

Boolean Affine Approximation with Binary Decision Diagrams

Kevin Henshall

Peter Schachte

Harald Søndergaard

Leigh Whiting

Department of Computer Science and Software Engineering
The University of Melbourne, Vic. 3010, Australia

kevin.henshall@gmail.com
schachte@csse.unimelb.edu.au
harald@csse.unimelb.edu.au
leighwhiting@gmail.com

Abstract

Selman and Kautz's work on knowledge compilation has established how approximation (strengthening and/or weakening) of a propositional knowledge-base can be used to speed up query processing, at the expense of completeness. In the classical approach, the knowledge-base is assumed to be presented as a propositional formula in conjunctive normal form (CNF), and Horn functions are used to over- and under-approximate it (in the hope that many queries can be answered efficiently using the approximations only). However, other representations are possible, and functions other than Horn can be used for approximations, as long as they have deduction-computational properties similar to those of the Horn functions. Zanuttini has suggested that the class of affine Boolean functions would be especially useful in knowledge compilation and has presented various affine approximation algorithms. Since CNF is awkward for presenting affine functions, Zanuttini considers both a sets-of-models representation and the use of modulo 2 congruence equations. Here we consider the use of reduced ordered binary decision diagrams (ROBDDs), a representation which is more compact than the sets of models and which (unlike modulo 2 congruences) can express any source knowledge-base. We present an ROBDD algorithm to find strongest affine upper-approximations of a Boolean function and we argue its correctness.

1 Introduction

A recurrent theme in artificial intelligence is the efficient use of (propositional) knowledge-bases. A promising approach, which was initially proposed by Selman & Kautz (1996), is to query (and perform deductions from) upper and lower approximations of the given knowledge-base. By choosing approximations that allow more efficient inference, it is often possible to quickly determine that some logical consequence of the knowledge-base entails the query, and therefore so does the original knowledge-base, avoiding the costly inference from the original. When this fails, it may be possible to quickly show that the query is not entailed by some implicant, and therefore not entailed by the full knowledge-base. Only when both of these fail must the full knowledge-base be used for inference. This approach to deduction is particularly

attractive if the knowledge-base is relatively stable (that is, many queries are handled between each time the knowledge-base changes), because in that case, the amortised cost of calculating the approximations is small.

It is usually assumed that Boolean functions are represented in clausal form, and that approximations are Horn (Selman & Kautz 1996, del Val 2005). The reason is that inference from Horn knowledge-bases may be exponentially more efficient than from unrestricted knowledge-bases. However, it has been noted that there are many other well-understood classes that have computational properties that include some of the attractive properties of the Horn class.

Zanuttini (2002a, 2002b) discusses the use of other classes of Boolean functions for approximation and points out that *affine* approximations have certain advantages over Horn approximations, most notably the fact that they do not blow out in size. This is certainly the case when affine functions are represented in the form of modulo-2 congruence equations. The more general sets-of-models representation is also considered by Zanuttini. In this paper, we consider a third, general, representation, namely reduced ordered binary decision diagrams (ROBDDs). We prove some important properties of affine functions and their ROBDD representation. Utilising these properties we design a new ROBDD algorithm for deriving strongest affine consequences (also known as affine envelopes). Schachte and Søndergaard (2006, 2007) have previously given ROBDD algorithms for finding monotone, Krom, and Horn envelopes, but also noticed that while those algorithms could be expressed as instances of a common scheme, the same scheme did not apply to affine functions. A different, less compositional, approach is needed in this case.

The rest of this paper proceeds as follows. In Section 2 we recapitulate the definition of the Boolean affine class, and we establish some of its important properties. We also briefly introduce ROBDDs, but mainly to fix our notation, as we assume that the reader is familiar with Boolean functions and their representation as decision diagrams. Section 3 recalls the model-based affine envelope algorithm, and develops an ROBDD-based algorithm, whose correctness rests on results established in Section 2.2. Section 4 describes our testing methodology, including our algorithm for generating random ROBDDs, and presents our results. Section 5 discusses related work and applications, and concludes.

2 Propositional Classes, Approximation and ROBDDs

We use ROBDDs (Brace, Rudell & Bryant 1990, Bryant 1986) to represent Boolean functions.

Horiyama & Ibaraki (2002) have recommended ROBDDs as suitable for implementing knowledge bases. Our choice of ROBDDs as a data structure is due to the fact that it offers a canonical representation for any Boolean function—a representation that is highly suitable for inductive reasoning.

Zanuttini (2002a) suggests using modulo 2 congruence equations to represent affine Boolean functions, and proves a polynomial complexity bound for computing affine envelopes in this representation. However, using a specialised representation has a cost in implementation complexity where affine and non-affine Boolean functions must be used together. For example, the algorithm for evaluating whether one ROBDD entails another is very straightforward, whereas evaluating whether a set of congruence equations entails a Boolean function in some other representation would be more complicated. Similarly, systems which repeatedly construct an affine approximation, manipulate it as a general Boolean function, and then approximate the result again, have much simpler implementations with a single universal representation than with a specialised affine representation plus a universal representation. For our purposes, computing envelopes as ROBDDs permits us to use the same representation for approximation to many different Boolean classes. Additionally, ROBDD-based inference is fast, and in particular, checking whether a valuation is a model, or finding a model, of an n -place function given by an ROBDD requires a path traversal of length no more than n .

