The Connectivity of Boolean Satisfiability:
Dichotomies for Formulas and Circuits

Konrad W. Schwerdtfeger
Institut für Theoretische Informatik, Leibniz Universität Hannover,
Appelstr. 4, 30167 Hannover, Germany
k.w.s@gmx.net

Abstract
For Boolean satisfiability problems, the structure of the solution space is characterized by the solution graph, where the vertices are the solutions, and two solutions are connected if they differ in exactly one variable. Motivated by research on heuristics and the satisfiability threshold, in 2006, Gopalan et al. studied connectivity properties of the solution graph and related complexity issues for CSPs [GKMP06, GKMP09]. They found dichotomies for the diameter of connected components and for the complexity of the st-connectivity question, and conjectured a trichotomy for the connectivity question. Their results were refined by Makino et al. [MTY07]. Recently, we were able to establish the trichotomy [Sch13].

Here, we consider connectivity issues of satisfiability problems defined by Boolean circuits and propositional formulas that use gates, resp. connectives, from a fixed set of Boolean functions. We obtain dichotomies for the diameter and the connectivity problems: on one side, the diameter is linear, and both problems are in P, while on the other, the diameter can be exponential, and the problems are PSPACE-complete.

Keywords: Boolean satisfiability, Boolean circuits, Post’s lattice, computational complexity, PSPACE-completeness, dichotomy theorems, graph connectivity.

1 Introduction
The Boolean satisfiability problem, as well as many related questions like equivalence, counting, enumeration, and numerous versions of optimization, are of great importance in both theory and applications of computer science.

Common to all these problems is that one asks questions about a Boolean relation given by some short description, e.g. a propositional formula, Boolean circuit, binary decision diagram, or Boolean neural network. For the usual formulas with the connectives $\land$, $\lor$ and $\neg$, several generalizations and restrictions have been considered. Most widely studied are Boolean constraint satisfactions problems (CSPs), that can be seen as a generalization of formulas in $CNF$ (conjunctive normal form), see Definition 3. Another generalization, that we will consider here, are formulas with connectives from an arbitrary fixed set of
Boolean functions $B$, known as $B$-formulas. This concept also applies to circuits, where the allowed gates implement the functions from $B$, called $B$-circuits. A further extension that allows for shorter representations, and in turn makes many problems harder, are quantifiers, which we will look at in Section 5.

Here we will investigate the structure of the solution space, which is of obvious relevance to these satisfiability related problems. Indeed, the solution space connectivity is strongly correlated to the performance of standard satisfiability algorithms like WalkSAT and DPLL on random instances: As one approaches the satisfiability threshold (the ratio of constraints to variables at which random $k$-CNF-formulas become unsatisfiable for $k \geq 3$) from below, the solution space fractures, and the performance of the algorithms breaks down [MMZ05, MMW07]. These insights mainly came from statistical physics, and lead to the development of the survey propagation algorithm, which has much better performance on random instances [MMW07]. This research was a motivation for Gopalan et al. to study connectivity properties of the solution space of Boolean CSPs [GKM09].

While the most efficient satisfiability solvers take CNF-formulas as input, one of the most important applications of satisfiability testing is verification and optimization in Electronic Design Automation (EDA), where the instances derive mostly from digital circuit descriptions [WLLH07]. Though many such instances can easily be encoded in CNF, the original structural information, such as signal ordering, gate orientation and logic paths, is lost, or at least obscured. Since exactly this information can be very helpful for solving these instances, considerable effort has been made recently to develop satisfiability solvers that work with the circuit description directly [WLLH07], which have far superior performance in EDA applications, or to restore the circuit structure from CNF [FM07]. This is one major motivation for our study.

A direct application of $st$-connectivity are reconfiguration problems, that arise when we wish to find a step-by-step transformation between two feasible solutions of a problem such that all intermediate results are also feasible. Recently, the reconfiguration versions of many problems such as INDEPENDENT-SET, VERTEX-COVER, SET-COVER GRAPH-$k$-COLORING, SHORTEST-PATH have been studied, and many complexity results were obtained, in some cases making use of Gopalan et al.’s results [IDH+11, KMM11]. Another related problem for which the solution space connectivity could be of interest is structure identification, where one is given a relation explicitly and seeks a short representation of some kind [CKZ08]; this problem is important especially in artificial intelligence.

Since many of the satisfiability related problems are hard to solve in general (they are NP- or even PSPACE-complete), one has tried to identify easier fragments and to classify restrictions in terms of their complexity. Possibly the best known result is Schaefer’s 1978 dichotomy theorem for CSPs, which states that for certain classes of allowed constraints the satisfiability of a CSP is in $P$, while it is NP-complete for all other classes [Sch78]. Analogously, Gopalan et al. in 2006 classified the complexity of connectivity questions for CSPs in Schaefer’s framework. In this paper, we consider the same connectivity issues as Gopalan et al., but for problems defined by Boolean circuits and propositional formulas.
that use gates, resp. connectives, from a fixed set of Boolean functions.

2 Propositional Formulas and their Solution Space Connectivity

Definition 1. An \( n \)-ary Boolean relation is a subset of \( \{0, 1\}^n \) \( (n \geq 1) \). The set of solutions of a propositional formula \( \phi \) with \( n \) variables defines in a natural way an \( n \)-ary Boolean relation \( R \), where the variables are taken in lexicographic order. The solution graph \( G(\phi) \) of \( \phi \) is the subgraph of the \( n \)-dimensional hypercube graph induced by the vectors in \( R \), i.e., the vertices of \( G(\phi) \) are the vectors in \( R \), and there is an edge between two vectors precisely if they differ in exactly one position.

Figure 1: Depictions of the subgraph of the 5-dimensional hypercube graph induced by a typical random Boolean relation with 12 elements. Left: highlighted on a orthographic hypercube projection. Center: highlighted on a “Spectral Embedding” of the hypercube graph by MATHEMATICA.

Figure 2: Subgraphs of the 8-dimensional hypercube graph induced by typical random relations with 40, 60 and 80 elements.

Definition 2. We use \( a, b, \ldots \) to denote vectors of Boolean values and \( x, y, \ldots \) to denote vectors of variables, \( a = (a_1, a_2, \ldots) \) and \( x = (x_1, x_2, \ldots) \). The Hamming distance \( |a - b| \) of two Boolean vectors \( a \) and \( b \) is the number of positions in which they differ. If \( a \) and \( b \) are solutions of \( \phi \) and lie in the same connected component of \( G(\phi) \), we write \( d_\phi(a, b) \) to denote the shortest-path distance between \( a \) and \( b \). The diameter of a connected component is the maximal
shortest-path distance between any two vectors in that component. The diameter of \( G(\phi) \) is the maximal diameter of any of its connected components.

In our proofs for \( B \)-formulas and \( B \)-circuits, we will use Gopalan et al.’s results for 3-CNF-formulas, so we also need to introduce some terminology for constraint satisfaction problems.

**Definition 3.** A CNF-formula is a Boolean formula of the form \( C_1 \land \ldots \land C_m \) (\( 1 \leq m < \infty \)), where each \( C_i \) is a clause, that is, a finite disjunction of literals (variables or negated variables). A \( k \)-CNF-formula (\( k \geq 1 \)) is a CNF-formula where each \( C_i \) has at most \( k \) literals.

