ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Submodular goal value of Boolean functions

Eric Bach [a], Jérémie Dusart [b], Lisa Hellerstein [c,*], Devorah Kletenik [d]

[a] Computer Science Dept., University of Wisconsin, 1210 W. Dayton St., Madison, WI 53706, USA
[b] The Caesarea Rothschild Institute, University of Haifa, Mount Carmel, Haifa 31905, Israel
[c] Department of Computer Science and Engineering, NYU Tandon School of Engineering, 6 Metrotech Center, Brooklyn, NY 11201, USA
[d] Department of Computer and Information Science, Brooklyn College, City University of New York, 2900 Bedford Avenue, Brooklyn, NY 11210, USA

ARTICLE INFO

Article history:
Received 13 February 2017
Received in revised form 23 August 2017
Accepted 12 October 2017
Available online xxxx

Keywords:
Boolean functions
Submodularity
Read-once formulas

ABSTRACT

Recently, Deshpande et al. introduced a new measure of the complexity of a Boolean function. We call this measure the "goal value" of the function. The goal value of f is defined in terms of a monotone, submodular utility function associated with f. As shown by Deshpande et al., proving that a Boolean function f has small goal value can lead to a good approximation algorithm for the Stochastic Boolean Function Evaluation problem for f. Also, if f has small goal value, it indicates a close relationship between two other measures of the complexity of f, its average-case decision tree complexity and its average-case certificate complexity. In this paper, we explore the goal value measure in detail. We present bounds on the goal values of arbitrary and specific Boolean functions, and present results on properties of the measure. We compare the goal value measure to other, previously studied, measures of the complexity of Boolean functions. Finally, we discuss a number of open questions suggested by our work.

1. Introduction

Deshpande et al. introduced a new measure of the complexity of a Boolean function [6]. We call this measure the "goal value" of the function.[1] In this paper, we explore properties of the goal value measure.

We begin with some background. The goal value measure was introduced as part of an effort to develop approximation algorithms for the *Stochastic Boolean Function Evaluation* (SBFE) problem. In this problem, we are given a Boolean function $f(x_1, \ldots, x_n)$, and we need to evaluate it on an initially unknown input x. For example, consider a situation where each bit x_i of x corresponds to the result of a medical test, and f is a function of the test results, indicating whether the patient has a certain disease.

In addition to f, the input to the SBFE problem includes a cost c_i and probability p_i for each variable x_i. The cost c_i is the cost associated with determining the value of x_i, and p_i is the independent probability that x_i is 1. We may choose which bits to "buy" sequentially, and the choice of the next bit can depend on the choices of the previous bits. The task in the SBFE problem is to choose the sequence of bits so as to minimize the expected cost incurred before we determine $f(x)$.

Once we know all bits of x, we have enough information to determine $f(x)$. In fact, once we know enough bits that they form a certificate of f (i.e., the value of $f(x)$ is already determined by those bits alone), we also have enough information to determine $f(x)$. Thus any set of bits that forms a certificate has the same value to us. However, if the bits we know do not

* Corresponding author.
E-mail address: lisa.hellerstein@nyu.edu (L. Hellerstein).

1 Deshpande et al. called it the *Q-value*, but we have chosen to use a more intuitive name.

https://doi.org/10.1016/j.dam.2017.10.022
0166-218X/© 2017 Elsevier B.V. All rights reserved.

form a certificate, then we cannot determine $f(x)$. Nevertheless, we may still want to think of these bits as having some value towards our goal of determining $f(x)$. We can therefore define a (non-traditional) utility function that quantifies the value of knowing particular subsets of the bits, by assigning a non-negative integer value to each partial assignment. For example, the utility function might assign a value of 4 to knowing that $x_1 = 1$ and $x_3 = 0$. The maximum value of this utility function should be achieved precisely on those sets of bits that form a certificate. Let us call this value Q.

Once we have such a utility function, there is a natural greedy approach to choosing which bits to "buy": buy bits one by one, each time buying the bit that achieves the largest expected increase in utility value, per unit cost, until you can determine $f(x)$. In general, this greedy approach will not perform well. However, if the utility function has two useful properties, namely monotonity and submodularity, then a result of Deshpande et al. implies that the expected cost of the greedy approach is at most a factor of $O(\log Q)$ times the minimum possible expected cost.[2]

We will define monotonicity and submodularity formally below. Intuitively, monotonicity means that utility increases (or stays the same) as more bits are acquired. Submodularity is a diminishing returns property: if we consider the value of a bit to be the increase in utility that it would provide, then postponing the purchase of a particular bit cannot increase its value.

Call a utility function with the above properties a *goal function* for f. The *goal value* of a Boolean function f is the smallest possible value of Q, of any goal function for f.