2.1 Boolean functions

Let $\mathcal{B} = \{0, 1\}$ and let \mathcal{V} be a denumerable set of variables. A *valuation* $\mu : \mathcal{V} \rightarrow \mathcal{B}$ is a (total) assignment of truth values to the variables in \mathcal{V} . Let $\mathcal{I} = \mathcal{V} \rightarrow \mathcal{B}$ denote the set of \mathcal{V} -valuations. A *partial valuation* $\mu : \mathcal{V} \rightarrow \mathcal{B} \cup \{\perp\}$ assigns truth values to some variables in \mathcal{V} , and \perp to others. Let $\mathcal{I}_p = \mathcal{V} \rightarrow \mathcal{B} \cup \{\perp\}$. We use the notation $\mu[x \mapsto i]$, where $x \in \mathcal{V}$ and $i \in \mathcal{B}$, to denote the valuation μ updated to map x to i , that is,

$$\mu[x \mapsto i](v) = \begin{cases} i & \text{if } v = x \\ \mu(v) & \text{otherwise} \end{cases}$$

A Boolean function over \mathcal{V} is a function $\varphi : \mathcal{I} \rightarrow \mathcal{B}$. We let \mathbf{B} denote the set of all Boolean functions over \mathcal{V} . The ordering on \mathcal{B} is the usual: $x \leq y$ iff $x = 0 \vee y = 1$. \mathbf{B} is ordered pointwise, so that the ordering relation corresponds exactly to classical entailment, \models . It is convenient to overload the symbols for truth and falsehood. Thus we let 1 denote the largest element of \mathbf{B} (that is, $\lambda\mu.1$) as well as of \mathcal{B} . Similarly 0 denotes the smallest element of \mathbf{B} (that is, $\lambda\mu.0$) as well as of \mathcal{B} . A valuation μ is a *model* for φ , denoted $\mu \models \varphi$, if $\varphi(\mu) = 1$. We let $models(\varphi)$ denote the set of models of φ . Conversely, the unique Boolean function that has exactly the set M as models is denoted $fn(M)$. A Boolean function φ is said to be *independent* of a variable x when for all valuations μ , $\mu[x \mapsto 0] \models \varphi$ iff $\mu[x \mapsto 1] \models \varphi$.

Existential quantification is defined as follows. Let φ be a Boolean function and $M = models(\varphi)$, then

$$\exists v(\varphi) = fn(\{\mu[v \mapsto 0] \mid \mu \in M\} \cup \{\mu[v \mapsto 1] \mid \mu \in M\})$$

and universal quantification can be defined dually (we will not need it here). Clearly $\exists v(\varphi)$ is independent of v .

In the context of an ordered set of n variables of interest, x_1, \dots, x_n , we may identify with μ the binary

sequence $bits(\mu)$ of length n :

$$\mu(x_1), \dots, \mu(x_n)$$

which we will write simply as a bit-string of length n . Similarly we may think of, and write, the set of valuations M as a set of bit-strings:

$$bits(M) = \{bits(\mu) \mid \mu \in M\}$$

As it could hardly create confusion, we shall present valuations variously as functions or bitstrings. We denote the *zero valuation*, which maps x_i to 0 for all $1 \leq i \leq n$, by $\vec{0}$.

We use the Boolean connectives \neg (negation), \wedge (conjunction), \vee (disjunction) and \oplus (exclusive or, or “xor”). These connectives operate on Boolean functions, that is, on elements of \mathbf{B} . Traditionally they are overloaded to also operate on truth values, that is, elements of \mathcal{B} . However, we deviate at this point, as the distinction between xor and its “bit-wise” analogue will be critical in what follows. Hence we denote the \mathcal{B} (bit) version by \oplus . We extend this to valuations and bit-strings in the natural way:

$$(\mu_1 \oplus \mu_2)(x) = \mu_1(x) \oplus \mu_2(x)$$

and we let \oplus_3 denote the “xor of three” operation $\lambda\mu_1\mu_2\mu_3.\mu_1 \oplus \mu_2 \oplus \mu_3$. We follow Zanuttini (2002a) in further overloading ‘ \oplus ’ and using the notation

$$M_\mu = \mu \oplus M = \{\mu \oplus \mu' \mid \mu' \in M\}$$

We read M_μ as “ M translated by μ ”. Note that for any set M , the function $\lambda\mu.M_\mu$ is an involution: $(M_\mu)_\mu = M$.

A final overloading results in the following definition. For $\varphi \in \mathbf{B}$, and $\mu \in \mathcal{I}$, let $\varphi \oplus \mu = fn(M_\mu)$ where $M = models(\varphi)$.

The distributed version \bigoplus of \oplus is defined by $\bigoplus\{\varphi_1, \dots, \varphi_n\} = \varphi_1 \oplus \dots \oplus \varphi_n$.

2.2 The affine class

An *affine* function is one whose set of models is closed under pointwise application of \oplus_3 (Schaefer 1978). Affine functions have a number of attractive properties, as we shall see. Syntactically, a Boolean function is affine iff it can be written as a conjunction of affine equations

$$c_1x_1 + c_2x_2 + \dots + c_nx_n = c_0$$

where $c_i \in \{0, 1\}$ for all $i \in \{0, \dots, n\}$.¹ This is well known, but for completeness we prove it below, as Proposition 2.2.

The affine class contains 1 and is closed under conjunction. Hence the concept of a unique best affine upper-approximation is well defined, and the function that takes a Boolean function and returns its best affine upper-approximation is an upper closure operator (Ward 1942, Ore 1943). For convenience, let us introduce a name for this operator:

Definition 2.1 Let φ be a Boolean function. The *affine envelope*, $aff(\varphi)$, of φ is defined:

$$aff(\varphi) = \bigwedge \{\psi \mid \varphi \models \psi \text{ and } \psi \text{ is affine}\} \quad \blacksquare$$

¹In some circles, such as the cryptography/coding community, the term “affine” is used only for a function that is 0 or 1, or can be written $c_1x_1 + c_2x_2 + \dots + c_nx_n + c_0$ (the latter is what Post (1941) called an “alternating” function). The resulting set of “affine” functions is not closed under conjunction.

There are numerous other classes of interest, including isotone, antitone, Krom, Horn, contra-dual Horn, k -Horn (Dechter & Pearl 1992), and k -quasi-Horn functions, for which the concept of an envelope is well-defined, as each class is closed under conjunction.²

Zanuttini (2002a) exploits the close connection between vector spaces and the sets of models of affine functions. A set $S \subseteq \mathcal{B}^k$ of bitstrings is a *vector space* iff $\vec{0} \in S$ and S is closed under \oplus . The next proposition suggests how one can simplify the task of doing model-closure under \oplus_3 .

Proposition 2.1 (Zanuttini 2002a) Given a non-empty set of models M and a valuation $\mu \in M$, M is closed under \oplus_3 iff M_μ is a vector space.

Proof: Let μ be an arbitrary element of M . Clearly M_μ contains $\vec{0}$, so the right-hand side of the claim amounts to M_μ being closed under \oplus .

