于任意 $y \in \underline{D}, v(y)=0$，于是有 $v(D)=0$。对于任意 $C_i, 1 \leqslant i \leqslant m$，注意到 $x_i \in C_i$ 且 $v(x_i)=1$，故 $v(C_i)=1$。于是可得

$$v(C_1 \wedge C_2 \wedge \cdots \wedge C_m)=v(C_1) \wedge \cdots \wedge v(C_m)=1$$

$$v(C_1 \wedge C_2 \wedge \cdots \wedge C_m \wedge D)=v(C_1) \wedge \cdots \wedge v(C_m) \wedge v(D)=0$$

故 $C_1 \wedge \cdots \wedge C_m \neq C_1 \wedge \cdots \wedge C_m \wedge D$，此与 D 是 S 中冗余子句矛盾。

由定理 6、定理 7 可得如下推论：

推论 设 S, x, D 如定义 5，则 x 是 D 中关于 S 的冗余文字当且仅当子句集 $\{x, C_1-D_1, \cdots, C_m-D_1\}$ 不可满足。

参 考 文 献

[1] 张健. 逻辑公式的可满足性判定-方法、工具及应用[M]. 北京：科学出版社，2000

[2] 秦永彬，张秋菊. 基于关键文字的求解 SAT 问题的启发式算法[J]. 计算机与数字工程，2010，38：1-4

[3] Liberatore P. Redundancy in logic I：CNF propositional formulae [J]. Artificial Intelligence，2005，163(30)：203-232

[4] Liberatore P. Redundancy in logic Ⅱ：2CNF and Horn propositional formulae [J]. Artificial Intelligence，2008，172(35)：265-299

[5] Gottlob G，Fermüller C G. Removing redundancy from a clause [J]. Artificial Intelligence，1993，61(27)：263-289

[6] 徐扬，邹开其. 布尔逻辑公式中文字和小项的可消性[J]. 西南交通大学学报，1990，1：107-112

[7] Ostrowski R，Mazure B，Saïs L，et al. Eliminating redundancies in SAT search trees [C]// Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence (ICTA' 2003). Sacramento，2003，5：100-104

[8] 王国俊. 数理逻辑引论与归结原理[M]. 北京：科学出版社，2006

(上接第 20 页)

$$0 \leqslant \sigma_{f,g} \leqslant 2^{4i-n+2n-2i}=2^{n+2i}$$

且 $\sigma_{f,g}=0$ 当且仅当 $\tilde{f} \cdot \tilde{g}=0$；当 $f=g$ 时，有 $\sigma_f=2^{n+2i}$。

本文借助于 Plateaued 函数的定义及函数分解理论得到了任一 Plateaued 函数与其对偶函数之间的一些关系，并运用 Plateaued 函数的对偶这一工具得到了任意两个 Plateaued 函数的互相关函数平方和指标的界。

参 考 文 献

[1] Dillon J F. Elementary Hadamard Difference sets [D]. University of Maryland，1974

[2] Canteaut A，Charpin P. Decomposing bent functions[J]. IEEE Transactions on Information Theory，2003，49(8)：2004-2019

[3] 卓泽朋，崇金凤，高胜，等. Bent 函数的对偶性[J]. 武汉大学学报：理学版，2012，58(1)：086-088

[4] Carlet C. Partially-bent functions[C]// Advances in Cryptology-CRYPTO' 92，Lecture Notes in Computer Science. Springer-Verlag，1993，740：280-291

[5] Zheng Y，Zhang X M. Plateaued functions[C]// Heidelberg，ed. Advances in Cryptology-ICICS' 99，Lecture Notes in Computer Science. Springer-Verlag，1999，1726：284-300

[6] Gong G，Khoo K. Additive Autocorrelation of Resilient Boolean Functions[C]// Selected Areas in Cryptography 2003，Lecture Notes in Computer Science. Springer-Verlag，2003，3006：275-290

[7] Zhou Yu，Xie Min，Xiao Guo-zhen. On the global avalanche characteristics of two Boolean functions and the higher order nonlinearity[J]. Information Sciences，2010，180：256-265

[8] Canteaut A，Carlet C，Charpin P，et al. On cryptographic properties of the cosets of R(1，m)[J]. IEEE Transactions on Information Theory，2001，47(4)：1494-1513

[9] Wang Wei-qiong，Xiao Guo-zhen. Decomposition and Construction of Plateaued Functions[J]. Chinese Journal of Electronics，2009，18(4)：686-688