Does SAT exhibit fractal behavior?

Joost J. Joosten[†] and Grant Olney Passmore[‡]

March 20, 2008

Abstract

In this paper we study the structure of the set SAT of all satisfiable propositional logical formulas. In particular we raise the question whether the distribution of SAT within the set $\mathcal{A}$ of all propositional formulas exhibits fractal behavior. This answer is of course relative to a metric on $\mathcal{A}$. We show that for one such metric there is strong evidence that the distribution does indeed behave wildly. Next we look at an alternative metric.

Keywords: theory of computation, boolean satisfiability

1 Introduction

Physics is full of (differential) equations whose solutions are either not expressible in terms of known analytical functions or give directly rise to uncontrollable dynamical behavior. Poincaré took an interesting turn at tackling these problems. Rather than being interested in the full solution or trajectory of such impenetrable dynamical systems, he considered qualitative and topological properties of them instead (e.g. those properties that do not change under smooth changes of coordinates).

We would like to make an analogy to computability. The set SAT of all satisfiable sets of clauses in propositional logic is well defined and actually elementary decidable. However, we know that it is likely – depending on if $\mathsf{P} = \mathsf{NP}$ – to be very hard to decide whether or not some set of clauses $\mathcal{S}$ belongs to SAT.

Maybe, in analogy with Poincaré's aforementioned approach to physics, we can single out some qualitative properties of SAT and UNSAT. If we can isolate topologically well-behaved fragments of the set of all sets of clauses, this might, for instance, give rise to a probabilistic approach to deal with SAT membership.

2 Structuring the set of sets of clauses

Let $\mathcal{A}$ denote the set of all sets of clauses. At times when no confusion can arise, we shall call a set of clauses also a *formula*. We shall define a notion of distance on $\mathcal{A}$. To this end, let us introduce a relation $\mathsf{One}(\varphi, \psi)$ that defines when two formulas φ and ψ have distance 1 from each other. We define $\mathsf{One}(\varphi, \psi)$ in two steps: first by defining a relation $\mathsf{One}{*}(\varphi, \psi)$, and then deriving $\mathsf{One}(\varphi, \psi)$ as its symmetric closure.

Definition 2.1. *Let* $\varphi = \{S_1, \ldots S_n\}$ $(n \geq 0)$ *and* $\psi = \{S'_1, \ldots S'_m\}$ $(m \geq 0)$ *be some given formulas in* $\mathcal{A}$. *We define the relation* $\mathsf{One}{*}(\varphi, \psi)$ *to hold if one of the following applies.*

1. *If* $m = n$ *and* $S_i = S'_i$ *for all but one fixed* j *such that* $(S_j \setminus \{l\}) \cup \{\neg l\} = S'_j$ *for some literal* $l \in S_j$.
2. *If* $m = n$ *and* $S_i = S'_i$ *for all but one fixed* j *such that* $S_j = S'_j \cup \{l\}$ *for some literal* $l \in S_j$.
3. *If* $\varphi = \psi \cup \{\{l\}\}$ *for some unit clause* $\{l\} \in \varphi$.

Definition 2.2. *We define the relation* $\mathsf{One}(\varphi, \psi)$ *to be the symmetric closure of* $\mathsf{One}{*}(\varphi, \psi)$.

Once we have defined the one-step distance, we can extend this to all sets of clauses in the following way.

[†]Institute for Logic, Language and Computation, University of Amsterdam, NL; email: jjoosten@phil.uu.nl

[‡]Laboratory for Foundations of Computer Science, University of Edinburgh, UK; email: g.passmore@ed.ac.uk

Definition 2.3. *For given formulas φ and ψ, we define the distance between them $d(\varphi, \psi)$ to be the length minus one of the shortest sequence*

$$\langle \varphi = \chi_0, \chi_1, \dots, \chi_{m-1} = \psi \rangle$$

such that for each consecutive χ_i, χ_{i+1}, we have $\mathsf{One}(\chi_i, \chi_{i+1})$. We refer to consecutive members of a sequence with this property as being "single steps."

It is important to observe that d is indeed a total function on $\mathcal{A} \times \mathcal{A}$.

