	まつ お けん し
氏　　名　（本籍）	松 尾 健 史　　（北 海 道）
学位の種類	博 士（情報科学）
学位記番号	情 博 第 283 号
学位受与年月日	平成１６年3月２５日
学位授与の要件	学位規則第４条第１項該当
研究科、専攻	東北大学大学院情報科学研究科（博士課程）情報基礎科学専攻
学位論文題目	ブール関数の表現の複雑さに関する研究
論文審査委員	（主 査） 東北大学教授　丸岡　章　　　東北大学教授　亀山　充隆 東北大学教授　阿曽　弘具 （工学研究科）

論 文 内 容 要 旨

第１章 序論

コンピュータサイエンスにおいて，様々な現実の問題に対し，コンピュータで何が出来て何が出来ないのか，出来るものについては，何が現実的な時間内でできて，何ができないのかを明らかにすることは重要である．一般に，コンピュータで解くべき問題は，ブール関数に関連するものが多い．これらの問題では，ブール関数 f をどのように表現してコンピュータに入力するかが問題になり，f の表現形式に依存して計算時間が大きく違ってくることもある．ブール関数の表現形式としては様々なものがある．それぞれの表現方法において，その表現形式の持つ特徴は異なる．例えば，表現サイズに関して，同じブール関数 f を表現するときも，選択する表現形式により，それぞれ関数を表現する記述長，すなわち，サイズが異なり，時には指数関数的な差が生じる場合もある．本研究の目的は，表現形式の特徴づけと，表現形式間の比較を理論的に解析することである．この研究の意義とは，予め様々な表現形式における問題の難しさ，あるいは表現形式間の比較をしておけば，ある応用の実現のためにどの表現形式を採用すべきかを考えるとき，何が可能で何が不可能かを予測することができ，開発の効率化が期待できるからである．

さて，本研究では，上で述べた目的の立場に立ち，特に DNF 式（積和標準形）に関連する研究を行った．まず，単調 DNF 式に基づく排他的論理和展開式（XOR-MDNF 式，$\oplus$MDNF 式）のような，DNF 式を拡張したと見なせる表現形式や，ホーン DNF 式やダブルホーン関数の様な，DNF 式にある制限を加えた表現形式において，それぞれの表現形式の特徴づけ，あるいは表現形式間の比較を行った．また，本研究では，現実的な時間内で正確に値を求めることが難しいと予想されている，DNF 式の充足割り当てを数え上げる問題に注目をし，その問題のある解法において，計算時間の短縮化が期待できる興味深い性質について研究を行った．

第２章 ブール関数の表現の複雑さ

第２章では，本研究の主題であるブール関数の表現の複雑さを議論する上で必要となる，ブール関数の基本事項や諸定義について述べた．また，DNF 式や，CNF 式（和積標準形）など，一般的に知られている表現形式において，その表現形式の持つ特徴や，他の表現形式との関係について，既知の結果を紹介した．さらに，計算学習理論に関する結果についても紹介した．

第３章 ホーン式と XOR-MDNF 式の関係

第３章では，人工知能をはじめ，コンピュータサイエンスで広く用いられているホーン式に注目した．なぜなら，もし，ある表現形式がホーン式と等価であるならば，その表現形式に基づいた学習の問題に対しても，ホーン式に関する多くの研究成果を利用することができ，さまざまな応用が期待できるからである．この章では，計

算学習の分野でしばしば用いられる，$\oplus$MDNF 式を取り上げ，その表現能力とホーン式の表現能力が，ある条件の下では互いに等価であることを示した．

$\oplus$MDNF 式とは，否定的なリテラルを含まない DNF 式（単調 DNF 式）$f_1, f_2, \ldots, f_d$（ただし，$f_1 > \cdots > f_d$）に対し，

$$f = f_1 \oplus f_2 \oplus \cdots \oplus f_d$$

と表現されるブール式である．ただし，同じ項は式 f の中に 2 回以上現れないものとする．ここで，2 つの関数 f と g に対し，不等式 $f > g$ とは $f^{-1}(1) \supset g^{-1}(1)$ を意味する．また，式を構成する DNF 式の個数 d を，f の深さと呼ぶ．$\oplus$1-MDNF 式とは，f の各 DNF 式 f_i がただ 1 つの項からなるブール式である．$\oplus$1-MDNF 式のクラスを $C_{\oplus 1-\mathrm{MDNF}}$ と書く．次に，ホーン DNF 式とはそれぞれの項が高々1 つの否定リテラルを持つ DNF 式である．また，ダブルホーン関数とは f と $\bar{f}$ がどちらもホーン DNF 式で表せる関数である．ダブルホーン DNF 式のクラスを C_{DH} と書く．1 決定リストとは，決定リスト $((t_1, a_1), \ldots, (t_s, a_s), a_{s+1})$ において，各 t_i が 1 つのリテラルからなる決定リストを表し，そのクラスを $C_{1-\mathrm{DL}}$ と書く．$w \in \{0,1\}^n$ に対して，w による f のリネーミング f^w とは，$f^w = f(x \oplus w)$ で与えられる関数である．関数クラス C に対し，C^R をリネーミングで閉じたクラスとする．すなわち，$C^R = \{f^w \mid f : \{0,1\}^n \to \{0,1\} \in C, w \in \{0,1\}^n, n \geq 1\}$.