For the ‘if’ direction, assume M_μ is closed under \oplus and consider $\mu_1, \mu_2, \mu_3 \in M$. Since $\mu \oplus \mu_2$ and $\mu \oplus \mu_3$ are in M_μ , so is $\mu_2 \oplus \mu_3$. And since furthermore $\mu \oplus \mu_1$ is in M_μ , so is $\mu \oplus \mu_1 \oplus \mu_2 \oplus \mu_3$. Hence $\mu_1 \oplus \mu_2 \oplus \mu_3$ is in M .

For the ‘only if’ direction, assume M is closed under \oplus_3 , and consider $\mu_1, \mu_2 \in M_\mu$. All of $\mu, \mu \oplus \mu_1$ and $\mu \oplus \mu_2$ are in M , and so $\mu \oplus (\mu \oplus \mu_1) \oplus (\mu \oplus \mu_2) = \mu \oplus \mu_1 \oplus \mu_2 \in M$. Hence $\mu_1 \oplus \mu_2 \in M_\mu$. ■

Proposition 2.2 A Boolean function is affine iff it can be written as a conjunction of equations

$$c_1x_1 + c_2x_2 + \dots + c_nx_n = c_0$$

where $c_i \in \{0, 1\}$ for all $i \in \{0, \dots, n\}$.

Proof: Assume the Boolean function φ is given as a conjunction of equations of the indicated form and let μ_1, μ_2 and μ_3 be models. That is, for each equation we have

$$\begin{aligned} c_1\mu_1(x_1) + c_2\mu_1(x_2) + \dots + c_n\mu_1(x_n) &= c_0 \\ c_1\mu_2(x_1) + c_2\mu_2(x_2) + \dots + c_n\mu_2(x_n) &= c_0 \\ c_1\mu_3(x_1) + c_2\mu_3(x_2) + \dots + c_n\mu_3(x_n) &= c_0 \end{aligned}$$

Adding left-hand sides and adding right-hand sides, making use of the fact that ‘ \cdot ’ distributes over ‘ $+$ ’, we get

$$c_1\mu(x_1) + c_2\mu(x_2) + \dots + c_n\mu(x_n) = c_0 + c_0 + c_0 = c_0$$

where $\mu = \mu_1 \oplus \mu_2 \oplus \mu_3$. As μ thus satisfies each equation, μ is a model of φ . This establishes the ‘if’ direction.

For the ‘only if’ part, note that by Proposition 2.1, we obtain a vector space M_μ from any non-empty set M closed under \oplus_3 by translating each element of M by $\mu \in M$. Now form a basis B for M_μ by taking one non- $\vec{0}$ vector at a time from M_μ and adding it to the set of basis vectors collected so far iff it is linearly independent of that set. Let $j = n - |B|$ (note that $0 \leq j \leq n$). B can be extended to a basis for \mathcal{B}^n by bringing B (read as a $|B| \times n$ matrix) into echelon form and adding j vectors $V = \{\vec{v}_1, \dots, \vec{v}_j\}$ (these

can be chosen from the natural basis for \mathcal{B}^n). From B and V we can compute a set of j linear equations

$$\begin{aligned} a_{11}x_1 \oplus \dots \oplus a_{1n}x_n &= 0 \\ a_{21}x_1 \oplus \dots \oplus a_{2n}x_n &= 0 \\ &\vdots \\ a_{j1}x_1 \oplus \dots \oplus a_{jn}x_n &= 0 \end{aligned} \quad (1)$$

that have exactly M_μ as their set of models. For each $i \in \{1, \dots, j\}$, the coefficients $\vec{a}_i = (a_{i1}, \dots, a_{in})$ are uniquely determined by the set of n equations

$$\begin{aligned} \vec{a}_i \cdot \vec{v}_i &= 1 \\ \vec{a}_i \cdot \vec{v}_k &= 0 & 1 \leq k \leq j, k \neq i \\ \vec{a}_i \cdot \vec{b} &= 0 & \vec{b} \in B \end{aligned}$$

This construction guarantees that $\vec{x} = (x_1, \dots, x_n)$ satisfies the conjunction of equations (1) iff \vec{x} is in the span of B (that is, in M_μ). Each function

$$f_i = \lambda \vec{x} \cdot \vec{a}_i \cdot \vec{x}$$

is linear, so for $\nu \in M_\mu$,

$$f_i(\nu \oplus \mu) = f_i(\nu) + f_i(\mu) = f_i(\mu)$$

Hence M can be described by the set of j affine equations

$$\begin{aligned} a_{11}x_1 \oplus \dots \oplus a_{1n}x_n &= f_1(\mu) \\ a_{21}x_1 \oplus \dots \oplus a_{2n}x_n &= f_2(\mu) \\ &\vdots \\ a_{j1}x_1 \oplus \dots \oplus a_{jn}x_n &= f_j(\mu) \end{aligned}$$

as desired. ■

Example 2.1 In \mathcal{B}^4 , the set of models $M = \{0100, 0111, 1001, 1010\}$ is closed under \oplus_3 and so determines an affine function. Choosing $\mu = 0100$ as translation, we have $M_\mu = \{0000, 0011, 1101, 1110\}$. One basis for M_μ is $\{0011, 1101\}$, which can be extended to a basis for \mathcal{B}^4 by adding $V = \{0100, 0001\}$. Hence M_μ can be described by the conjunction

$$\begin{aligned} a_{11}x_1 \oplus \dots \oplus a_{14}x_4 &= 0 \\ a_{21}x_1 \oplus \dots \oplus a_{24}x_4 &= 0 \end{aligned}$$

where the coefficients are determined by solving

$$\begin{aligned} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

In other words, M_μ is described by

$$\begin{aligned} x_1 \oplus x_2 &= 0 \\ x_1 \oplus x_3 \oplus x_4 &= 0 \end{aligned}$$

In the case of $(x_1, x_2, x_3, x_4) = \mu = (0, 1, 0, 0)$, the left-hand sides evaluate to 1 and 0, respectively. Hence M is described by

$$\begin{aligned} x_1 \oplus x_2 &= 1 \\ x_1 \oplus x_3 \oplus x_4 &= 0 \quad \blacksquare \end{aligned}$$

²Other classes that are commonly considered in AI are not closed under conjunction and therefore do not have well-defined concepts of (unique) envelopes. Examples are the *unate* functions (a unate function is one that can be turned into an isotone function by systematic negation of zero or more variables) and the *renamable Horn* functions (a renamable Horn function is similarly one that can be turned into a Horn function by systematic negation of zero or more variables). For example, $x \rightarrow y$ and $x \leftarrow y$ both are unate, while $x \leftrightarrow y$ is not, so the ‘unate envelope’ of the latter is not well-defined.