Lemma 2.4. *$d : \mathcal{A} \times \mathcal{A} \to \mathbb{N}$ is a total function.*

Proof. Given any fixed $\varphi, \psi \in \mathcal{A}$ it suffices to show that *at least one* path of single steps exists between $\varphi = \{S_1, \dots, S_n\}$ and $\psi = \{S'_1, \dots, S_m\}$. Consider the sequence that is built by removing each literal appearing in an S_i in φ one by one until φ has been transformed into the set containing only the empty clause. Then build up each S'_j in ψ step by step in the analogous manner. □

Lemma 2.5. *$d : \mathcal{A} \times \mathcal{A} \to \mathbb{N}$ is a metric function.*

Proof. It is clear that (i) $d(\varphi, \psi) \geq 0$, (ii) $d(\varphi, \psi) = 0$ iff $\varphi = \psi$, and (iii) $d(\varphi, \psi) = d(\psi, \varphi)$. The triangle inequality, $d(\varphi, \lambda) \leq d(\varphi, \psi) + d(\psi, \lambda)$ also holds for all $\varphi, \psi, \lambda \in \mathcal{A}$ as a path from φ to ψ and a path from ψ to λ can be composed to obtain a path from φ to λ. □

Thus, we obtain a topology on $\mathcal{A}$, $\langle \mathcal{A}, \mathcal{O} \rangle$, generated by the collection of open n-balls $\mathcal{O}^*$ defined as follows:

$$\mathcal{O}^* = \bigcup_{\varphi \in \mathcal{A}} \{ \bigcup_{n \in \mathbb{N}} \{ \{ \psi \in \mathcal{A} \mid d(\varphi, \psi) < n \} \} \}.$$

Definition 2.6. *We call a Cauchy sequence $c = \langle \varphi_0, \varphi_1, \dots \rangle$ trivial iff*

$$\exists n \in \mathbb{N} \; \forall m \in \mathbb{N} \; (m \geq n \Rightarrow \varphi_m = \varphi_{m+1}).$$

Lemma 2.7. *The only Cauchy sequences over $\langle \mathcal{A}, \mathcal{O} \rangle$ are trivial.*

We immediately obtain the following corollary.

Corollary 2.8. *$\langle \mathcal{A}, \mathcal{O} \rangle$ is a complete and separable metric space. Equivalently, $\langle \mathcal{A}, \mathcal{O} \rangle$ is a Polish space.*

Proof. Completeness follows from the fact that the triviality of a Cauchy sequence implies its limit is contained in the space. Separability follows simply by the fact that $\mathcal{A}$ is countable. □

Though Corollary 2.8 seems to imply that our space is quite nice topologically, it is rather not. In particular, we now observe that the topology generated by d is in fact *discrete*.

Theorem 2.9. *$\langle \mathcal{A}, \mathcal{O} \rangle$ is equivalent to the discrete topology on $\mathcal{A}$. That is, $\mathcal{O} = 2^{\mathcal{A}}$.*

Proof. The proof is easy, by showing that every singleton pointset is open. □

Nevertheless, d still endows $\mathcal{A}$ with some rather interesting and seemingly chaotic structure. We now turn to analysing the qualitative interaction between UNSAT and SAT under $\langle \mathcal{A}, \mathcal{O} \rangle$.

2.1 Unsatisfiable formulas

The simplest unsatisfiable formula is $\bot$, the set containing only the empty clause. It is easy to observe that all formulas distance one from $\bot$ are satisfiable. In this sense, we can see $\bot$ as a little island within a surrounding sea of satisfiable formulas. Let us make this intuitive notion precise and then address the question as to whether or not more islands exist.