A \( k \)-clause is a disjunction of \( k \) variables or negated variables. For \( 0 \leq i \leq k \), let \( D_i \) be the set of all satisfying truth assignments of the \( k \)-clause whose first \( i \) literals are negated, and let \( S_k = \{ D_0, \ldots, D_k \} \). Thus, \( \text{CNF}(S_k) \) is the collection of \( k \)-CNF-formulas.

For a finite set of Boolean relations \( S \), a CNF(\( S \))-formula (with constants) over a set of variables \( V \) is a finite conjunction \( C_1 \land \ldots \land C_m \), where each \( C_i \) is a constraint application (constraint for short), i.e., an expression of the form \( R(\xi_1, \ldots, \xi_k) \), with a \( k \)-ary relation \( R \in S \), and each \( \xi_j \) is a variable in \( V \) or one of the constants 0, 1.

Gopalan et al. studied the following two decision problems for CNF(\( S \))-formulas:

- the connectivity problem \( \text{Conn}(S) \): given a CNF(\( S \))-formula \( \phi \), is \( G(\phi) \) connected? (if \( \phi \) is unsatisfiable, then \( G(\phi) \) is considered connected)
- the \( st \)-connectivity problem \( \text{st-Conn}(S) \): given a CNF(\( S \))-formula \( \phi \) and two solutions \( s \) and \( t \), is there a path from \( s \) to \( t \) in \( G(\phi) \)?

**Lemma 4.** [GKMP09, Lemm 3.6] \( \text{st-Conn}(S_3) \) and \( \text{Conn}(S_3) \) are PSPACE-complete.

**Proof.** \( \text{st-Conn}(S_3) \) and \( \text{Conn}(S_3) \) are in PSPACE: Given a CNF(\( S_3 \))-formula \( \phi \) and two solutions \( s \) and \( t \), we can guess a path of length at most \( 2^n \) between them and verify that each vertex along the path is indeed a solution. Hence \( \text{st-Conn}(S_3) \) is in NPSPACE=PSPACE. For \( \text{Conn}(S_3) \), by reusing space we can check for all pairs of vectors whether they are satisfying and, if they both are, whether they are connected in \( G(\phi) \).

We can not state the full proof for the PSPACE-hardness here. It consists of a direct reduction from the computation of a space-bounded Turing machine \( M \). The input-string \( w \) of \( M \) is mapped to a CNF(\( S_3 \))-formula and two satisfying assignments \( s \) and \( t \), corresponding to the initial and accepting configuration respectively, s.t. \( s \) and \( t \) are connected in \( G(\phi) \) iff \( M \) accepts \( w \).

**Lemma 5.** [GKMP09, Lemm 3.7] For \( n \geq 2 \), there is an \( n \)-ary Boolean function \( f \) with \( f(1, \ldots, 1) = 1 \) and a diameter of at least \( 2^{\lfloor \frac{n}{2} \rfloor} \).

**Proof.** It is easy to construct such a function, see [GKMP09].
3 Circuits, Formulas, and Post’s Lattice

An $n$-ary Boolean function is a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$. Let $B$ be a finite set of Boolean functions.

A $B$-circuit $C$ with input variables $x_1, \ldots, x_n$ is a directed acyclic graph, augmented as follows: Each node (here also called gate) with indegree 0 is labeled with an $x_i$ or a 0-ary function from $B$, each node with indegree $k > 0$ is labeled with a $k$-ary function from $B$. The edges (here also called wires) pointing into a gate are ordered. One node is designated the output gate.

Given values $a_1, \ldots, a_n \in \{0, 1\}$ to $x_1, \ldots, x_n$, $C$ computes an $n$-ary function $f_C$ as follows: A gate $v$ labeled with a variable $x_i$ returns $a_i$, a gate $v$ labeled with a function $f$ computes the value $f(b_1, \ldots, b_k)$, where $b_1, \ldots, b_k$ are the values computed by the predecessor gates of $v$, ordered according to the order of the wires. For a more formal definition see [Vol99].

A $B$-formula is defined inductively: A variable $x$ is a $B$-formula. If $\phi_1, \ldots, \phi_m$ are $B$-formulas, and $f$ is an $n$-ary function from $B$, then $f(\phi_1, \ldots, \phi_n)$ is a $B$-formula; here, we identify the function $f$ and the symbol representing it in a formula.

It is easy to see that the functions computable by a $B$-circuit, as well as the functions definable by a $B$-formula, are exactly those that can be obtained from $B$ by superposition, together with all projections [BCRV03]. By superposition, we mean substitution (that is, composition of functions), permutation and identification of variables, and introduction of fictive variables (variables on which the value of the function does not depend). This class of functions is denoted by $[B]$. $B$ is closed (or said to be a clone) if $[B] = B$. A base of a clone $F$ is any set $B$ with $[B] = F$.

Already in the early 1920s, Emil Post extensively studied Boolean functions, identified all closed classes, found a finite base for each of them, and detected their inclusion structure [Pos41]. The closed classes form a lattice, called Post’s lattice, depicted in Figure 4; the definitions and bases of the classes are given in Table 1. A modern proof can be found in [Zve05]. To define the clones, we introduce some special properties of Boolean functions: An $n$-ary Boolean function $f$ is said to be

- $a$-reproducing, if $f(a, \ldots, a) = a$, for $a \in \{0, 1\}$,
- monotonic, if $a_1 \leq b_1, \ldots, a_n \leq b_n$ implies $f(a_1, \ldots, a_n) \leq f(b_1, \ldots, b_n)$,
- self-dual, if $f(x_1, \ldots, x_n) = f(x_1, \ldots, \overline{x_n})$,
- affine, if $f(x_1, \ldots, x_n) = x_i_1 \oplus \ldots \oplus x_i_m \oplus c$ with $i_1, \ldots, i_m \in \{1, \ldots, n\}$ and $c \in \{0, 1\}$,
- $c$-separating, if there exists an $i \in \{1, \ldots, n\}$ s.t. $a_i = c$ for all $a \in f^{-1}(c)$, for $c \in \{0, 1\}$,
- $c$-separating of degree $m$, if for all $U \subseteq f^{-1}(c)$ of size $|U| = m$ there exists an $i \in \{1, \ldots, n\}$ s.t. $a_i = c$ for all $a \in U$, for $c \in \{0, 1\}$ and $m \geq 2$.

Not surprisingly, the complexity of problems defined by $B$-formulas and $B$-circuits depends on $[B]$, and the complexity of numerous problems for $B$-circuits and $B$-formulas has been classified by means of Post’s lattice [RW00, Sch07],
starting with satisfiability: Analogously to Schaefer, Lewis in 1978 found a dichotomy for $B$-formulas [Lew79]: if $[B]$ contains the function $x \land \overline{y}$, SAT is NP-complete, else it is in P.

While for $B$-circuits the complexity of every decision problem solely depends on $[B]$ (up to AC$^0$ isomorphism), for formulas this need not be the case, since the transformation of a $B$-formula into a $B'$-formula might require an exponential increase in the formula size even if $[B] = [B']$, as the $B'$-representation of some function from $B$ may need to use some input variable more than once [Mic12]. For example, let $h(x, y) = x \land y$; then there is no shorter $\{h\}$-representation of the function $x \land y$ than $h(x, h(x, y))$.

