Efficient Computation of Approximate Isomorphisms between Boolean Functions

Hans Ulrich Simon

November 26, 2015

Abstract

Arvind and Vasudev [2] have introduced the notion of an approximate isomorphism between two Boolean functions f and g. They furthermore designed two algorithms that construct an approximate isomorphism when given oracle access to f and g. The first of these algorithms is specialized to Boolean functions which are computable by constant-depth circuits. The second one applies to any pair of Boolean functions. It runs in exponential time and achieves optimality up to a factor of order $\sqrt{n}$. In this paper, we present an improved analysis and come up with a variant of the second algorithm that runs in polynomial time and achieves optimality up to a factor of (approximately) 2.

1 Introduction

Two Boolean functions are said to be isomorphic if they are equal up to a permutation of the variables. The problem of deciding whether two functions $f, g : \{0,1\}^n \to \{0,1\}$ are isomorphic is known to be coNP-hard even when f and g are given as DNF formulas. The problem is in Σ_2^p but it is not known to be in coNP. Agrawal and Thierauf [1] have shown that the problem is not Σ_2^p-hard unless the polynomial hierarchy collapses to Σ_3^p.

Arvind and Vasudev [2] have introduced the notion of approximate isomorphisms where the Boolean functions f, g are said to be $(1-\rho)$-approximately isomorphic if g, after applying an appropriate permutation to its variables, assigns the same binary labels as f on a fraction $\rho \in [0,1]$ of the Boolean domain. Clearly $\rho = 1$ means that f and g are (fully) isomorphic. The notion of an approximate isomorphism naturally leads to the following maximization problem: given oracle access to f and g, find a permutation of the variables of g which makes the agreement rate ρ with the function f as large as possible.

Arvind and Vasudev have designed two approximation algorithms. The first one is specialized to Boolean functions f, g which are computable by constant-depth circuits, say by circuits of depth d. It achieves the following: if f, g are (fully) isomorphic, then it runs in time $2^{\log^{O(d)}(n/\varepsilon)\sqrt{n}}$ and, with a probability of at least $1 - 2^{-\Omega(n)}$, it returns a permutation of the variables of g that leads to an agreement rate of $1 - \varepsilon$ with the function f. The second

algorithm in [2] applies to any pair of Boolean functions. It runs in exponential time and returns a permutation σ of the variables of g so that the resulting agreement rate with f differs from the optimal one at most by a factor of order $\sqrt{n}$. In this paper, we present an improved analysis and come up with an algorithm that runs in polynomial time and achieves optimality up to a factor of (approximately) 2. See Theorem 3.4 for the precise result.

The remainder of the paper is organized as follows. In Section 2, we specify the main problem more formally and fix some notations. In Section 3, we present our main results. In Section 4, we briefly discuss how small the agreement rate between f and g^σ can possibly become when the *worst* permutation σ is chosen for two (fully) isomorphic functions f and g.

2 The Problem MAX-BFI

The constants 0 and 1 are called *binary labels* or simply *labels* in this paper. For any Boolean function $g : \{0,1\}^n \to \{0,1\}$, any vector $x = (x[1], \ldots, x[n])$ and any permutation σ of $1, \ldots, n$, we introduce the notations

$$x^\sigma = (x[\sigma(1)], \ldots, x[\sigma(n)]) \quad \text{and } g^\sigma(x) = g(x^\sigma) \ .$$

Two Boolean functions $f, g : \{0,1\}^n \to \{0,1\}$ are said to be *isomorphic* if there exists a permutation σ such that $f = g^\sigma$. Boolean Function Isomorphism (BFI) is the problem of deciding whether two Boolean functions, given by oracle access[1], are isomorphic.