Definition 2.10. *Let P be a property on $\mathcal{A}$ and let $x \in \mathcal{A}$ with $P(x)$. With $\vec{x}_P$ we denote the set of all points that can be reached from x using only distance one steps to other points with property P. This set is inductively defined as follows:*

1. *$P(x) \Rightarrow x \in \vec{x}_P$,*
2. *$y \in \vec{x}_P \;\; \wedge \;\; \mathsf{One}(y, z) \;\; \wedge \;\; P(z) \;\; \Rightarrow \;\; z \in \vec{x}_P$.*

We say that x and y are connected *if $y \in \vec{x}_P$. A set B is called connected whenever all of its points are.*

Definition 2.11. *Let P be a property on $\mathcal{A}$. We say that P contains an* island *if there is some x and y with $P(x), P(y)$, such that $x \notin \vec{y}_P$.*

Lemma 2.12. UNSAT *contains an island.*

Proof. Consider $\bot$, the set with just the empty clause. □

Lemma 2.13. *For any two $x, y \in$ UNSAT $\setminus \bot$ we have $x \in \vec{y}_P$.*

Proof. We can define a sequence of unsatisfiable formulas that bring us from x to y by taking only single steps as follows. Start with x, first add $\{p\}$ to x, where $\{p\}$ is some unit clause appearing in neither x nor y. Next, add $\{\neg p\}$ to x. Finally remove one by one the rest of x until $\{p\}$ and $\{\neg p\}$ are all that remain. As at all times, both $\{p\}$ and $\{\neg p\}$ are contained in the set of clauses, each intermediate formula in the sequence is unsatisfiable. Finally, build up y step by step. The last two steps will be to throw away the spurious $\{\neg p\}$ and $\{p\}$ to end with y. □

Corollary 2.14. UNSAT *consists of two islands.*

Theorem 2.15. SAT *is connected.*

Proof. Consider $x, y \in$ SAT. We shall define again a path in SAT from x to y consisting of only single steps. As x is satisfiable, we can find a satisfying assignment α. So, for each clause, there is at least one literal satisfied by α. In each clause in x, select one such satisfied literal. Now, from each clause throw away all non-selected literals by taking only single steps. At this stage, we have a set of unit clauses. Now, select one such unit clause, $\{p\}$, and reduce the remaining formula to contain only $\{p\}$ by taking single steps.

As y is also satisfiable, there is some assignment β satisfying it. It is clear that we can build up y from the single unit clause $\{p\}$ in a way similar to (but reverse of) the manner in which we came from x to $\{p\}$. If $\beta \models \neg p$ we might first need to flip the polarity in our unit clause $\{p\}$ if p occurs at all in some clause in y. This is admissible as flipping the polarity of a literal is indeed a single step (see case (1) of the definition of One$*$).

Note that the formula $\top$, that is the empty set, is also connected to any other satisfiable formula. □

Note that the distances between x and y using paths within SAT or UNSAT are linear in the sum of the number of literals appearing in x and y.

2.2 Metric relations between SAT and UNSAT

In the previous subsection we focused solely upon the topological structure of SAT and UNSAT. Let us now make some quantitative observations.

Lemma 2.16. $\forall\, x{\in}$SAT $\exists\, y{\in}$UNSAT $d(x, y) \leq 2$.

Proof. In two steps we adjoin two new unit clauses with complementary literals. □

So, in this sense the satisfiable formulas are thin within the unsatisfiable ones. The next lemma says that, in a certain sense, the set UNSAT is not thin in SAT.

Lemma 2.17. $\forall\, n{\in}\mathbb{N}\, \exists\, x{\in}$UNSAT $\forall\, y{\in}$SAT $d(x, y) \geq 2n$.

Proof. Consider the set consisting of n distinct pairs of complementary unit clauses. □

The combination of 2.16 and 2.17 suggests some wild topological structure on SAT, especially around the area where there are about as many satisfiable as non-satisfiable formulas, see [2] [1].

Question 2.18. *Is there a way to capture this wild structure?*

3 Towards a more general space of propositions

3.1 Probabilistic propositions

As we wish to employ analytical techniques to investigate SAT, we seek an isometric embedding of $\mathcal{A}$ into a continuous Polish space. We first define $\mathcal{P}$.

Definition 3.1. *Let $\mathcal{L}$ be the set of propositional literals. Let $\mathcal{P}^* = \mathcal{L} \times [0,1]$. If $x = \langle l, r\rangle \in \mathcal{P}^*$ we call x a "probabilistic literal." We define $\mathcal{P}$ to be the set of all sets of clauses φ built from probabilistic literals with the property that for all $l \in \mathcal{L}$,*

$$\{r \mid \exists S_i \in \varphi \ (\langle l, r\rangle \in S_i)\}$$

is a singleton.