Zanuttini (2002a) shows that the complexity of generating the equational form from an affine function's set of models is $\mathcal{O}(n^4)$.

It follows from the syntactic characterisation that the number of models possessed by an affine function is either 0 or a power of 2. Other properties will now be established that are used in the justification of the affine envelope algorithm of Section 3.

The first property is that if a Boolean function φ has two models that differ for exactly one variable v , then its affine envelope is independent of v . To state this precisely we introduce a concept of a ‘‘characteristic’’ valuation for a variable.

Definition 2.2 In the context of a set of variables V , let $v \in V$. The *characteristic valuation* for v , χ_v , is defined by

$$\chi_v(x) = \begin{cases} 1 & \text{if } x = v \\ 0 & \text{otherwise} \quad \blacksquare \end{cases}$$

Note that $\mu \oplus \chi_v$ is the valuation which agrees with μ for all variables except v . Moreover, if $\mu \models \varphi$, then both of μ and $\mu \oplus \chi_v$ are models of $\exists v(\varphi)$.

Proposition 2.3 Let φ be a Boolean function whose set of models forms a vector space, and assume that for some valuation μ and some variable v , μ and $\mu \oplus \chi_v$ both satisfy φ . Then φ is independent of v .

Proof: The set M of models contains at least two elements, and since it is closed under \oplus , χ_v is a model. Hence for every model ν of φ , $\nu \oplus \chi_v$ is another model. It follows that φ is independent of v . \blacksquare

Proposition 2.4 Let φ be a Boolean function. If, for some valuation μ and some variable v , μ and $\mu \oplus \chi_v$ both satisfy φ , then $\text{aff}(\varphi) = \exists v(\text{aff}(\varphi))$.

Proof: Let μ be a model of φ , with $\mu \oplus \chi_v$ also a model. For every model ν of φ , we have that $\nu \oplus \mu \oplus (\mu \oplus \chi_v)$ satisfies $\text{aff}(\varphi)$, that is, $\nu \oplus \chi_v \models \text{aff}(\varphi)$. Now since both ν and $\nu \oplus \chi_v$ satisfy $\text{aff}(\varphi)$, it follows that $\exists v(\text{aff}(\varphi))$ cannot have a model that is not already a model of $\text{aff}(\varphi)$ (and the converse holds trivially). Hence $\text{aff}(\varphi) = \exists v(\text{aff}(\varphi))$. \blacksquare

Lemma 2.5 Let φ be a satisfiable Boolean function with set M of models, and let $\mu \models \text{aff}(\varphi)$. Then there is an odd positive integer k and a subset M' of M , such that $|M'| = k$ and $\mu = \bigoplus M'$.

Proof: Define

$$\begin{aligned} M_0 &= M \\ M_i &= M_{i-1} \cup \{\mu_1 \oplus \mu_2 \oplus \mu_3 \mid \mu_1, \mu_2, \mu_3 \in M_{i-1}\} \end{aligned}$$

for $i > 0$. Then $\{M_i\}_{i \geq 0}$ is an increasing sequence of sets of models, stabilising in a finite number of steps, that is, for some non-negative j ,

$$M_i = M_j = \text{models}(\text{aff}(\varphi))$$

for all $i \geq j$.

An induction on i now shows that for all i and all $\mu \in M_i$, μ can be written as a sum $\bigoplus M'$ of an odd number of models of $M_0 = M$ (‘‘odd plus odd plus odd equals odd’’). In particular this holds for μ in M_j , that is, for each model of $\text{aff}(\varphi)$. \blacksquare

Proposition 2.6 For all Boolean functions φ and variables v , $\text{aff}(\exists v(\varphi)) = \exists v(\text{aff}(\varphi))$.

Proof: We need to show that the models of $\text{aff}(\exists v(\varphi))$ are exactly the models of $\exists v(\text{aff}(\varphi))$. Clearly $\text{aff}(\exists v(\varphi))$ is 0 iff φ is 0 iff $\exists v(\text{aff}(\varphi))$ is 0. So we can assume that $\text{aff}(\exists v(\varphi))$ is satisfiable—let $\mu \models \text{aff}(\exists v(\varphi))$. By Lemma 2.5, for some positive odd number k ,

$$\mu = \mu_1 \oplus \mu_2 \oplus \cdots \oplus \mu_k$$

with μ_1, \dots, μ_k being different models of $\exists v(\varphi)$. These k models can be partitioned into two sets, according to whether they satisfy φ ; let

$$\begin{aligned} M &= \{\mu_i \mid 1 \leq i \leq k, \mu_i \models \varphi\} \\ M' &= \{\mu_i \mid 1 \leq i \leq k, \mu_i \not\models \varphi\} \end{aligned}$$

Then both M and M'_v consist entirely of models of φ . Hence, depending on the parity of M' 's cardinality, either μ or $\mu \oplus \chi_v$ is a model of $\text{aff}(\varphi)$ (or both are). In either case, $\mu \models \exists v(\text{aff}(\varphi))$.