As in [2], we replace the strict notion of an isomorphism by a measure ranging over the interval $[0,1]$ that quantifies "how isomorphic" two Boolean functions are. To this end, we define

$$\rho(f,g) = 2^{-n} \cdot |\{x \in \{0,1\}^n : f(x) = g(x)\}| \quad \text{and } \rho^*(f,g) = \max_\sigma \rho(f, g^\sigma) \ .$$

We will refer to $\rho(f,g)$ as the *agreement rate* of f and g. Note that f and g are (fully) isomorphic iff $\rho^*(f,g) = 1$. In the sequel, we will discuss the following optimization problem:

MAX-BFI Given oracle access to $f, g : \{0,1\}^n \to \{0,1\}$, compute a maximizer σ^* of $\rho(f, g^\sigma)$, i.e. compute a permutation σ^* such that $\rho(f, g^{\sigma^*}) = \rho^*(f,g)$.

In the following section, we will present a randomized approximation algorithm for this problem.

Notational Conventions We denote by $\mathcal{S}_n$ the set of permutations of $1, \ldots, n$. The notation $s \in_R S$ for some finite set S means that s is chosen uniformly at random from S. Probabilities (resp. expected values) involving parameters $s \in_R S$ are then written in the form $\Pr_{s \in_R S}[\cdot]$ (resp. $\mathbb{E}_{s \in_R S}[\cdot]$).

[1] An oracle for $f : \{0,1\}^n \to \{0,1\}$ returns $f(x)$ when called on x

3 The Main Results

The key result in this section, Theorem 3.2 below, states that random permutations achieve, on the average, an agreement rate that differs from the optimal one by factor 1/2 only. But before we state and prove this formally, we discuss a very simple optimization problem in two real variables that will play a prominent role in the proof of the key result:

Lemma 3.1. *Let α, β be constants such that $0 \leq \alpha \leq 1-\beta \leq 1$. Let the function h be given by $h(a,b) = ab + (1-a)(1-b)$. Then,*

$$\left\{ \min_{a,b} h(a,b) \text{ **s.t.** } \alpha \leq a, b \leq 1-\beta \right\} \geq \frac{\alpha+\beta}{2} \quad . \tag{1}$$

Proof. For each fixed a, the function $h_a(b) = h(a,b)$ is linear in b. Thus $h_a(b)$ is monotonically increasing or monotonically decreasing. Hence $h_a(b)$ is minimized for some $b \in \{\alpha, 1-\beta\}$. For reasons of symmetry, the analogous remark is valid with the roles of a and b exchanged. Therefore at least one of the sets $\{(\alpha, 1-\beta), (1-\beta, \alpha)\}$ and $\{(\alpha,\alpha), (1-\beta, 1-\beta)\}$ must contain an optimal solution (a^*, b^*) of the minimization problem on the left-hand side in (1). In the former case,

$$h(a^*, b^*) = \alpha(1-\beta) + (1-\alpha)\beta = \alpha + \beta - 2\alpha\beta$$

whereas, in the latter case,

$$h(a^*, b^*) \geq \min\{\alpha^2 + (1-\alpha)^2, \beta^2 + (1-\beta)^2\} \quad .$$

Our assumptions on α and β imply that $\alpha, \beta \geq 0$ and $\alpha + \beta \leq 1$. Since the geometric mean is upper-bounded by the arithmetic mean, we have $\alpha\beta \leq (\alpha+\beta)^2/4 \leq (\alpha+\beta)/4$ and $\alpha(1-\alpha) \leq 1/4$. Thus $\alpha + \beta - 2\alpha\beta \geq (\alpha+\beta)/2$. Moreover,

$$\alpha^2 + (1-\alpha)^2 = 1 - 2\alpha(1-\alpha) \geq 1 - \frac{1}{2} = \frac{1}{2} \geq \frac{\alpha+\beta}{2}$$

and, for reasons of symmetry, the analogous inequality holds with the roles of α and β exchanged. In any case, one gets $h(a^*, b^*) \geq (\alpha+\beta)/2$, as desired. □

One can easily extend the proof of Lemma 3.1 and show that the optimal solution (a^*, b^*) is an element of $\{(\alpha, 1-\beta), (1-\beta, \alpha)\}$ if $\alpha, \beta \leq 1/2$, whereas it equals (α, α) (resp. (β, β)) if $\alpha > 1/2$ (resp. if $\beta > 1/2$). Since we do not need this extension in the sequel, we omit the details.