We call members of $\mathcal{P}$ "probabilistic propositions." Thus, each $l \in \mathcal{L}$ exists in $\mathcal{P}^*$ continuum many times, once for each $r \in [0,1]$, and $\mathcal{P}$ is the set of all sets of clauses built from literals in $\mathcal{P}^*$ with the restriction that no classical propositional literal appears in the same probabilistic proposition with more than one real valuation attached to it. If $x = \langle l, r\rangle \in \mathcal{P}^*$, we let $\pi_0(x) = l$ and $\pi_1(x) = r$.

On the way to defining our *metric* $\partial : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$, we fix a bijection from between $\mathcal{L}$ and $\mathbb{N}$.

Definition 3.2. *Let $\gamma : \mathcal{L} \to \mathbb{N}$ be some bijection.*

We now recall a space familiar to functional analysts, the countable direct sum of Euclidean space.

Definition 3.3. *We first define its domain $\mathcal{E}_\infty$ and then derive its topology from the standard metric.*

$$\mathcal{E}_\infty = \oplus_{n\in\mathbb{N}} \mathbb{R}.$$

Now, members of $\mathcal{E}_\infty$ are simply infinite dimensional vectors with only finitely many nonzero real components. The standard metric $d_\infty : \mathcal{E}_\infty \times \mathcal{E}_\infty \to \mathbb{R}$ is the straight-forward generalization of the Pythagorean Theorem one would expect. That is, given $x, y \in \mathcal{E}_\infty$

$$d_\infty(x,y) = \sqrt[2]{\sum_{n\in\mathbb{N}} (x[n] - y[n])^2}.$$

Observe that since any two $x, y \in \mathcal{E}_\infty$ agree for all but finitely many components, $d_\infty(x,y)$ always converges. It is readily observed to be a metric.

As the members of $\mathcal{E}_\infty$ are simply infinite dimensional vectors with only finitely many nonzero real components, this space provides a wonderful framework for embedding our probabilistic propositions. This is because probabilistic propositions, while they may be of any arbitrary finite size, are themselves only each built from finitely many probabilistic literals. The next step towards constructing our metric will be to define the notion of a *literal valuation vector* for each $\varphi \in \mathcal{P}$.

Definition 3.4. *Given $\varphi \in \mathcal{P}$, we define the literal valuation vector for φ to be the (countably) infinite dimensional vector $\sigma(\varphi)$ with the following property:*

$$\sigma(\varphi)[n] = r \quad if \quad \exists S_i \in \varphi \ s.t. \ \exists \langle l, r\rangle \in S_i \ (\gamma(l) = n),$$

$$\sigma(\varphi)[n] = 0 \quad otherwise.$$

This vector is well-defined, since by definition if a probabilistic proposition $\varphi \in \mathcal{P}$ contains both $\langle l, r\rangle$ and $\langle l, r'\rangle$ in any of its clauses, then $r = r'$. It is easy to construct two different probabilistic propositions that nevertheless yield the same literal valuation vector. Our metric accommodates this fact by combining geometric aspects of the literal valuation vector together with topological aspects of the metric d.

We may now define our metric $\partial : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$.

Definition 3.5. *Let $\varphi, \psi \in \mathcal{P}$. We define the ∂-distance between φ and ψ, $\partial(\varphi, \psi)$, as follows:*

$$\partial(\varphi,\psi) = d_\infty(\sigma(\varphi), \sigma(\psi)) + d(\pi_{\mathcal{A}}(\varphi), \pi_{\mathcal{A}}(\psi)).$$

where $\pi_{\mathcal{A}} : \mathcal{P} \to \mathcal{A}$ is the "forgetful projection map" that takes a probabilistic proposition to its classical counterpart (by ignoring the literal valuations).