4 Computational and Structural Dichotomies

Now we consider the connectivity problems for $B$-formulas and $B$-circuits:

- $\text{BF-Conn}(B)$: Given a $B$-formula $\phi$, is $G(\phi)$ connected?
- $\text{st-BF-Conn}(B)$: Given a $B$-formula $\phi$ and two solutions $s$ and $t$, is there a path from $s$ to $t$ in $G(\phi)$?

The corresponding problems for circuits are denoted $\text{Circ-Conn}(B)$ resp. $\text{st-Circ-Conn}(B)$.

**Theorem 6.** Let $B$ be a finite set of Boolean functions.

1. If $B \subseteq M$, $B \subseteq L$, or $B \subseteq S_0$, then
   (a) $\text{st-Circ-Conn}(B)$ and $\text{Circ-Conn}(B)$ are in $P$,
   (b) $\text{st-BF-Conn}(B)$ and $\text{BF-Conn}(B)$ are in $P$,
   (c) the diameter of every function $f \in [B]$ is linear in the number of variables of $f$.

2. Otherwise,
   (a) $\text{st-Circ-Conn}(B)$ and $\text{Circ-Conn}(B)$ are PSPACE-complete,
   (b) $\text{st-BF-Conn}(B)$ and $\text{BF-Conn}(B)$ are PSPACE-complete,
   (c) there are functions $f \in [B]$ such that their diameter is exponential in the number of variables of $f$.

The proof follows from the Lemmas in the next subsections. By the following Proposition, we can relate the complexity of $B$-formulas and $B$-circuits.

**Proposition 7.** Every $B$-formula can be transformed into an equivalent $B$-circuit in polynomial time.

**Proof.** A $B$-formula already is a suitable encoding for a special $B$-circuit with outdegree of at most one. \qed
4.1 The Easy Side of the Dichotomy

Lemma 8. If $B \subseteq M$, the solution graph of any $n$-ary function $f \in [B]$ is connected, and $d_f(a, b) = |a - b| \leq n$ for any two solutions $a$ and $b$.

Proof. The table of all closed classes of Boolean functions shows that $f$ is monotonic in this case. Thus, either $f = 0$, or $(1, \ldots, 1)$ must be a solution, and every other solution $a$ is connected to $(1, \ldots, 1)$ in $G(\phi)$ since $(1, \ldots, 1)$ can be reached by flipping the variables assigned 0 in $a$ one at a time to 1. Further, if $a$ and $b$ are solutions, $b$ can be reached from $a$ in $|a - b|$ steps by first flipping all variables that are assigned 0 in $a$ and 1 in $b$, and then flipping all variables that are assigned 1 in $a$ and 0 in $b$. \hfill \Box

Lemma 9. If $B \subseteq S_0$, the solution graph of any function $f \in [B]$ is connected, and $d_f(a, b) \leq |a - b| + 2$ for any two solutions $a$ and $b$.

Proof. Since $f$ is 0-separating, there is an $i$ such that $a_i = 0$ for every vector $a$ with $f(a) = 0$, thus every $b$ with $b_i = 1$ is a solution. It follows that every solution $t$ can be reached from any solution $s$ in at most $|s - t| + 2$ steps by first flipping the $i$-th variable from 0 to 1 if necessary, then flipping all other variables in which $s$ and $t$ differ, and finally flipping back the $i$-th variable if necessary. \hfill \Box

Lemma 10. If $B \subseteq L$,

1. st-Circ-Conn$(B)$ and Circ-Conn$(B)$ are in $P$,
2. st-BF-Conn$(B)$ and BF-Conn$(B)$ are in $P$,
3. for any function $f \in [B]$, $d_f(a, b) = |a - b|$ for any two solutions $a$ and $b$ that lie in the same connected component of $G(\phi)$.

Proof. Since every function $f \in L$ is linear, $f(x_1, \ldots, x_n) = x_{i_1} \oplus \cdots \oplus x_{i_m} \oplus c$, and any two solutions $s$ and $t$ are connected iff they differ only in fictional variables: If $s$ and $t$ differ in at least one non-fictional variable (i.e., an $x_i \in \{x_{i_1}, \ldots, x_{i_m}\}$), to reach $t$ from $s$, $x_i$ must be flipped eventually, but for every solution $a$, any vector $b$ that differs from $a$ in exactly one non-fictional variable is no solution. If $s$ and $t$ differ only in fictional variables, $t$ can be reached from $s$ in $|s - t|$ steps by flipping one by one the variables in which they differ.

Since $\{x \oplus y, 1\}$ is a base of $L$, every $B$-circuit $C$ can be transformed in polynomial time into an equivalent $\{x \oplus y, 1\}$-circuit $C'$ by replacing each gate of $C$ with an equivalent $\{x \oplus y, 1\}$-circuit. Now one can decide in polynomial time whether a variable $x_i$ is fictional by checking for $C'$ whether the number of “backward paths” from the output gate to gates labeled with $x_i$ is odd, so st-Circ-Conn$(B)$ is in $P$.

$G(\phi)$ is connected iff at most one variable is non-fictional, thus Circ-Conn$(B)$ is in $P$.

By Proposition 7, st-BF-Conn$(B)$ and BF-Conn$(B)$ are in $P$ also. \hfill \Box

This completes the proof of the easy side of the dichotomy.
4.2 The Hard Side of the Dichotomy

**Proposition 11.** \( \text{st-Circ-Conn}(B) \) and \( \text{Circ-Conn}(B) \), as well as \( \text{st-BF-Conn}(B) \) and \( \text{BF-Conn}(B) \), are in \( \text{PSPACE} \) for any finite set \( B \) of Boolean functions.

*Proof.* This follows as in Lemma 4.

**Proposition 12.** For 1-reproducing 3-CNF-formulas, the problems \( \text{st-Conn} \) and \( \text{Conn} \) are \( \text{PSPACE-complete} \).

*Proof.* We chose the variables in the proof of Lemma 4 such that the accepting configuration of the Turing machine corresponds to the \((1,\ldots,1)\) vector.

An inspection of Post’s lattice shows that if \( B \not\subseteq M, B \not\subseteq L, \) and \( B \not\subseteq S_0 \), then \( |B| \geq S_{12}, |B| \geq D_1, \) or \( |B| \geq S_{12}^k, \forall k \geq 2, \) so we have to prove \( \text{PSPACE-completeness} \) and show the existence of \( B \)-formulas with an exponential diameter in these cases.

In the proofs, we will use the following abbreviations: If we have the \( n \) variables \( x_1, \ldots, x_n \), we write \( \mathbf{x} \) for \( x_1 \land \cdots \land x_n \) and \( \mathbf{f} \) for \( \mathbf{x} \land \cdots \land \mathbf{x} \). Also, we write \( (\mathbf{x} = c_1 \cdots c_n) \) for \( x_1 \leftrightarrow c_1 \land \cdots \land x_n \leftrightarrow c_n \), where \( c_1, \ldots, c_n \in \{0,1\} \) are constants; e.g., we write \( (\mathbf{x} = 101) \) for \( x_1 \land \mathbf{f}_2 \land x_3 \). Further, we use \( \mathbf{x} \in \{a,b,\ldots\} \) for \( (\mathbf{x} = a) \lor (\mathbf{x} = b) \lor \ldots \). Finally, if we have two vectors of Boolean values \( a \) and \( b \) of length \( n \) and \( m \) resp., we write \( a \cdot b \) for their concatenation \((a_1,\ldots,a_n,b_1,\ldots,b_m)\).