We are now well prepared for the key result in this section:

Theorem 3.2. *For each pair $f, g : \{0,1\}^n \to \{0,1\}$ of Boolean functions, we have*

$$\mathbb{E}_{\sigma \in_R \mathcal{S}_n}[\rho(f, g^\sigma)] \geq \frac{\rho^*(f,g)}{2} \quad .$$

Proof. For the sake of brevity, let $B = \{0,1\}^n$ and, for $i = 0, \ldots, n$, let $B_i \subseteq B$ be the subset consisting of all points with Hamming weight i. Let $s_i(f) = |B_i \cap f^{-1}(1)|$ denote the number of points in B_i to which f assigns the label 1. Let $X(u,\sigma)$ be the function that evaluates to 1 if $f(u) = g^\sigma(u) = g(u^\sigma)$ and that evaluates to 0 otherwise. If u is chosen uniformly at random from B_i and σ is chosen uniformly at random from $\mathcal{S}_n$, then (u, u^σ) is uniformly distributed over $B_i \times B_i$. From this and from the fact that $|B_i| = \binom{n}{i}$, it follows that

$$\mathbb{E}_{u \in_R B_i, \sigma \in_R \mathcal{S}_n}[X(u,\sigma)] = \frac{s_i(f)s_i(g)}{\binom{n}{i}^2} + \frac{\left(\binom{n}{i} - s_i(f)\right)\left(\binom{n}{i} - s_i(g)\right)}{\binom{n}{i}^2} \ .$$

Let $Y_i(\sigma)$ count the total number of agreements between f and g^σ on B_i, i.e., $Y_i(\sigma) = \sum_{u \in B_i} X(u,\sigma)$. Note that, for each $\sigma_0 \in \mathcal{S}_n$, we have

$$Y_i(\sigma_0) = \binom{n}{i} \cdot \Pr_{u \in B_i}[X(u,\sigma) = 1 | \sigma = \sigma_0] = \binom{n}{i} \cdot \mathbb{E}_{u \in B_i}[X(u,\sigma) | \sigma = \sigma_0] \ .$$

It follows that

$$\begin{aligned}
\mathbb{E}_{\sigma \in \mathcal{S}_n}[Y_i(\sigma)] &= \frac{1}{n!} \cdot \sum_{\sigma_0 \in \mathcal{S}_n} \mathbb{E}_{\sigma \in \mathcal{S}_n}[Y_i(\sigma) | \sigma = \sigma_0] \\
&= \frac{1}{n!} \cdot \sum_{\sigma_0 \in \mathcal{S}_n} Y_i(\sigma_0) \\
&= \binom{n}{i} \cdot \frac{1}{n!} \cdot \sum_{\sigma_0 \in \mathcal{S}_n} \mathbb{E}_{u \in B_i}[X(u,\sigma) | \sigma = \sigma_0] \\
&= \binom{n}{i} \cdot \mathbb{E}_{u \in B_i, \sigma \in \mathcal{S}_n}[X(u,\sigma)] \ .
\end{aligned}$$

Let now $Y(\sigma)$ count the total number of agreements between f and g^σ on B, i.e., $Y(\sigma) = \sum_{i=0}^n Y_i(\sigma)$. In the sequel, we show that $\mathbb{E}_{\sigma \in_R \mathcal{S}_n}[Y(\sigma)] \geq \rho^*(f,g)2^{n-1}$ (from which the theorem is immediate). We now bring Lemma 3.1 into play. Let $a_i = s_i(f)/\binom{n}{i}$ and $b_i = s_i(g)/\binom{n}{i}$ so that $\mathbb{E}_{u \in B_i, \sigma \in \mathcal{S}_n}[X(u,\sigma)] = a_i b_i + (1-a_i)(1-b_i)$. Let σ^* be the permutation such that $\rho^*(f,g) = \rho(f, g^{\sigma^*})$. Let $0 \leq \alpha_i \leq 1$ (resp. $0 \leq \beta_i \leq 1$) denote the fraction of points in B_i to which f and g^{σ^*} both assign label 1 (resp. label 0). Note that