Theorem 3.6. *$\partial : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$ is a metric on $\mathcal{P}$.*

Proof. We verify each metric axiom in turn: (i)-(iii) are immediate. (iv) Given $\varphi, \psi, \lambda \in \mathcal{P}$, we have

$$\begin{aligned}
\partial(\varphi,\lambda) &= d_\infty(\sigma(\varphi),\sigma(\lambda)) + d(\pi_{\mathcal{A}}(\varphi),\pi_{\mathcal{A}}(\lambda)) \\
&\leq (d_\infty(\sigma(\varphi),\sigma(\psi)) + d_\infty(\sigma(\psi),\sigma(\lambda))) \\
&\quad + (d(\pi_{\mathcal{A}}(\varphi),\pi_{\mathcal{A}}(\psi)) + d(\pi_{\mathcal{A}}(\psi),\pi_{\mathcal{A}}(\lambda))) \\
&\leq (d_\infty(\sigma(\varphi),\sigma(\psi)) + d(\pi_{\mathcal{A}}(\varphi),\pi_{\mathcal{A}}(\psi))) \\
&\quad + (d_\infty(\sigma(\psi),\sigma(\lambda)) + d(\pi_{\mathcal{A}}(\psi),\pi_{\mathcal{A}}(\lambda))) \\
&\leq \partial(\varphi,\psi) + \partial(\psi,\lambda).
\end{aligned}$$

□

Theorem 3.7. *There is an isometry $\Pi : \mathcal{A} \to \mathcal{P}$.*

Proof. Let Π map each $\varphi \in \mathcal{A}$ to the equivalent formula over $\mathcal{P}$ in which each literal occurring in φ is assigned the valuation 0. □

Theorem 3.8. *$\langle \mathcal{P}, \partial \rangle$ is complete.*

Proof. Consider any ∂ Cauchy sequence $c = \langle x_0, x_1, \ldots \rangle$. If c is trivial, its limit must be in $\mathcal{P}$. Consider c nontrivial. As $\partial(x_i, x_{i+1}) = d_\infty(\sigma(x_i), \sigma(x_{i+1})) + d(\pi_\mathcal{A}(x_i), \pi_\mathcal{A}(x_{i+1}))$, it follows, as d takes values only in $\mathbb{N}$, that an index $n \in \mathbb{N}$ must exist s.t. $\forall m \in \mathbb{N}\ (m \geq n \to \pi_\mathcal{A}(x_m) = \pi_\mathcal{A}(x_{m+1}))$. That is, the Cauchy sequence $c' = \langle x_n, x_{n+1}, \ldots \rangle$ must converge solely w.r.t. the metric d_∞ upon the corresponding literal valuation vectors in the space $\mathcal{E}_\infty$, and moreover we see that the underlying classical proposition (e.g. that given by $\pi_\mathcal{A}(x_i)$ for each element of c') must after index n remain constant. It then follows that the number of nonzero entries in the literal valuation vectors for each element of c' must remain constant, and thus the limit w.r.t. d_∞ of the sequence of literal valuation vectors lies in $\mathcal{E}_\infty$. But then since for all $\vec{v} \in \mathcal{E}_\infty$ there exists some $\varphi \in \mathcal{P}$ s.t. $\sigma(\varphi) = \vec{v}$, we have

$$lim\ c \ = \ lim\ c' \ = \ lim\ \langle x_n, x_{n+1}, \ldots \rangle = \psi$$

s.t. $\forall i \in \mathbb{N}\ (\pi_\mathcal{A}(\psi) = \pi_\mathcal{A}(x_i)) \wedge \sigma(\psi) = lim\ \sigma(x_n)$. But, such a ψ exists uniquely in $\mathcal{P}$. Thus, every ∂ Cauchy sequence in $\mathcal{P}$ converges to a limit in $\mathcal{P}$. □

Theorem 3.9. *$\langle \mathcal{P}, \partial \rangle$ is separable.*

Proof. Consider the space of forumlas generated by the collection of probabilistic literals whose literal valuation vectors contain only rational values. □

Thus, our space $\langle \mathcal{P}, \partial \rangle$ is an uncountable continuous Polish space that admits an isometric embedding of our natural discrete metric space. We present this space as an avenue for future investigations.

References

[1] A. Alou, A. Ramani, I. Markov, K. Sakallah, *Solving Difficult Instances of Boolean Satisfiability in the Presence of Symmetry.* IEEE Transactions on CAD, 22(9):1117-1137, September 2003.

[2] P. Beame, *Phase Transitions in Proof Complexity and Satisfiability Search.* AAAS Annual Meeting, Seattle, February 2004.