**Lemma 13.** If \( |B| \geq S_{12} \),

1. \( \text{st-BF-Conn}(B) \) and \( \text{BF-Conn}(B) \) are \( \text{PSPACE-complete} \),
2. \( \text{st-Circ-Conn}(B) \) and \( \text{Circ-Conn}(B) \) are \( \text{PSPACE-complete} \),
3. for \( n \geq 3 \), there is an \( n \)-ary function \( f \in |B| \) with diameter of at least \( 2^{\frac{n+1}{2}} \).

*Proof.* 1. We reduce the problems for 1-reproducing 3-CNF-formulas to the ones for \( B \)-formulas: We map a 1-reproducing 3-CNF-formula \( \phi \) and two solutions \( s \) and \( t \) of \( \phi \) to a \( B \)-formula \( \phi' \) and two solutions \( s' \) and \( t' \) of \( \phi' \) such that \( s' \) and \( t' \) are connected in \( G(\phi') \) iff \( s \) and \( t \) are connected in \( G(\phi) \), and such that \( G(\phi') \) is connected iff \( G(\phi) \) is connected.

First for any 1-reproducing formula \( \psi \), we define a connectivity-equivalent formula \( T_\psi \in S_{12} \) using the standard connectives, then we show how to transform \( \phi \) into the \( B \)-formula \( \phi' \) that will be equivalent to \( T_\phi \).

Let \( \psi \) be a 1-reproducing formula over the variables \( x_1, \ldots, x_n \). We define the formula \( T_\psi \) over the \( n + 1 \) variables \( x_1, \ldots, x_n \) and \( y \) as

\[
T_\psi = \psi \land y,
\]

where \( y \) is a new variable. All solutions \( a \) of \( T_\psi(x,y) \) have \( a_{n+1} = 1 \), so \( T_\psi \) is 1-separating and 0-reproducing. Moreover, \( T_\psi \) is still 1-reproducing, and thus in \( S_{12} \). For any two solutions \( s \) and \( t \) of \( \psi(x) \), \( s' = s \cdot 1 \) and \( t' = t \cdot 1 \) are solutions.
of \( T_\psi(x, y) \), and it is easy to see that they are connected in \( G(T_\psi) \) iff \( s \) and \( t \) are connected in \( G(\psi) \), and that \( G(T_\psi) \) is connected iff \( G(\psi) \) is connected.

Now we know that for any 1-reproducing 3-CNF-formula \( \phi \), \( T_\phi \) can be expressed as a \( B \)-formula \( \phi' \) since \( T_\phi \in S_{12} \). However, the transformation could lead to an exponential increase in the formula size (see Section 3), so we have to show how to construct \( \phi' \) in polynomial time. We do this by parenthesizing the conjunctions of \( \phi \) such that we get a tree of \( \land \)'s of depth logarithmic in the size of \( \phi \), and then replacing each clause \( C_i \) with some \( B \)-formula \( \xi_{C_i} \), and each expression \( \phi_1 \land \phi_2 \) with a \( B \)-formula \( \xi_{\land}(\phi_1, \phi_2) \), s.t. the resulting formula is equivalent to \( T_\phi \). This can increase the formula size by only a polynomial in the original size even if \( \xi_{\land} \) uses some input variable more than once. This is a standard-technique for such proofs in Post’s framework, see e.g. [BCRV03]. Here we easily see that we can simply replace each clause \( C_i \) of \( \phi \) with some \( B \)-formula equivalent to \( T_{C_i} \), and each \( \land \) with a \( B \)-formula equivalent to \( T_\land \) since \( (\psi_1 \land y) \land (\psi_2 \land y) \land y \equiv \psi_1 \land \psi_2 \land y \), but in the next proofs this will not be obvious, so we formalize the procedure.

Let \( \phi = C_1 \land \cdots \land C_n \) be a 1-reproducing 3-CNF-formula. Since \( \phi \) is 1-reproducing, every clause \( C_i \) of \( \phi \) is itself 1-reproducing, and we can express \( T_{C_i} \) through a \( B \)-formula \( T_{C_i}^* \). Also, we can express \( T_\land(x_1, x_2) = x_1 \land x_2 \land y \) through a \( B \)-formula \( T_\land^* \) since \( \land \) is 1-reproducing. Now let \( \phi' = \text{Tr}(T_{C_1}^*, \ldots, T_{C_n}^*) \), where \( \text{Tr} \) is the following recursive algorithm that takes a list of formulas as input,

Algorithm \( \text{Tr}(\psi_1, \ldots, \psi_m) \)

1. if \( m = 1 \) return \( \psi_1 \)
2. else if \( m \) is even, return
   \[ \text{Tr}(T_{\land}^*[x_1/\psi_1, x_2/\psi_2], T_{\land}^*[x_1/\psi_3, x_2/\psi_4], \ldots, T_{\land}^*[x_1/\psi_{m-1}, x_2/\psi_m]) \]
3. else return
   \[ \text{Tr}(T_{\land}^*[x_1/\psi_1, x_2/\psi_2], T_{\land}^*[x_1/\psi_3, x_2/\psi_4], \ldots, T_{\land}^*[x_1/\psi_{m-2}, x_2/\psi_{m-1}], \psi_m) \).

Here \( \psi[x_i/\xi] \) denotes the formula obtained by substituting the formula \( \xi \) for the variable \( x_i \) in the formula \( \psi \). Note that in every \( T_\psi^* \) we have the same variable \( y \).

Since the recursion terminates after a number of steps logarithmic in the number of clauses of \( \phi \), and every step increases the total formula size by only a constant factor, the algorithm runs in polynomial time. We show \( \phi' = T_\phi \) by induction. The basis is clear. Since \( T_\psi \equiv T_\psi^* \), it suffices to show that \( T_\land[x_1/T_{\psi_1}, x_2/T_{\psi_2}] \equiv T_{\psi_1 \land \psi_2} \):

\[ T_\land[x_1/T_{\psi_1}, x_2/T_{\psi_2}] = (\psi_1 \land y) \land (\psi_2 \land y) \land y \equiv \psi_1 \land \psi_2 \land y = T_{\psi_1 \land \psi_2}. \]

2. This follows from 1. by Proposition 7.
3. By Lemma 5 there is an 1-reproducing \((n-1)\)-ary function \( f \) with diameter of at least \( 2^\lceil \frac{n-1}{2} \rceil \). Let \( f \) be represented by a formula \( \phi \); then, \( T_\phi \) represents an \( n \)-ary function of the same diameter in \( S_{12} \). \( \square \)

Lemma 14. If \([B] \supseteq D_1\),
1. ST-BF-CONN(B) and BF-CONN(B) are PSPACE-complete, 
2. ST-CIRC-CONN(B) and CIRC-CONN(B) are PSPACE-complete,
3. for \( n \geq 5 \), there is an \( n \)-ary function \( f \in [B] \) with diameter of at least \( 2^{\frac{n^2}{2}} \).

Proof. 1. This proof is similar to the previous one, but the construction is more intricate; for every 1-reproducing 3-CNF formula we have to construct a self-dual function s.t. the connectivity is retained. For clarity, we do the construction in two steps.