$$\sum_{i=0}^n (\alpha_i + \beta_i)\binom{n}{i} = \rho^*(f,g)2^n \ .$$

Clearly $\alpha_i + \beta_i \leq 1$. Furthermore $a_i, b_i \geq \alpha_i$ and, symmetrically, we have $1 - a_i, 1 - b_i \geq \beta_i$ (because, otherwise, we could not get the rates α_i resp. β_i of agreement between f and g^{σ^*}). From Lemma 3.1 (applied to $\alpha_i, \beta_i, a_i, b_i$ in place of α, β, a, b), we conclude that $\mathbb{E}_{u \in_R B_i, \sigma \in_R \mathcal{S}_n}[X(u,\sigma)] = a_i b_i + (1-a_i)(1-b_i) \geq (\alpha_i + \beta_i)/2$. Thus $\mathbb{E}_{\sigma \in \mathcal{S}_n}[Y_i(\sigma)] \geq \binom{n}{i}(\alpha_i + \beta_i)/2$ and

$$\mathbb{E}_{\sigma \in \mathcal{S}_n}[Y(\sigma)] = \sum_{i=0}^n \mathbb{E}_{\sigma \in \mathcal{S}_n}[Y_i(\sigma)] \geq \sum_{i=0}^n \frac{\alpha_i + \beta_i}{2}\binom{n}{i} = \rho^*(f,g)2^{n-1} \ ,$$

which completes the proof of the theorem. □

Let σ denote a random permutation. The following result states that the random variable $\rho(f, g^\sigma)$ has a chance of at least $\varepsilon/2$ for not falling below $\rho^*(f, g)$ by more than factor $(1-\varepsilon)/2$:

Corollary 3.3. *For each pair $f, g : \{0,1\}^n \to \{0,1\}$ of Boolean functions and for all $0 \leq \varepsilon \leq 1$, the following holds:*

$$\Pr_{\sigma \in \mathcal{S}_n}\left[\rho(f, g^\sigma) > \frac{1-\varepsilon}{2} \cdot \rho^*(f, g)\right] \geq \frac{\varepsilon}{2} \ .$$

Proof. Let $0 \leq X \leq 2c$ be a random variable whose expectation is at least c. Solving

$$\begin{aligned} c \leq \mathbb{E}[X] &= \mathbb{E}[X | X > (1-\varepsilon)c] \cdot \Pr[X > (1-\varepsilon)c] + \\ &\quad \mathbb{E}[X | X \leq (1-\varepsilon)c] \cdot \Pr[X \leq (1-\varepsilon)c] \\ &\leq 2c \cdot \Pr[X > (1-\varepsilon)c] + (1-\varepsilon)c \end{aligned}$$

for $\Pr[X > (1-\varepsilon)c]$, we get $\Pr[X > (1-\varepsilon)c] \geq \varepsilon/2$. Since $\rho(f, g^\sigma)$ is a random variable with expectation at least $\rho^*(f, g)/2$ and with values ranging from 0 to $\rho^*(f, g)$, we may apply the above reasoning to $X = \rho(f, g^\sigma)$ and $c = \rho^*(f, g)/2$, which yields the corollary. □

Corollary 3.3 implies that a collection of random permutations is likely to contain a "good permutation" σ that satisfies the condition in the $\Pr[\cdot]$-expression of Corollary 3.3. In order to actually find a permutation whose performance is not much worse than the performance of the good one, we can use empirical estimates for the various ρ-values. Thus the following result should not come as a surprise:

