$2k$-approximation algorithm for minimizing DNF representation of k-interval Boolean functions

Filip Bártek

February 26, 2015

1 Introduction

In the following text, an algorithm will be shown that computes a small DNF representation of Boolean function given as a set of k intervals. An approximation ratio of $2k$ will be proved.

2 Definitions

Definition 2.1 (k-interval Boolean function)**.** Let $a_1, b_1, \ldots, a_k, b_k$ be n-bit numbers such that $0 \leq a_1$, $a_1 \leq b_1$, $b_1 \leq a_2 - 2$, …, $b_k \leq 2^n - 1$. Then $f^n_{[a_1,b_1],\ldots,[a_k,b_k]} : \mathbf{B}^n \to \mathbf{B}$ is a function defined as follows:

$$f^n_{[a_1,b_1],\ldots,[a_k,b_k]}(x) = \begin{cases} 1 & \text{if } x \in [a_i, b_i] \text{ for some } i \\ 0 & \text{otherwise} \end{cases}$$

Note that the adjacent intervals are required to be separated by at least one false point.

3 Algorithm

3.1 Description

Input Numbers $a_1, b_1, \ldots, a_k, b_k$ that satisfy the inequalities in definition 2.1

Output A set of ternary vectors

Procedure The algorithm goes through all the intervals $[a_i, b_i]$. For each i, the longest common prefix of a_i and b_i is computed. Let j be its length. Note that $a^{[j+1]} = 0$ and $b^{[j+1]} = 1$. Now let $a'' = a_i^{[j+2,n]}$ and $b'' = b_i^{[j+2,n]}$. Optimally span the suffix interval $[a'', 1^{n-j-1}]$ and the prefix interval $[0^{n-j-1}, b'']$ using the (linear time) algorithm introduced in [2]. Prepend $a_i^{[1,j+1]}$ and $b_i^{[1,j+1]}$ to the respective ternary vectors and add them to the output spanning set.

3.2 Correctness

Theorem 3.2.1. *The algorithm spans exactly* $f^n_{[a_1,b_1],\ldots,[a_k,b_k]}$.

Proof. This is easy to see from the fact that the subintervals form a partition of the multi-interval (i.e. the set of all true points) and that each of them is spanned exactly by the suffix or prefix procedure. □

3.3 Approximation ratio

Theorem 3.3.1. *Let* $\mathcal{T}_{opt}$ *be an optimal spanning set of* $f^n_{[a_1,b_1],\ldots,[a_k,b_k]}$ *and let* $\mathcal{T}_{approx}$ *be the spanning set returned by the algorithm. We claim that:*

$$|\mathcal{T}_{approx}| \leq 2k|\mathcal{T}_{opt}| \tag{1}$$

Proof. Let $\mathcal{T}_x$ be the largest (n-bit) spanning set of a "suffix" or "prefix" subinterval added in the algorithm. Without loss of generality, let the respective subinterval be "prefix" $[b_i^{[1,j+1]}, b_i]$. From [1, p. 36] we know that there is an orthogonal set of $[b_i^{[1,j+1]}, b_i]$ of size $|\mathcal{T}_x|$, and moreover that its orthogonality only depends on the false point $b+1$. Note, however, that $b+1$ is also a false point in $f^n_{[a_1,b_1],\ldots,[a_k,b_k]}$. Thus we obtain an orthogonal set of size $|\mathcal{T}_x|$ for the k-interval function, limiting the size of its optimal spanning set $|\mathcal{T}_{opt}| \geq |\mathcal{T}_x|$.

Since $\mathcal{T}_x$ is the largest of the $2k$ partial sets used to span the function in the approximation algorithm, we know that $|\mathcal{T}_{approx}| \leq 2k|\mathcal{T}_x|$.

Joining the inequalities together we conclude: $|\mathcal{T}_{approx}| \leq 2k|\mathcal{T}_x| \leq 2k|\mathcal{T}_{opt}|$. □

References

[1] Jakub Dubovský. A construction of minimum DNF representations of 2-interval functions. Master's thesis, Charles University in Prague, 2012.

[2] Baruch Schieber, Daniel Geist, and Ayal Zaks. Computing the minimum DNF representation of boolean functions defined by intervals. *Discrete Applied Mathematics*, 149(1–3):154 – 173, 2005. Boolean and Pseudo-Boolean Functions Boolean and Pseudo-Boolean Functions.

Glossary

DNF disjunctive normal form. 1