For a 1-reproducing formula \( \psi \) over the \( n \) variables \( x_1, \ldots, x_n \), we construct a formula \( T_{\psi}^- \in D_1 \) with three new variables \( (y_1,y_2,y_3) = y \),

\[
T_{\psi}^- = (\psi(x) \land y) \lor \left( \overline{\psi(x)} \land \overline{y} \right) \lor y \in \{100,010,001\}.
\]

Observe that \( T_{\psi}^- (x,y) \) is self-dual: for any solution ending with 111, the inverse vector (that ends with 000) is no solution; all vectors ending with 100, 010, or 001 are solutions and their inverses are no solutions. Also, \( T_{\psi}^- \) is still 1-reproducing, and it is 0-reproducing since \( \psi(0\cdots0) \equiv \overline{\psi(1\cdots1)} \equiv 0 \).

Further, for any two solutions \( s \) and \( t \) of \( \psi(x) \), \( s' = s \cdot 111 \) and \( t' = t \cdot 111 \) are solutions of \( T_{\psi}^- (x,y) \) and are connected in \( G(T_{\psi}^-) \) iff \( s \) and \( t \) are connected in \( G(\psi) \): Every solution \( \alpha \) of \( \psi \) corresponds to a solution \( \alpha \cdot 111 \) of \( T_{\psi}^- \), and the connectivity does not change by padding the vectors with 111, and since there are no solutions of \( T_{\psi}^- \) ending with 110, 101, or 011, every other solution of \( T_{\psi}^- \) differs in at least two variables from the solutions \( \alpha \cdot 111 \) that correspond to solutions of \( \psi \).

Note that exactly one connected component is added in \( G(T_{\psi}^-) \) to the components corresponding to those of \( G(\psi) \): It consists of all vectors ending with 000, 100, 010, or 001 (any two vectors ending with 000 are connected e.g. via those ending with 001). It follows that \( G(T_{\psi}^-) \) is always unconnected. To fix this, we modify \( T_{\psi}^- \) to a function \( T_\psi \) by adding \( 1\cdots1\cdot110 \) as a solution, thereby connecting \( 1\cdots1\cdot111 \) (which is always a solution because \( T_{\psi}^- \) is 1-reproducing) with \( 1\cdots1\cdot1\cdot110 \), and thereby with the additional component of \( T_{\psi}^- \). To keep the function self-dual, we must in turn remove \( 0\cdots0\cdot001 \), which does not alter the connectivity. Formally,

\[
T_\psi = (T_{\psi}^- \lor (x \land (y = 110))) \land \neg(x \land (y = 001))
\]

\[
= (\psi(x) \land y) \lor \left( \overline{\psi(x)} \land \overline{y} \right) \lor (y \in \{100,010,001\} \land \neg(x \land (y = 001)) \lor (x \land (y = 110)).
\]  

Now \( G(T_\psi) \) is connected iff \( G(\psi) \) is connected.

Next we use the algorithm \( Tr \) from the previous proof to transform any 1-reproducing 3-CNF-formula \( \phi \) into a \( B \)-formula \( \phi' \) equivalent to \( T_\psi \), but with the definition (1) of \( T \). Again, we have to show \( T \land [x_1/T_{\psi 1},x_2/T_{\psi 2}] \equiv T_{\psi 1 \land \psi 2} \).
Here,
\[ T \land [x_1/T_{\psi_1}, x_2/T_{\psi_2}] = (T_{\psi_1} \land T_{\psi_2} \land y) \lor (T_{\psi_1} \land T_{\psi_2} \land \neg y) \]
\[ \lor (y \in \{100,010,001\} \land \neg (T_{\psi_1} \land T_{\psi_2} \land (y = 001))) \]
\[ \lor (T_{\psi_1} \land T_{\psi_2} \land (y = 110)). \]

We consider the parts of the formula in turn: For any formula \( \xi \) we have
\[ T_{\land} [x_{\xi}/T_{\psi_1}, x_{\xi}/T_{\psi_2}] = \left( T_{\psi_1} \land T_{\psi_2} \land \psi_{\xi} \right) \]
\[ \lor (y \in \{100,010,001\} \land \neg (T_{\psi_1} \land T_{\psi_2} \land (y = 001))) \]
\[ \lor (T_{\psi_1} \land T_{\psi_2} \land (y = 110)). \]

For the second line, we observe
\[ T_{\land} [x_{\psi}/T_{\psi_1}, x_{\psi}/T_{\psi_2}] = \left( T_{\psi_1} \lor T_{\psi_2} \land \psi \right) \]
\[ \lor (y \in \{100,010,001\} \land \neg (T_{\psi_1} \lor T_{\psi_2} \land (y = 001))) \]
\[ \lor (T_{\psi_1} \lor T_{\psi_2} \land (y = 110)). \]

Since \( T_{\land} [x_{\psi}/T_{\psi_1}, x_{\psi}/T_{\psi_2}] = (T_{\psi_1} \lor T_{\psi_2} \land \psi) \) for any \( \psi \), the third line becomes
\[ \lor (x_{\psi_1} \lor x_{\psi_2} \land (y = 110)). \]

Now \( T_{\land} [x_1/T_{\psi_1}, x_2/T_{\psi_2}] \) equals
\[ T_{\psi_1} \land T_{\psi_2} = \left( \psi_1(x_{\psi_1}) \land \psi_2(x_{\psi_2}) \land y \right) \]
\[ \lor (y \in \{100,010,001\} \land \neg (\psi_1(x_{\psi_1}) \land \psi_2(x_{\psi_2}) \land (y = 001))) \]
\[ \lor (x_{\psi_1} \land x_{\psi_2} \land (y = 110)). \]

2. This follows from 1. by Proposition 7.
3. By Lemma 5 there is an 1-reproducing \((n-3)\)-ary function \(f\) with diameter of at least \(2\left\lceil \frac{n-3}{2} \right\rceil\). Let \(f\) be represented by a formula \(\phi\); then, \(T_\phi\) represents an \(n\)-ary function of the same diameter in \(D_1\).

**Lemma 15.** If \([B] \supseteq S_{02}^k\).

1. st-BF-Conn\((B)\) and BF-Conn\((B)\) are PSPACE-complete,
2. st-Circ-Conn\((B)\) and Circ-Conn\((B)\) are PSPACE-complete,
3. for \(n \geq k + 4\), there is an \(n\)-ary function \(f \in [B]\) with diameter of at least \(2\left\lceil \frac{n-4}{2} \right\rceil\).

**Proof.**
1. This proof is analogous to the previous one. For a 1-reproducing formula \(\psi\) over the \(n\) variables \(x_1, \ldots, x_n\), we construct the formula \(T_\psi^\sim \in S_{02}^k\) with the additional variables \(y, z\) and \((z_1, \ldots, z_{k+1}) = z\),

\[
T_\psi^\sim = (\psi \land y \land z) \lor z \notin \{0 \cdots 0, 10 \cdots 0, 010 \cdots 0, \ldots, 0 \cdots 01\}.
\]

\(T_\psi^\sim\) is 0-separating of degree \(k\) since all vectors that are no solutions of \(T_\psi^\sim\) end with a vector \(b \in \{0 \cdots 0, 10 \cdots 0, 010 \cdots 0, \ldots, 0 \cdots 01\} \subset \{0, 1\}^{k+1}\) and thus any \(k\) of them have at least one common variable assigned 0. Also, \(T_\psi^\sim\) is 0-reproducing and still 1-reproducing.