Theorem 3.4. *There is an efficient randomized algorithm APPROX-ISO which, given input parameters $0 < \gamma, \varepsilon, \delta < 1$ and given oracle access to two Boolean functions $f, g : \{0,1\}^n \to \{0,1\}$, achieves the following. APPROX-ISO calls the oracles $O\left(\frac{\ln(1/\gamma)\ln(1/(\varepsilon\gamma))}{\varepsilon\delta^2}\right)$ times and, with a probability of at least $1-\gamma$, APPROX-ISO returns a permutation σ such that*

$$\rho(f, g^\sigma) \geq \frac{1-\varepsilon}{2}\rho^*(f, g) - \delta \ . \tag{2}$$

Proof. The algorithm APPROX-ISO does the following:

1. Set $m = \left\lceil \frac{2\ln(2/\gamma)}{\varepsilon} \right\rceil$ and draw m permutations, say $\sigma_1, \ldots, \sigma_m$, independently at random.
2. Set
$$m' = \left\lceil \frac{2}{\delta^2} \ln\left(\frac{4m}{\gamma}\right) \right\rceil = O\left(\frac{\ln(1/(\varepsilon\gamma))}{\delta^2}\right) \tag{3}$$
and draw m' points, say $x_1, \ldots, x_{m'}$ from $\{0,1\}^n$ independently at random.
3. Call the f-oracle on $x_1, \ldots, x_{m'}$. For every $j \in \{1, \ldots, m\}$ and every $j' \in \{1, \ldots, m'\}$, call the g-oracle on $x_{j'}^{\sigma_j}$.

4. For $j = 1, \ldots, m$, let $X_j = |\{j' \in \{1, \ldots, m'\} : \ f(x_{j'}) = g^{\sigma_j}(x_{j'})\}|$. Note that X_j is binomially distributed with parameters m' and $p_j = \rho(f, g^{\sigma_j})$ so that $\mathbb{E}[X_j] = p_j m' = \rho(f, g^{\sigma_j})m'$.

5. Return the empirically best performing permutation, i.e. return σ_k for an index k such that $X_k = \max_{j=1,\ldots,m} X_j$.

Clearly APPROX-ISO runs in polynomial time. We now prove that APPROX-ISO returns a good permutation with high probability. First observe that, according to Corollary 3.3, each $\sigma \in \{\sigma_1, \ldots, \sigma_m\}$ has a chance of at least $\varepsilon/2$ to be chosen such that

$$\rho(f, g^{\sigma}) > \frac{1-\varepsilon}{2} \cdot \rho^*(f, g) \ . \tag{4}$$

The above choice of m implies that $(1 - \varepsilon/2)^m < \exp(-\varepsilon m/2) \leq \gamma/2$. Thus, with a probability of at least $1 - \gamma/2$, condition (4) holds for at least one choice of $\sigma \in \{\sigma_1, \ldots, \sigma_m\}$, say for $\sigma = \sigma_{k^*}$. The well-known Chernoff-Hoeffding bound [3, 4] and the choice of m' in (3) imply that, for each fixed $j \in \{1, \ldots, m\}$, we have

$$\Pr[|X_j/m' - p_j| > \delta/2] \leq 2\exp(-2(\delta/2)^2 m') = 2\exp(-\delta^2 m'/2) \leq \gamma/(2m) \ .$$

It follows from the union bound that, with a probability of at least $1 - \gamma/2$, we have that

$$\forall j = 1, \ldots, m : |X_j/m' - p_j| \leq \delta/2 \ . \tag{5}$$

The algorithm APPROX-ISO does not necessarily return σ_{k^*} but it returns the "empirical champion" σ_k. Condition (5) implies that $p_{k^*} = \rho(f, g^{\sigma_{k^*}})$ can exceed $p_k = \rho(f, g^{\sigma_k})$ by at most δ. Noting that the number of oracle calls is $m' + mm'$ and putting everything together, the theorem follows. □