本研究では，$C_{\mathrm{DH}} = C_{\oplus 1-\mathrm{MDNF}}$，すなわち，$\oplus$1-MDNF 式のクラスとダブルホーン関数のクラスが一致することを示した．また，Eiter らの $C^R_{\mathrm{DH}} = C_{1-\mathrm{DL}}$ の結果より，$C^R_{\oplus 1-\mathrm{MDNF}} = C_{1-\mathrm{DL}}$ であることも示せた．

次に，与えられたホーン DNF 式を $\oplus$MDNF 表現に変換する問題について考えた．$\oplus$MDNF 表現のサイズがホーン DNF 表現のサイズに関して指数的になることがあるので，明らかにこの問題に対する多項式アルゴリズムは存在しない．そこで，この問題を，求める $\oplus$MDNF 式 $f = f_1 \oplus \cdots \oplus f_d$ の各 DNF 式 f_i を列挙する問題と考え，アルゴリズムの効率を，f_{i-1} が出力されてから f_i が出力されるまでの時間で評価することにする．任意の $1 \leq i \leq d$ に対し，f_i が出力される時間が，入力サイズとそれまでの出力 $f_1, \ldots, f_{i-1}$ のサイズの多項式の時間で求められるとき，この問題は逐次多項式時間で解けると言う．本研究では，ホーン DNF 式を $\oplus$MDNF 式に変換する逐次多項式時間アルゴリズムを与えた．

第 4 章 XOR-MDNF 式の多項式しきい値関数表現

ブール関数の多項式しきい値関数（PTF）という表現形式に注目する．PTF とは，ブール関数をある実数係数の多重線形多項式の符号で表す表現形式である．すなわち，

$$p(x) = \sum_{S \subseteq \{1,\ldots,n\}} C_S \prod_{i \in S} x_i,\ C_S\text{は適当な実数}$$

によって，$x \in \{+1,-1\}^n$ を $\mathrm{sgn}(p(x)) \in \{+1,-1\}$ に写像する関数である．PTF 表現に関する研究は，ブール関数の複雑さ，あるいは，計算学習理論の分野において，幅広く行われており，ある表現形式を PTF 表現で表すとき，PTF 表現のサイズが入力サイズの多項式程度しか大きくならないことを示すことは，PTF の研究により得られた多くの研究成果を，この表現形式においても利用できるようになるため，意義のあることである．本章では，第 3 章で注目した $\oplus$MDNF 式を PTF で表現するときに必要なサイズについて論じた．まず，Jackson の結果と，Freund と Schapire の結果を組み合わせることにより，任意の DNF 式 f は，f のサイズより多項式程度しか大きくならないサイズの PTF 表現を持つということが示せる．この結果は，定数個の DNF 式を排他的論理和展開式で結んで得られる関数もまた，多項式サイズの PTF 表現を持つことを意味している．よって，$\oplus$MDNF 式 $f = f_1 \oplus \cdots \oplus f_d$ の深さ d が定数ならば，多項式サイズの PTF 表現を持つことが言える．もっと詳しく言えば，サイズ $O(n^d \cdot (\mathrm{size}_{\oplus\mathrm{MDNF}}(f))^{2d})$ の PTF で表される．ここで，$\mathrm{size}_{\oplus\mathrm{MDNF}}(f)$ は f の $\oplus$MDNF 表現のサイズとする．さて，ここで深さ 2（$d = 2$）の場合を考える．このとき上で述べた結果は，深さ 2 の $\oplus$MDNF 式 $f = f_1 \oplus f_2$ は，サイズ $O(n^2 \cdot (\mathrm{size}_{\oplus\mathrm{MDNF}}(f))^4)$ の PTF で表現可能であることを示している．本研究では，Jackson の結果を拡張することで，この上界を改善することができた．すなわち，深さ 2 の $\oplus$MDNF 式 f は，サイズ $O(n \cdot (\mathrm{size}_{\oplus\mathrm{MDNF}}(f))^4)$ の PTF で表現できることを示した．