Further, for any two solutions \(s\) and \(t\) of \(\psi(x)\), \(s' = s \cdot 1 \cdots 0\) and \(t' = t \cdot 1 \cdots 0\) are solutions of \(T_\psi^\sim(x, y, z)\) and are connected in \(G(T_\psi^\sim)\) iff \(s\) and \(t\) are connected in \(G(\psi)\).

But again, we have produced an additional connected component (consisting of all vectors not ending with 10 \cdots 0, 010 \cdots 0, \ldots, 0 \cdots 01, or 0 \cdots 0). To connect it to a component corresponding to one of \(\psi\), we add 1 \cdots 1 \cdot 10 \cdots 0 as a solution,

\[
T_\psi = (\psi \land y \land z) \lor z \notin \{0 \cdots 0, 10 \cdots 0, 010 \cdots 0, \ldots, 0 \cdots 01\} \\
\lor (x \land y \land (z = 10 \cdots 0)).
\]

Now \(G(T_\psi)\) is connected iff \(G(\psi)\) is connected.

Again we show that the algorithm \(T_r\) works in this case. Here,

\[
T_\land [x_1/T_{\psi_1}, x_2/T_{\psi_2}] = (T_{\psi_1}(x_{\psi_1}) \land T_{\psi_2}(x_{\psi_2}) \land y \land z) \\
\lor z \notin \{0 \cdots 0, 10 \cdots 0, 010 \cdots 0, \ldots, 0 \cdots 01\} \\
\lor (T_{\psi_1}(x_{\psi_1}) \land T_{\psi_2}(x_{\psi_2}) \land y \land (z = 10 \cdots 0))
\]

which is equivalent to

\[
T_{\psi_1 \land \psi_2} = (\psi_1(x_{\psi_1}) \land \psi_2(x_{\psi_2}) \land y \land z) \\
\lor z \notin \{0 \cdots 0, 10 \cdots 0, 010 \cdots 0, \ldots, 0 \cdots 01\} \\
\lor (x_{\psi_1} \land y \land (z = 10 \cdots 0)).
\]

2. This follows from 1. by Proposition 7.

3. By Lemma 5 there is an 1-reproducing \((n - k - 2)\)-ary function \(f\) with diameter of at least \(2\left\lceil \frac{n-k-2}{2} \right\rceil\). Let \(f\) be represented by a formula \(\phi\); then, \(T_\phi\) represents an \(n\)-ary function of the same diameter in \(S_{02}^k\).
This completes the proof of Theorem 6.

5 The Connectivity of Quantified Formulas

Definition 16. A quantified $B$-formula $\phi$ (in prenex normal form) is an expression of the form

$$Q_1 y_1 \ldots Q_m y_m \varphi(y_1, \ldots, y_m, x_1, \ldots, x_n),$$

where $\varphi$ is a $B$-formula, and $Q_1, \ldots, Q_m \in \{\exists, \forall\}$ are quantifiers. $x_1, \ldots, x_n$ are called the free variables of $\phi$.

For quantified $B$-formulas, we define the connectivity problems

- $\text{QBF-Conn}(B)$: Given a quantified $B$-formula $\phi$, is $G(\phi)$ connected?
- $\text{st-QBF-Conn}(B)$: Given a quantified $B$-formula $\phi$ and two solutions $s$ and $t$, is there a path from $s$ to $t$ in $G(\phi)$?

Theorem 17. Let $B$ be a finite set of Boolean functions.

1. If $B \subseteq M$ or $B \subseteq L$, then
   (a) $\text{st-QF-Conn}(B)$ and $\text{QBF-Conn}(B)$ are in $P$,
   (b) the diameter of every quantified $B$-formula is linear in the number of free variables.

2. Otherwise,
   (a) $\text{st-QBF-Conn}(B)$ and $\text{QBF-Conn}(B)$ are PSPACE-complete,
   (b) there are quantified $B$-formulas with at most one quantifier such that their diameter is exponential in the number of free variables.

Proof. 1. For $B \subseteq M$, any quantified $B$-formula $\phi$ represents a monotonic function: Using $\exists y \psi(y, x) = \psi(0, x) \lor \psi(1, x)$ and $\forall y \psi(y, x) = \psi(0, x) \land \psi(1, x)$ recursively, we can transform $\phi$ into an equivalent $M$-formula since $\land$ and $\lor$ are monotonic. Thus as in Lemma 8, $\text{st-QBF-Conn}(B)$ and $\text{QBF-Conn}(B)$ are trivial, and $d_f(a, b) = |a - b|$ for any two solutions $a$ and $b$.

For a quantified $B$-formula $\phi = Q_1 y_1 \ldots Q_m y_m \varphi$ with $B \subseteq L$, we first remove the quantifications over all fictional variables of $\varphi$ (and eliminate the fictional variables if necessary). If quantifiers remain, $\phi$ is either tautological (if the rightmost quantifier is $\exists$) or unsatisfiable (if the rightmost quantifier is $\forall$), so the problems are trivial, and $d_f(a, b) = |a - b|$ for any two solutions $a$ and $b$. Otherwise, we have a quantifier-free formula and the proof follows from Lemma 10.

2. Again as in Lemma 4, it follows that $\text{st-QBF-Conn}(B)$ and $\text{QBF-Conn}(B)$ are in PSPACE, since the formula value problem for quantified $B$-formulas is in PSPACE [SM73].

An inspection of Post’s lattice shows that if $B \not\subseteq M$ and $B \not\subseteq L$, then $|B| \supseteq S_{12}, |B| \supseteq D_1$, or $|B| \supseteq S_{02}$, so we have to prove PSPACE-completeness.
and show the existence of $B$-formulas with an exponential diameter in these cases.

For $|B| \supseteq S_{12}$ and $|B| \supseteq D_{1}$, the statements for the PSPACE-hardness and the diameter obviously carry over from Theorem 6.

For $B \supseteq S_{02}$, we proceeded again similar as in the proof of Lemma 13. We give a reduction from the problems for 1-reproducing 3-CNF-formulas: We map a 1-reproducing 3-CNF-formula $\phi$ and two solutions $s$ and $t$ of $\phi$ to a quantified $B$-formula $\phi'$ and two solutions $s'$ and $t'$ of $\phi'$ such that $s'$ and $t'$ are connected in $G(\phi')$ iff $s$ and $t$ are connected in $G(\phi)$, and such that $G(\phi')$ is connected iff $G(\phi)$ is connected.