4 Existence of Bad Permutations

If the permutation σ is chosen by "nature" (at random), then, as we have seen in the previous section, the random variable $\rho(f, g^{\sigma})$ has a significant chance of being close to $\rho^*(f, g)/2$. What can we say if σ and f, g are chosen by the "devil"? In other words, do there exist triples (f, g, σ) such that σ is a terribly bad choice in its role of an approximate isomorphism between f and g? In this section, we briefly show that such triples exist, indeed. The key observations are as follows:

Lemma 4.1. *Let σ be a permutation of $1, \ldots, n$, and let G be the subgroup of the symmetric group that is generated by σ. If each orbit $G(x) = \{x, \sigma(x), \sigma^2(x), \ldots\}$ with $x \in \{0,1\}^n \setminus \{\vec{0}, \vec{1}\}$ has an even cardinality, then there exists a Boolean function $f : \{0,1\}^n \to \{0,1\}$ such that $\rho(f, f^{\sigma}) = 2^{-(n-1)}$ (whereas $\rho^*(f, f) = 1$).*

Proof. The Boolean domain decomposes into orbits. We define f orbitwise. As for the trivial orbits $G(\vec{0}) = \{\vec{0}\}$ and $G(\vec{1}) = \{\vec{1}\}$, we make an arbitrary choice, say $f(\vec{0}) = f(\vec{1}) = 0$. Consider now an arbitrary orbit $G(x)$ with x being different from $\vec{0}$ and from $\vec{1}$. By assumption $|G(x)|$ is even. For every $y = \sigma^k(x)$, we may therefore define $f(y) = 0$ iff k is even. It follows from this construction that f and f^σ totally disagree on any non-trivial orbit, i.e., they agree only on $\vec{0}$ and $\vec{1}$. Therefore $\rho(f, f^\sigma) = 2^{-(n-1)}$ which concludes the proof. □

We briefly note that the proof of Lemma 4.1 actually shows a stronger statement: for each permutation σ, there exists a function f such that $\rho(f, f^\sigma)$ equals 2^{-n} times the number of orbits (w.r.t. the group generated by σ) of odd cardinality.

The following result shows that the assumption imposed on σ in Lemma 4.1 can be satisfied:

Lemma 4.2. *Let n be a power of 2, and let G be the subgroup of the symmetric group that is generated by the cyclic permutation $\sigma = \langle 1\ 2\ \ldots\ n \rangle$. Then each orbit $G(x)$ with $x \in \{0,1\}^n \setminus \{\vec{0}, \vec{1}\}$ has an even cardinality.*

Proof. The order of the cyclic permutation σ is n. Alternatively, the order of σ can be specified as follows. It equals the least $k \geq 1$ such that every $x \in \{0,1\}^n$ is a fix point of σ^k. Let $k(x)$ denote the smallest $k \geq 1$ such that x is a fix point of σ^k. Thus the order of σ equals the least common multiple of the numbers in $\{k(x) : x \in \{0,1\}^n\}$. Clearly $k(x) = |G(x)|$. We may conclude that n, which is assumed to be a power of 2, equals the least common multiple of the cardinalities of the orbits. Thus, the cardinality of every orbit must be a power of 2. The assertion of the lemma is now immediate because the case $|G(x)| = 2^0 = 1$ can occur only for the vectors $x \in \{\vec{0}, \vec{1}\}$. □

Conclusions and Open Problems We have seen that there is a very simple randomized polynomial time approximation algorithm for MAX-BFI that differs from optimality not much more than by a factor of 1/2. This raises the question whether the factor 1/2 is a barrier that, perhaps, cannot be broken by any polynomial time approximation algorithm. Another straightforward question is whether we can find a *deterministic* polynomial-time approximation algorithm for MAX-BFI that achieves a reasonably good agreement rate.

References

[1] Manindra Agrawal and Thomas Thierauf. The formula isomorphism problem. *SIAM Journal on Computing*, 30(3):990–1009, 2000.

[2] Vikraman Arvind and Yadu Vasudev. Isomorphism testing of Boolean functions computable by constant-depth circuits. *Information and Computation*, 239:3–12, 2014.

[3] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.*, 23(4):493–507, 1952.

[4] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.