Intuitively, we can think of a goal function for f as being a monotone, submodular surrogate for f. It has an expanded domain (including partial assignments to the variables) and an expanded range (the set $\{0, 1, \ldots, Q\}$, rather than just $\{0, 1\}$). The goal value of f measures how much we need to expand the range in order to achieve both monotonicity and submodularity. Deshpande et al. showed that, in addition to having implications for Boolean function evaluation, proving that a Boolean function f has small goal value indicates a close relationship between two other measures of the complexity of f: its average-case decision tree complexity and its average-case certificate complexity.

In this paper, however, our focus is not on the implications of goal value. Instead, we focus on understanding the fundamental properties of the goal value measure itself. We prove new bounds on the goal values of general and specific Boolean functions. We also consider the related 1-goal value and 0-goal value measures. Our results suggest a number of interesting open questions.

We begin by presenting necessary definitions and explaining the known connections between goal value, decision tree complexity, and Boolean function evaluation.

We then prove some generic properties of goal functions. While a single Boolean function can have different goal value functions, we show that two distinct Boolean functions can only have the same goal function if they are complements of each other. An interesting open question is whether each Boolean function has a unique *optimal* goal function.

We show that the goal value of every Boolean function f is at least n, if f depends on all n of its input variables. We note that this lower bound is tight for certain Boolean functions, such as AND, OR and XOR functions. We also present tight goal value bounds for Boolean k-of-n functions.

Deshpande et al. showed that the goal value of every Boolean function f is upper bounded by $\mathrm{ds}(f) \cdot \mathrm{cs}(f)$, where $\mathrm{ds}(f)$ is the minimum number of terms in a DNF formula for the function, and $\mathrm{cs}(f)$ is the minimum number of clauses in a CNF formula for the function [6]. We show that if f is a Boolean read-once function, then the goal value of f is actually equal to $\mathrm{ds}(f) \cdot \mathrm{cs}(f)$. Thus for read-once functions, there is a close relationship between goal value and the traditional complexity measures $\mathrm{ds}(f)$ and $\mathrm{cs}(f)$.

We show that the goal value of any Boolean function f is at most $2^n - 1$. We do not know how close goal values can come to this upper bound. Deshpande et al. showed that two specific Boolean functions have goal values lower bounded by $2^{n/2}$. Using our bound on the goal value of read-once functions, we show that there is a read-once function whose goal value is $3^{\frac{n}{3}}(\frac{n}{3})$, which is equal to $2^{\alpha n}(n/3)$, where $\alpha = \frac{\log_2 3}{3} \approx .528$.

We give an (exponential-sized) integer program for arbitrary Boolean function f whose solution is an optimal goal function for f.

2. Terminology and notation

Boolean function definitions: Let $V = \{x_1, \ldots, x_n\}$. A partial assignment is a vector $b \in \{0, 1, *\}^n$. For partial assignments $a, b \in \{0, 1, *\}^n$, a is an *extension* of b, written $a \succeq b$, if $a_i = b_i$ for all $b_i \neq *$. We also say that b is *contained* in a.

A literal is a Boolean variable x_i or its negation $\neg x_i$. A term is a conjunction ($\wedge$) of literals. A clause is a disjunction ($\vee$) of literals. A DNF formula is either the constant 0 or 1, or a formula $T_1 \vee T_2 \vee \ldots \vee T_k$ where each T_i is a term. A CNF formula is either the constant 0 or 1, or a formula $C_1 \wedge C_2 \wedge \cdots \wedge C_k$ where each C_i is a clause.

Given Boolean function $f : \{0, 1\}^n \to \{0, 1\}$, a partial assignment $b \in \{0, 1, *\}^n$ is a *0-certificate* of f if $f(a) = 0$ for all $a \in \{0, 1\}^n$ such that $a \succeq b$. It is a *1-certificate* if $f(a) = 1$ for all $a \in \{0, 1\}^n$ such that $a \succeq b$. It is a *certificate* if it is either a 0-certificate or a 1-certificate. We say that b *contains* a variable x_i if $b_i \neq *$. We will occasionally treat a certificate b as being the set of variables x_i such that $b_i \neq *$, together with the assignments given to them by b. The *size* of a certificate b is the

2 This follows from an $O(\log Q)$ approximation bound for Stochastic Submodular Cover, using the Adaptive Greedy algorithm of Golovin and Krause [11]. This bound was originally claimed by Golovin and Krause, but an error in their proof was recently found [1]. Deshpande et al. provided an alternative proof of the bound [6].

Please cite this article in press as: E. Bach, et al., Submodular goal value of Boolean functions, Discrete Applied Mathematics (2017), https://doi.org/10.1016/j.dam.2017.10.022.