For a 1-reproducing formula $\psi$ over the $n$ variables $x_{1},...,x_{n}$, we define a formula $T_{\psi} \in S_{02}$ with the two additional variables $y$ and $z$,

$$T_{\psi} = (\psi \land y) \lor z.$$ 

$T_{\psi}$ is 0-separating since all vectors that are no solutions have $z = 0$. Also, $T_{\psi}$ is 0-reproducing and still 1-reproducing. Here again, we use the algorithm $Tr$ from the proof of Lemma 13 to transform any 1-reproducing 3-CNF-formula $\phi$ into a $B$-formula $\phi'$ equivalent to $T_{\phi}$. Again, we show

$$T_{\wedge}[x_{1}/T_{\psi_{1}}, x_{2}/T_{\psi_{2}}] = (((\psi_{1} \land y) \lor z) \land ((\psi_{2} \land y) \lor z) \land y) \lor z$$

$$\equiv (((\psi_{1} \land y) \land (\psi_{2} \land y) \land y) \lor z$$

$$\equiv (\psi_{1} \land \psi_{2} \land y) \lor z = T_{\psi_{1} \land \psi_{2}}.$$ 

Now let

$$\phi' = \forall z \phi';$$

then, for any two solutions $s$ and $t$ of $\phi(x)$, $s' = s \cdot 1$ and $t' = t \cdot 1$ are solutions of $\phi'(x,y)$, and they are connected in $G(\phi')$ iff $s$ and $t$ are connected in $G(\phi)$, and $G(\phi')$ is connected iff $G(\phi)$ is connected.

The proof of Lemma 5 shows that there is an 1-reproducing $(n-1)$-ary function $f$ with diameter of at least $2\lfloor \frac{n-1}{2} \rfloor$. Let $f$ be represented by a formula $\phi$; then $\phi'$ is a quantified $B$-formula with $n$ free variables and one quantifier with the same diameter.

Remark 18. An analog to 17 also holds for quantified circuits as defined in [RW00, Section 7].

6 Conclusions and Future Work

While the classification for CSPs required an essential enhancement of Schaefer’s framework and the introduction of new classes of CNF(S)-formulas, for $B$-formulas and $B$-circuits the connectivity issues fit entirely into Post’s framework, although the proofs were quite novel, and made substantial use of Gopalan et al.’s results for 3-CNF-formulas.
Also, the results for $B$-formulas and $B$-circuits are much simpler: here we have one common dichotomy for the diameter and both connectivity questions, and satisfiability is always as least as easy as connectivity, while for CSPs, the connectivity problem exhibits a trichotomy, and $st$-connectivity can be easier than satisfiability (for safely tight, non-Schaefer sets of relations), see [Sch13].

As Gopalan et al. stated in [GKMP06], we also believe that “connectivity properties of Boolean satisfiability merit study in their own right”, which is substantiated by the recent interest in reconfiguration problems. Moreover, we imagine our results could aid the advancement of circuit based SAT solvers.

As mentioned in the introduction, besides CSPs, $B$-formulas and circuits, there are further ways to look at Boolean satisfiability, and investigating the connectivity in these settings might be worthwhile as well. E.g., disjunctive normal forms with special connectivity properties were studied by Ekin et al. already in 1997 for their “important role in problems appearing in various areas including in particular discrete optimization, machine learning, automated reasoning, etc.” [EHK99].

Other connectivity-related problems already mentioned by Gopalan et al. are counting the number of components and approximating the diameter. Also, with regard to reconfiguration problems, one could try to find the shortest path between two solutions, or the optimal path according to some measure.

Furthermore, our definition of connectivity is not the only sensible: One could regard two solutions connected whenever their Hamming distance is at most $d$, for any fixed $d \geq 1$; this was already considered related to random satisfiability, see [ART06]. This generalization seems meaningful as well as challenging.

Finally, a most interesting subject are CSPs over larger domains; in 1993, Feder and Vardi conjectured a dichotomy for the satisfiability problem over arbitrary finite domains [FV98], and while the conjecture was proved for domains of size three in 2002 by Bulatov [Bul02], it remains open to date for the general case. Close investigation of the solution space might lead to valuable insights here.

For $k$-colorability, which is a special case of the general CSP over a $k$-element set, the connectivity problems and the diameter were already studied by Bonsma and Cereceda [BC09], and Cereceda, van den Heuvel, and Johnson [CvdHJ11]. They showed that for $k = 3$ the diameter is at most quadratic in the number of vertices and the $st$-connectivity problem is in P, while for $k \geq 4$, the diameter can be exponential and $st$-connectivity is PSPACE-complete in general.

Acknowledgments.

I am grateful to Heribert Vollmer for pointing me to these interesting themes.
Figure 4: Post's lattice with our results for the connectivity problems and the diameter. For comparison, the satisfiability problem (without quantifiers) is NP-complete for the bold circled classes, and in P for the other ones.
| Class | Definition | Base |
|-------|------------|------|
| BF    | All Boolean functions | \{x \land y, \lnot x\} |
| R₀    | \{ f ∈ BF | f is 0-reproducing} | \{x \land y, x \lor y\} |
| R₁    | \{ f ∈ BF | f is 1-reproducing} | \{x \lor y, x \leftrightarrow y\} |
| R₂    | \text{R}_0 \cap \text{R}_1 | \{x \lor y, x \land (y \leftrightarrow z)\} |
| M     | \{ f ∈ BF | f is monotone} | \{x \land y, x \lor y, 0, 1\} |
| M₀    | M \cap \text{R}_0 | \{x \land y, x \lor y, 0\} |
| M₁    | M \cap \text{R}_1 | \{x \land y, x \lor y, 1\} |
| M₂    | M \cap \text{R}_2 | \{x \land y, x \lor y\} |
| S₀    | \{ f ∈ BF | f is 0-separating} | \{x \to y\} |
| S₀ⁿ   | \{ f ∈ BF | f is 0-separating of degree n\} | \{x \to y, \text{dual}(T^{n+1}_n)\} |
| S₁    | \{ f ∈ BF | f is 1-separating} | \{x \to y\} |
| S₁ⁿ   | \{ f ∈ BF | f is 1-separating of degree n\} | \{x \to y, T^{n+1}_n\} |
| S₀₂   | S₀ \cap \text{R}_2 | \{x \lor (y \land z), \text{dual}(T^{n+1}_n)\} |
| S₀₁   | S₀ \cap \text{M} | \{\text{dual}(T^{n+1}_n), 1\} |
| S₀₀   | S₀ \cap \text{R}_2 \cap \text{M} | \{x \lor (y \land z), \text{dual}(T^{n+1}_n)\} |
| S₀₀₀  | S₀ \cap \text{R}_2 \cap \text{M} | \{x \lor (y \land z)\} |
| S₀₁₂  | S₁ \cap \text{R}_2 | \{x \lor (y \lor z), T^{n+1}_n\} |
| S₁₂   | S₁ \cap \text{R}_2 | \{x \lor (y \lor z)\} |
| S₁₁   | S₁ \cap \text{M} | \{T^{n+1}_n, 0\} |
| S₁₁₀  | S₁ \cap \text{R}_2 \cap \text{M} | \{x \lor (y \lor z), 0\} |
| S₁₀   | S₁ \cap \text{R}_2 \cap \text{M} | \{x \lor (y \lor z), T^{n+1}_n\} |
| D     | \{ f ∈ BF | f is self-dual\} | \{\text{maj}(x, \lnot y, \lnot z)\} |
| D₁    | D \cap \text{R}_2 | \{\text{maj}(x, y, \lnot z)\} |
| D₂    | D \cap \text{M} | \{\text{maj}(x, y, z)\} |
| L     | \{ f ∈ BF | f is linear\} | \{x \lor y, 1\} |
| L₀    | L \cap \text{R}_0 | \{x \lor y\} |
| L₁    | L \cap \text{R}_1 | \{x \lor y\} |
| L₂    | L \cap \text{R}_2 | \{x \lor y\} |
| L₃    | L \cap D | \{x \lor y \lor z\} |
| E     | \{ f ∈ BF | f is constant or a conjunction\} | \{x \land y, 0, 1\} |
| E₀    | E \cap \text{R}_0 | \{x \land y, 0\} |
| E₁    | E \cap \text{R}_1 | \{x \land y, 1\} |
| E₂    | E \cap \text{R}_2 | \{x \land y\} |
| V     | \{ f ∈ BF | f is constant or a disjunction\} | \{x \lor y, 0, 1\} |
| V₀    | V \cap \text{R}_0 | \{x \lor y, 0\} |
| V₁    | V \cap \text{R}_1 | \{x \lor y, 1\} |
| V₂    | V \cap \text{R}_2 | \{x \lor y\} |
| N     | \{ f ∈ BF | f is essentially unary\} | \{\lnot x, 0, 1\} |
| N₂    | N \cap D | \{\lnot x\} |
| I     | \{ f ∈ BF | f is constant or a projection\} | \{x, 0, 1\} |
| I₀    | I \cap \text{R}_0 | \{x, 0\} |
| I₁    | I \cap \text{R}_1 | \{x, 1\} |
| I₂    | I \cap \text{R}_2 | \{x\} |