Conversely, let $\mu \models \exists v(\text{aff}(\varphi))$. Then either μ or $\mu \oplus \chi_v$ is a model of $\text{aff}(\varphi)$ (or both are). Hence μ (or $\mu \oplus \chi_v$ as the case may be) can be written as a sum of k models μ_1, \dots, μ_k (k odd) of φ . It follows that both $\mu_1 \oplus \mu_2 \oplus \cdots \oplus \mu_k$ and $\mu_1 \oplus \mu_2 \oplus \cdots \oplus \mu_k \oplus \chi_v$ are models of $\exists v(\varphi)$. Hence $\mu \models \text{aff}(\exists v(\varphi))$. \blacksquare

These propositions are important because they allow an aggressive approach to the elimination of variables in an affine envelope algorithm. It should be noted that both aff and $\exists v$ are upper closure operators, but there was no *a priori* reason to assume that they commute (Ore 1943). Indeed, there are natural classes of Boolean functions for which envelopes are well-defined, but where approximation into the class does not commute with existential quantification. As an example take the class of positive functions. A function is *positive* iff it evaluates to 1 when all variables are 1. This class is closed under conjunction, so we can define $\text{pos}(\varphi)$ to be the positive envelope of φ . The reader can now verify that in \mathcal{B}^2 , for example,

$$\begin{aligned} \text{pos}(\exists x(\neg x \wedge \neg y)) &= \text{pos}(\neg y) = x \vee \neg y \\ &\neq 1 = \exists x(x \leftrightarrow y) = \exists x(\text{pos}(\neg x \wedge \neg y)) \end{aligned}$$

Hence approximation and variable elimination do not commute in this case.

2.3 ROBDDs

We briefly recall the essentials of ROBDDs (Bryant 1992). Let the set \mathcal{V} of propositional variables be equipped with a total ordering \prec . *Binary decision diagrams (BDDs)* are defined inductively as follows:

- 0 is a BDD.
- 1 is a BDD.
- If $x \in \mathcal{V}$ and R_1 and R_2 are BDDs then $\text{ite}(x, R_1, R_2)$ is a BDD.

Let $R = \text{ite}(x, R_1, R_2)$. We say a BDD R' *appears in* R iff $R' = R$ or R' appears in R_1 or R_2 . We define $\text{vars}(R) = \{v \mid \text{ite}(v, -, -) \text{ appears in } R\}$. The meaning of a BDD is given as follows.

$$\begin{aligned} \llbracket 0 \rrbracket &= 0 \\ \llbracket 1 \rrbracket &= 1 \\ \llbracket \text{ite}(x, R_1, R_2) \rrbracket &= (x \wedge \llbracket R_1 \rrbracket) \vee (\bar{x} \wedge \llbracket R_2 \rrbracket) \end{aligned}$$

A BDD is an *Ordered binary decision diagram (OBDD)* iff it is 0 or 1 or if it is $\text{ite}(x, R_1, R_2)$, R_1 and R_2 are OBDDs, and $\forall x' \in \text{vars}(R_1) \cup \text{vars}(R_2) : x \prec x'$.

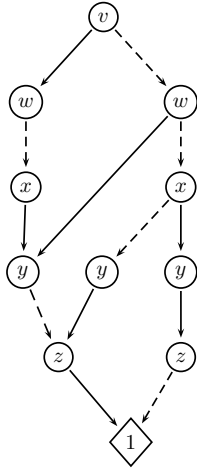


Figure 1: An example of our diagrammatic representation of an ROBDD. Our diagrams leave out the 0 sink and all arcs to it.

An OBDD R is a *Reduced Ordered Binary Decision Diagram* (ROBDD (Bryant 1986, Bryant 1992)) iff for all BDDs R_1 and R_2 appearing in R , $R_1 = R_2$ when $\llbracket R_1 \rrbracket = \llbracket R_2 \rrbracket$. Practical implementations (Brace et al. 1990) use a function $\text{mknd}(x, R_1, R_2)$ to create all ROBDD nodes as follows:

1. If $R_1 = R_2$, return R_1 instead of a new node, as $\llbracket \text{ite}(x, R_1, R_2) \rrbracket = \llbracket R_1 \rrbracket$.
2. If an identical ROBDD was previously built, return that one instead of a new one; this is accomplished by keeping a hash table, called the *unique table*, of all previously created nodes.
3. Otherwise, return $\text{ite}(x, R_1, R_2)$.

This ensures that ROBDDs are strongly canonical: a shallow equality test is sufficient to determine whether two ROBDDs represent the same Boolean function.

Figure 1 shows an example of an ROBDD. In general we depict the ROBDD $\text{ite}(x, R_1, R_2)$ as a directed acyclic graph rooted in x , with a solid arc from x to the dag for R_1 and a dashed line from x to the dag for R_2 . However, to avoid unnecessary clutter, we omit the 0 node (sink) and all arcs leading to that sink. The ROBDD in Figure 1 denotes the function which has five models: $\{00011, 00110, 01001, 01101, 10101\}$.

As a typical example of an ROBDD algorithm, Algorithm 1 generates the disjunction of two given ROBDDs. We present algorithms in Haskell style, using pattern matching and guarded equations. This operation will be used by the affine approximation algorithm presented in Section 3.

Algorithm 2 is used to extract a model from an ROBDD. For an unsatisfiable ROBDD (that is, 0) we return \perp . Although presented here in recursive fashion, it is better implemented in an iterative manner whereby we traverse through the ROBDD, one pointer moving down the “else” branch at each node, a second pointer trailing immediately behind. If a 1 sink is found, we return the path traversed thus far and note that any further variables which we are yet to encounter may be assigned any value. If a 0 sink is found, we use the trailing pointer to step up a level, follow the “then” branch for one step and continue searching for a model by following “else” branches. This method relies on the fact that ROBDDs are “reduced”, so that if no 1 sink can be reached from a node, then the node itself is the 0 sink.

$$\mu = 01100$$

$$M = \begin{Bmatrix} 01011 \\ 01100 \\ 10111 \\ 11001 \end{Bmatrix} \quad M_\mu = \begin{Bmatrix} 00111 \\ 00000 \\ 11011 \\ 10101 \end{Bmatrix}$$

$$N = \begin{Bmatrix} 00111 \\ 00000 \\ 11011 \\ 10101 \\ 11100 \\ 10010 \\ 01110 \\ 01001 \end{Bmatrix} \quad N_\mu = \text{aff}(M) = \begin{Bmatrix} 01011 \\ 01100 \\ 10111 \\ 11001 \\ 10000 \\ 11110 \\ 00010 \\ 00101 \end{Bmatrix}$$

Figure 2: Steps in Algorithm 3

We shall later use the following obvious corollary of Proposition 2.3:

Corollary 2.7 Let ROBDD R represent a function whose set of models form a vector space. Then every path from R ’s root node to the 1 sink contains the same sequence of variables, namely $\text{vars}(R)$ listed in variable order. ■

3 Finding Affine Envelopes for ROBDDs

Zanuttini (2002a) gives an algorithm, here presented as Algorithm 3, for finding the affine envelope, assuming a Boolean function φ is represented as a set of models. This algorithm is justified by Proposition 2.1.