第 5 章 DNF 式の充足数え上げ問題

第 5 章では，#P 完全問題に属し，現実的な時間内で解くことが難しいと考えられている，DNF 式の充足割り

当てを数え上げる問題について考えた．n 変数 DNF 式 F の項 $T_1,\cdots,T_m$ が与えられているとする．項 T_i を充足する入力変数の割り当ての集合を A_i とし，$S\subseteq\{1,\ldots,m\}$ に対して $a_S=|\bigcap_{i\in S}A_i|$ とする．ここで，$|A|$ は集合 A の元の個数を表す，このとき，DNF 式 $T_1\vee\cdots\vee T_m$ を充足する割り当ての個数 $|\bigcup_{i=1}^{m}A_i|$ は，包除原理の式

$$|\bigcup_{i=1}^{m}A_i|=\sum_{k=1}^{m}(-1)^{k-1}\left\{\sum_{S\subseteq\{1,\ldots,m\},|S|=k}a_S\right\}$$

から，厳密に求められる．しかし，この方法だと，m 項 DNF 式に対し，包除原理の式に現れる 2^m-1 項全ての計算が必要となるため，指数的な時間がかかり，かつ，一般には 1 つの項でも欠けると正確に求められない．これに対し，Kahn らは，F を充足する割り当ての個数を具体的に求める方法を与えたものではないが，$|S|\le\log n+1$ となる全ての S について a_S の値が分かれば，F を充足させる割り当ての個数が一意に定まることを示した．本研究では，この結果に基づき，各項に現れるリテラルの個数の最大値（以下，幅）を t に制限したとき，$|S|\le\log t+2$ である全ての S について a_S が分かれば，F を充足させる割り当てが一意に定まることを示した．さらに，幅を制限しない場合は $|S|\le\log n$ となる全ての S における a_S だけでは充足する個数を一意に定めれられないこと，また，幅を t に制限した場合は $|S|\le\log t+1(t\le n/2)$ となる全ての S における a_S だけでは充足する個数を一意に定めれられないことを示し，Kahn らの結果と，幅を t に制限した本研究の結果が，それぞれ最適であることを示した．

第 6 章 結論

第 6 章は，本論文のまとめであり，本研究で得られた結果とその意義について述べた．

論文審査の結果の要旨

ブール関数を表す表現形式にはブール式，ホーン式，決定木，決定リスト，決定ダイアグラムなどいろいろある．VLSIの設計から人工知能の分野に至る様々な問題は，ブール関数を用いて定式化されることが多く，その都度適当な表現形式が選ばれ，その表現形式のもとで議論されてきた．しかし，表現形式により問題を解くのに要する計算時間に指数関数的な違いが生じることがあるにもかかわらず，各表現形式の特性について系統的に論じられることは少なかった．本論文は，ブール関数の種々の表現形式について，2つの表現形式間の関係や別の表現形式への等価変換など，計算の複雑さの視点から系統的に論じるもので，全編6章よりなる．

第1章は序論である．

第2章では，ブール関数に関する基本的な事項及びその表現形式について述べている．

第3章では，単調DNF式を排他的論理和演算で結合した表現形式（以下$\oplus$MDNF式）について述べた後，ブール関数fとその否定がいずれもホーン式で表される関数のクラスと，各単調DNF式がちょうど1つの項からなる$\oplus$MDNF式のクラスが一致することを証明している．また，与えられたホーン式を$\oplus$MDNF式に変換する逐次多項式時間アルゴリズムを与えている．これは優れた成果である．

第4章では，ブール関数を多重線形多項式の符号で表す表現形式であるPTF表現を取り上げ，$\oplus$MDNF式をPTF表現で表したときのサイズの上界を評価するとともに，特に，$\oplus$MDNF式が2つの単調DNF式からなる場合について，その上界を改良している．

第5章では，与えられたDNF式の充足割り当ての個数を数えあげる問題を取りあげている．ここで，充足割り当てとは，ブール式を充足する変数への割り当てである．この問題は，現実的な時間では計算できないと予想されているが，DNF式の各項のサイズがt以下の場合，そのDNF式への充足割り当ての個数は，$\log t+2$個以下の項を同時に満たす割り当ての個数をすべて見ると一意に決まること，および，この値を$\log t+1$には改善できないことを示している．これは優れた成果である．

第6章は結論である．

以上要するに本論文は，ブール関数の表現形式について計算量の観点から検討し，DNF式や$\oplus$MDNF式の表現の複雑さに関する成果を得たもので，情報基礎科学並びに計算量理論の発展に寄与するところが少なくない．

よって，本論文は博士（情報科学）の学位論文として合格と認める．