Table 1: List of all closed classes of Boolean functions with definitions and bases. 
\((T^n_k)\) denotes the threshold function, \(T^n_k(x_1, \ldots, x_n) = 1 \iff \sum_{i=1}^n x_i \geq k,\) and \(\text{dual}(f)(x_1, \ldots, x_n) = \overline{f(\overline{x_1}, \ldots, \overline{x_n})}\).
References

[ART06] Dimitris Achlioptas and Federico Ricci-Tersenghi, *On the solution-space geometry of random constraint satisfaction problems*, Proceedings of the thirty-eighth annual ACM symposium on Theory of computing, ACM, 2006, pp. 130–139. 15

[BC09] Paul Bonsma and Luis Cereceda, *Finding paths between graph colourings: Pspace-completeness and superpolynomial distances*, Theoretical Computer Science 410 (2009), no. 50, 5215–5226. 15

[BCRV03] Elmar Böhler, Nadia Creignou, Steffen Reith, and Heribert Vollmer, *Playing with boolean blocks, part i: Posts lattice with applications to complexity theory*, SIGACT News, 2003. 5, 9

[Bu102] Andrei A Bulatov, *A dichotomy theorem for constraints on a three-element set*, Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on, IEEE, 2002, pp. 649–658. 15

[CKZ08] Nadia Creignou, Phokion Kolaitis, and Bruno Zanuttini, *Structure identification of boolean relations and plain bases for co-clones*, Journal of Computer and System Sciences 74 (2008), no. 7, 1103–1115. 2

[CvdHJ11] Luis Cereceda, Jan van den Heuvel, and Matthew Johnson, *Finding paths between 3-colorings*, Journal of graph theory 67 (2011), no. 1, 69–82. 15

[EHK99] Oya Ekin, Peter L Hammer, and Alexander Kogan, *On connected boolean functions*, Discrete Applied Mathematics 96 (1999), 337–362. 15

[FM07] Zhaohui Fu and Sharad Malik, *Extracting logic circuit structure from conjunctive normal form descriptions*, VLSI Design, 2007. Held jointly with 6th International Conference on Embedded Systems., 20th International Conference on, IEEE, 2007, pp. 37–42. 2

[FV98] Tomás Feder and Moshe Y Vardi, *The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory*, SIAM Journal on Computing 28 (1998), no. 1, 57–104. 15

[GKMP06] Parikshit Gopalan, Phokion G. Kolaitis, Elitza N. Maneva, and Christos H. Papadimitriou, *The connectivity of boolean satisfiability: Computational and structural dichotomies*, ICALP’06, 2006, pp. 346–357. 1, 15

[GKMP09] Parikshit Gopalan, Phokion G. Kolaitis, Elitza Maneva, and Christos H. Papadimitriou, *The connectivity of boolean satisfiability: Computational and structural dichotomies*, SIAM J. Comput. 38 (2009), no. 6, 2330–2355. 1, 2, 4

[IDH+11] Takehiro Ito, Erik D. Demaine, Nicholas J. A. Harvey, Christos H. Papadimitriou, Martha Sideri, Ryuhei Uehara, and Yushi Uno, *On the complexity of reconfiguration problems*, Theor. Comput. Sci. 412 (2011), no. 12-14, 1054–1065. 2

[KMM11] Marcin Kaminski, Paul Medvedev, and Martin Milanic, *Shortest paths between shortest paths and independent sets*, Combinatorial Algorithms, Springer, 2011, pp. 56–67. 2

[Lew79] Harry R Lewis, *Satisfiability problems for propositional calculi*, Mathematical Systems Theory 13 (1979), no. 1, 45–53. 6

[Mic12] Thomas Michael, *On the applicability of post’s lattice*, Information Processing Letters 112 (2012), no. 10, 386–391. 6

[MMW07] Elitza Maneva, Elchanan Mossel, and Martin J Wainwright, *A new look at survey propagation and its generalizations*, Journal of the ACM (JACM) 54 (2007), no. 4, 17. 2
[MMZ05] Marc Mézard, Thierry Mora, and Riccardo Zecchina, *Clustering of solutions in the random satisfiability problem*, Physical Review Letters **94** (2005), no. 19, 197205. 2

[MTY07] Kazuhsa Makino, Suguru Tamaki, and Masaki Yamamoto, *On the boolean connectivity problem for horn relations*, Proceedings of the 10th international conference on Theory and applications of satisfiability testing, SAT’07, 2007, pp. 187–200. 1

[Pos41] Emil L Post, *The two-valued iterative systems of mathematical logic*(am-5), vol. 5, Princeton University Press, 1941. 5

[RW00] Steffen Reith and Klaus W Wagner, *The complexity of problems defined by boolean circuits*, 2000. 5, 14

[Sch78] Thomas J. Schaefer, *The complexity of satisfiability problems*, STOC ’78, 1978, pp. 216–226. 2

[Sch07] Henning Schnoor, *Algebraic techniques for satisfiability problems*, Ph.D. thesis, Universität Hannover, 2007. 5

[Sch13] Konrad W. Schwerdtfeger, *A computational trichotomy for connectivity of boolean satisfiability*, ArXiv CoRR abs/1312.4524 (2013). 1, 15

[SM73] L. J. Stockmeyer and A. R. Meyer, *Word problems requiring exponential time*, Proceedings of the Fifth Annual ACM Symposium on Theory of Computing, STOC ’73, 1973, pp. 1–9. 13

[Vol99] Heribert Vollmer, *Introduction to circuit complexity: A uniform approach*, Springer-Verlag New York, Inc., 1999. 5

[WLLH07] Chi-An Wu, Ting-Hao Lin, Chih-Chun Lee, and Chung-Yang Ric Huang, *Qutesat: a robust circuit-based sat solver for complex circuit structure*, Proceedings of the conference on Design, automation and test in Europe, EDA Consortium, 2007, pp. 1313–1318. 2

[Zve05] Igor’ E. Zverovich, *Characterizations of closed classes of boolean functions in terms of forbidden subfunctions and post classes*, Discrete Appl. Math. **149** (2005), no. 1-3, 200–218. 5