Example 3.1 To see Algorithm 3 in action, assume that φ has four models, $M = \{01011, 01100, 10111, 11001\}$, and refer to Figure 2. We randomly pick $\mu = 01100$ and obtain M_μ as shown. The first round of completion under ‘ \oplus ’ adds three bit-strings: $\{11100, 10010, 01110\}$, and another round adds 01001 to produce N . Finally, “adding back” $\mu = 01100$ yields the affine envelope $N_\mu = \text{aff}(M)$. ■

We are interested in developing an algorithm for ROBDDs. We can improve on Algorithm 3 and at the same time make it more suitable for ROBDD manipulation. The idea is to build the result N step by step, by picking the models ν of M_μ one at a time and computing $N := N \cup N_\nu$ at each step. We can start from $N = \{\vec{0}\}$, as $\vec{0}$ has to be in M_μ . This leads to Algorithm 4.

This formulation is well suited to ROBDDs, as the operation N_ν , that is, taking the xor of a model ν with each model of the ROBDD N can be implemented by traversing N and, for each v -node with $\nu(v) = 1$, swapping that node’s children. And we can do better, utilising two observations.

First, during its construction, there is no need to traverse the ROBDD N for each individual model ν . A full traversal of N will find all its models systematically, eliminating a need to remove them one by one.

Second, the ROBDD being constructed can be simplified aggressively during its construction, by utilising Propositions 2.4 and 2.6. Namely, as we traverse ROBDD R systematically, many paths from

Algorithm 1 The “or” operator for ROBDDs

```
or(1, _) = 1
or(0, _) = 0
or(_, 1) = 1
or(_, 0) = 0
or(ite(x, T, E), ite(x', T', E'))
  | x < x' = mknd(x, or(T, ite(x', T', E')), or(E, ite(x', T', E')))
  | x' < x = mknd(x', or(ite(x, T, E), T'), or(ite(x, T, E), E'))
  | otherwise = mknd(x, or(T, T'), or(E, E'))
```

Algorithm 2 get_model algorithm for ROBDDs

```
get_model(0) = ⊥
get_model(1) = λv.⊥
get_model(ite(x, T, E)) =
  let μ = get_model(T)
  in
    if μ = ⊥ then
      get_model(E)[x ↦ 0]
    else μ[x ↦ 1]
```

Algorithm 3 The sets-of-models based affine envelope algorithm

```
Input: The set  $M$  of models for function  $\varphi$ .
Output:  $\text{aff}(M)$  — the set of models of  $\varphi$ 's affine envelope.
if  $M = \emptyset$  then
  return  $M$ 
end if
 $N \leftarrow \emptyset$ 
choose  $\mu \in M$ 
 $New \leftarrow M_\mu$ 
repeat
   $N \leftarrow N \cup New$ 
   $New \leftarrow \{\mu_1 \oplus \mu_2 \mid \mu_1, \mu_2 \in N\} \setminus N$ 
until  $New = \emptyset$ 
return  $N_\mu$ 
```

Algorithm 4 A variant of Algorithm 3

```
Input: The set  $M$  of models for function  $\varphi$ .
Output:  $\text{aff}(M)$  — the set of models of  $\varphi$ 's affine envelope.
if  $M = \emptyset$  then
  return  $M$ 
end if
 $N \leftarrow \{\vec{0}\}$ 
choose  $\mu \in M$ 
 $R \leftarrow M_\mu \setminus \{\vec{0}\}$ 
for all  $\nu \in R$  do
   $N \leftarrow N \cup N_\nu$ 
end for
return  $N_\mu$ 
```

the root to the 1 sink will be found that do not contain every variable in $\text{vars}(R)$. Each such path corresponds to a model *set* of cardinality 2^k , k being the number of “skipped” variables. Proposition 2.4 tells us that, eventually, the affine envelope will be independent of all such “skipped” variables, and Proposition 2.6 guarantees that variable elimination can be interspersed arbitrarily with the process of “xor-ing” models, that is, we can eliminate variables aggressively.

This leads to Algorithm 5. The algorithm combines several operations in an effort to amortise their cost. In what follows we step through the details of the algorithm.

The `to_aff` function finds an initial model μ of R , before translating R through the call to `translate`. This initial call to `translate` has the effect of “xor-ing” μ with all of the models of R . Once translated, the xor closure is taken, before translating again using the initial model μ to obtain the affine closure.

`translate` is responsible for computing the xor of a model with an ROBDD. Its operation relies on the observation that for a given node v in the ROBDD, if $\mu(v) = 1$, then the operation is equivalent to exchanging the “then” and “else” branches of v .

`xor_close` is used to compute the xor-closure of an ROBDD R . The third argument passed to `trav` is an accumulator in which the result is constructed. As in Algorithm 4, we know that $\bar{0}$ will be a model of the result, so we initialise the accumulator as (the ROBDD for) $\bigwedge\{\bar{v} \mid v \in \text{vars}(R)\}$.

`trav` implements a recursive traversal of the ROBDD, and when a model is found in μ , we “extend” the affine envelope to include the newly found model. Namely, `extend`(R, S, μ) produces (the ROBDD for) $R \vee S_\mu$. Note that once a model is found during the traversal, `trav` checks if μ is already present within the xor-closure, and if it is not, invokes `extend` accordingly. This simple check avoids making unnecessary calls to `extend`.

The `cons` function represents a special case of `mknd`. It takes an additional argument in μ and uses it to determine whether to restrict away the corresponding node being constructed. The correctness of `cons` rests on Propositions 2.4 and 2.6, which guarantee that affine approximation can be interspersed with variable elimination, so that the latter can be performed aggressively.

Finally, once a model is found during a traversal, `extend` is used to build up the affine closure of the ROBDD. In the context of the initial call `extend`(S, S, μ), Corollary 2.7 ensures that the pattern of the last equation for `extend` is sufficient: If neither argument is a sink, the two will have the same root variable.

Example 3.2 Consider the ROBDD R (shown again in Figure 3(a)), whose set of models is $\{00011, 00110, 01001, 01101, 10101\}$. Picking $\mu = 00011$ and translating gives R_μ , shown in Figure 3(b). This ROBDD represents a set of vectors $\{00000, 00101, 01010, 01110, 10110\}$ which is to be extended to a vector space.

The algorithm now builds up S , the xor-closure of R_μ , by taking one vector v at a time from R_μ and extending S to a vector space that includes v . S begins as the zero vector.

The first step of the algorithm just adds 00101 to the existing zero vector (Figure 3(c)). The next step comes across the vector 01X10 (which actually represents two valuations) and existentially quantifies away the variable x (Figure 4(a)). Note that the variable z

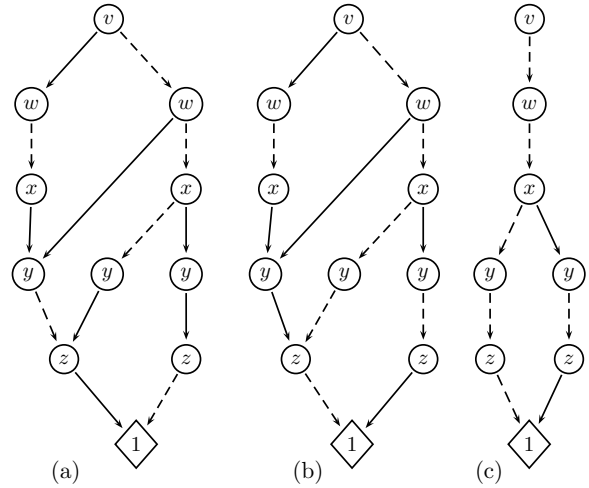


Figure 3: (a): The ROBDD R from Figure 1. (b): The translated version R_μ . (c): The vector space S that has been extended to cover 00101.

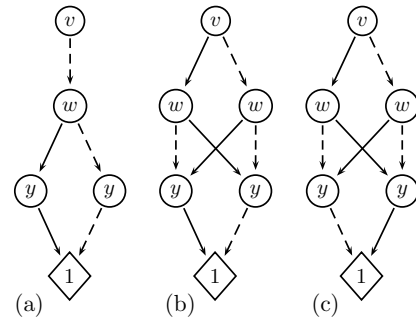


Figure 4: (a): The vector space S after being extended to cover 01X10. (b): S after extending to cover 10110. (c): S translated to give the affine closure of R .

also disappears: this is due to the extension required to include 01X10 that adds enough valuations such that z is “covered” by the vector space.

Extending to cover 10110 simply requires every model to be copied, with v mapped to 1 (Figure 4(b)). Finally, translating back by μ produces A , the affine closure of R , shown in Figure 4(c). ■

4 Experimental Evaluation

To evaluate Algorithms 3 and 5 we generated random Boolean functions using Algorithm 6. We generated random Boolean functions of n variables, with an additional parameter to control the density of the generated function, that is, to set the likelihood of a random valuation being a model. For Algorithm 3 we extracted models from the generated ROBDDs, so that both algorithms were tested on identical Boolean functions.

The function `gen_rand_bdd`(n, pr) builds, in the form of an ROBDD R , a random Boolean function with the property that the likelihood of an arbitrary valuation satisfying R is 2^{-pr} . It invokes `rand_bdd`($0, n-1, pr$). This recursive algorithm builds a ROBDD of $(n-pr)$ variables and at depth $(n-pr)$, a random choice is made as to whether to continue generating the random function or to simply join the

Algorithm 5 Affine envelopes for ROBDDs

Input: An ROBDD R .

Output: The affine envelope of R .

```
to_aff(0) = 0
to_aff(R) = let  $\mu = \text{get\_model}(R)$  in translate(xor_close(translate(R,  $\mu$ ),  $\mu$ )

translate(0,  $\perp$ ) = 0
translate(1,  $\perp$ ) = 1
translate(ite( $x, T, E$ ),  $\mu$ )
  | ( $\mu(x) = 0$ ) = cons( $x$ , translate( $T, \mu$ ), translate( $E, \mu$ ),  $\mu$ )
  | ( $\mu(x) = 1$ ) = cons( $x$ , translate( $E, \mu$ ), translate( $T, \mu$ ),  $\mu$ )

xor_close(R) = trav(R,  $\lambda v. \perp$ ,  $\bigwedge \{\bar{v} \mid v \in \text{vars}(R)\}$ )

trav(0,  $\perp$ ,  $S$ ) =  $S$ 
trav(1,  $\mu$ ,  $S$ )
  | ( $\mu \models S$ ) =  $S$ 
  | otherwise = extend( $S, S, \mu$ )
trav(ite( $x, T, E$ ),  $\mu$ ,  $S$ ) = trav( $T, \mu[x \mapsto 1]$ , trav( $E, \mu[x \mapsto 0]$ ,  $S$ ))

cons( $x, T, E, \mu$ )
  | ( $\mu(x) = \perp$ ) = or( $T, E$ )
  | otherwise = mknd( $x, T, E$ )

extend(1,  $\perp$ ,  $\perp$ ) = 1
extend( $\perp$ , 1,  $\perp$ ) = 1
extend(0,  $S, \mu$ ) = translate( $S, \mu$ )
extend(ite( $x, T, E$ ), 0,  $\mu$ ) = cons( $x$ , extend( $T, 0, \mu$ ), extend( $E, 0, \mu$ ),  $\mu$ )
extend(ite( $x, T, E$ ), ite( $x, T', E'$ ),  $\mu$ )
  | ( $\mu(x) = 1$ ) = mknd( $x$ , extend( $T, E', \mu$ ), extend( $E, T', \mu$ ))
  | otherwise = cons( $x$ , extend( $T, T', \mu$ ), extend( $E, E', \mu$ ),  $\mu$ )
```

Algorithm 6 Generation of random Boolean functions as ROBDDs

Input: The number n of variables in the random function,

pr a calibrator set so that the probability
of a valuation being a model is 2^{-pr} .

Output: A random Boolean function represented as an ROBDD.

```
gen_rand_bdd( $n, pr$ ) = rand_bdd(0,  $n - 1, pr$ )

rand_bdd( $m, n, pr$ )
  | ( $m = n$ ) = mknd( $m$ , rand_sink(), rand_sink())
  | otherwise = mknd( $m, T, E$ )
    where
       $T = \text{if } (m > n - pr) \wedge \text{cointoss}() \text{ then rand\_bdd}(m + 1, n, pr) \text{ else } 0$ 
       $E = \text{if } (m > n - pr) \wedge \text{cointoss}() \text{ then rand\_bdd}(m + 1, n, pr) \text{ else } 0$ 

rand_sink() = if cointoss() then 1 else 0

cointoss() returns 1 or 0 with equal probability.
```

Variables	Algorithm 3	Algorithm 5
12	0.021	0.017
15	5.991	0.272
18	—	0.407
21	—	1.710
24	—	14.967

Table 1: Average time in milliseconds to compute one affine envelope

branch with a 0 sink. If the choice is to continue, then the algorithm recursively applies `rand_bdd($m + 1, n, pr$)` to the branch.

By building a “complete” ROBDD of $(n - pr)$ variables, we were able to distribute the number of models for a given number of variables. In this way, we were able to compare the various algorithms for differing model distributions.

Table 1 shows the average time (in milliseconds) taken by each of the algorithms over 10,000 repetitions with the probability $1/1024$ of a valuation being a model. Timing data were collected on a machine running Solaris 9, with two Intel Xeon CPUs running at 2.8GHz and 4GB of memory. Only one CPU was used and tests were run under minimal load on the system. Our implementation of Algorithm 3 uses sorted arrays of bitstrings (so that search for models is logarithmic). As the number of models grows exponentially with the number of variables, it is not surprising that memory consumption exceeded available space, so we were unable to collect timing data for more than 15 variables.

5 Conclusion

Approximation and the generation of envelopes for Boolean formulas is used extensively in the querying of knowledge bases. Previous research has focused on the use of Horn approximations represented in conjunctive normal form (CNF). In this paper, following the suggestion of Zanuttini, we instead focused on the class of affine functions, using an approximation algorithm suggested by Zanuttini (2002a). Our initial implementation using a naive sets-of-models (as arrays of bitstrings) representation was disappointing, as even for functions with very few models, the affine envelope often has very many models (in fact, the affine envelope of very many functions is 1), so storing sets of models as an array becomes prohibitive even for functions over rather few variables.

ROBDDs have proved to be an appropriate representation for many applications of Boolean functions. Functions with very many models, as well as very few, have compact ROBDD representations. Thus we have developed a new affine envelope algorithm using ROBDDs. Our approach is based on the same principle as Zanuttini’s, but takes advantage of some useful characteristics of ROBDDs. In particular, Propositions 2.4 and 2.6 allow us to project away variables aggressively, often significantly reducing the sizes of the representations being manipulated earlier than would happen otherwise.

Further research in this area includes implementing Zanuttini’s suggested modulo 2 congruence equations representation and comparing to our ROBDD implementation. This also includes evaluating the cost of determining whether a set of congruence equations entail a given general Boolean function. It would also be interesting to compare affine approximation to approximation to other classes for information loss to

evaluate whether affine functions really are as suitable for knowledge-base approximations as Horn or other functions.

References

- Brace, K., Rudell, R. & Bryant, R. (1990), Efficient implementation of a BDD package, in ‘Proc. Twenty-seventh ACM/IEEE Design Automation Conf.’, pp. 40–45.
- Bryant, R. (1986), ‘Graph-based algorithms for Boolean function manipulation’, *IEEE Trans. Computers* **C-35**(8), 677–691.
- Bryant, R. (1992), ‘Symbolic Boolean manipulation with ordered binary-decision diagrams’, *ACM Computing Surveys* **24**(3), 293–318.
- Dechter, R. & Pearl, J. (1992), ‘Structure identification in relational data’, *Artificial Intelligence* **58**, 237–270.
- del Val, A. (2005), ‘First order LUB approximations: Characterization and algorithms’, *Artificial Intelligence* **162**, 7–48.
- Horiyama, T. & Ibaraki, T. (2002), ‘Ordered binary decision diagrams as knowledge-bases’, *Artificial Intelligence* **136**, 189–213.
- Ore, O. (1943), ‘Combinations of closure relations’, *Ann. Math.* **44**(3), 514–533.
- Post, E. (1941), *The Two-Valued Iterative Systems of Mathematical Logic*, Princeton University Press. Reprinted in M. Davis, Solvability, Provability, Definability: The Collected Works of Emil L. Post, pages 249–374, Birkhäuser, 1994.
- Schachte, P. & Søndergaard, H. (2006), Closure operators for ROBDDs, in E. A. Emerson & K. Namjoshi, eds, ‘Proc. Seventh Int. Conf. Verification, Model Checking and Abstract Interpretation’, Vol. 3855 of *Lecture Notes in Computer Science*, Springer, pp. 1–16.
- Schachte, P. & Søndergaard, H. (2007), Boolean approximation revisited, in I. Miguel & W. Ruml, eds, ‘Abstraction, Reformulation and Approximation: Proc. SARA 2007’, Vol. 4612 of *Lecture Notes in Artificial Intelligence*, Springer, pp. 329–343.
- Schaefer, T. J. (1978), The complexity of satisfiability problems, in ‘Proc. Tenth Ann. ACM Symp. Theory of Computing’, pp. 216–226.
- Selman, B. & Kautz, H. (1996), ‘Knowledge compilation and theory approximation’, *Journal of the ACM* **43**(2), 193–224.
- Ward, M. (1942), ‘The closure operators of a lattice’, *Ann. Math.* **43**(2), 191–196.
- Zanuttini, B. (2002a), Approximating propositional knowledge with affine formulas, in ‘Proc. Fifteenth European Conf. Artificial Intelligence (ECAI’02)’, IOS Press, pp. 287–291.
- Zanuttini, B. (2002b), Approximation of relations by propositional formulas: Complexity and semantics, in S. Koenig & R. Holte, eds, ‘Abstraction, Reformulation and Approximation: Proc. SARA 2002’, Vol. 2371 of *Lecture Notes in Artificial Intelligence*, Springer, pp. 242–255.