WORDBOX: A case study in SAT Encodings

Victor S. Miller

IDA Center for Communications Research, Princeton, NJ 08540
USA

April 27, 2021

Abstract

The great effectiveness of SAT solvers for large, difficult, SAT problems is a recent welcome development. It is well know to practitioners in that field that the encoding of a problem can have a significant and radical effect on the practical solution time.

In particular, many combinatorial problems have a straightforward "easy" encoding which is logically sufficient to specify the problem. However, SAT solvers are very poor at inferring "global" knowledge about the problem. In fact, with some problems, such as the "Pigeon Hole Principal", in which global knowledge makes the problem trivial, branching SAT solvers must take exponential time. Examples of global knowledge are graph connectedness and parity. Therefore, in order to find an improved encoding global knowledge should be used.

In this note we consider the encoding a recreational Mathematics problem known as *Wordbox*. One is given a list of words, and an $m \times n$ rectangular grid. The object of the problem is to label the grid points with letters so that one can trace out all of the words in the list by moving adjacent grid points (using up/down, right/left moves). We show that this problem is an instance of the *labeled graph homomorphism* problem, in which we're given two finite undirected graphs G and H, along with labels $\ell(v)$ for every node of G. The object of the problem is to find a map $f : V(G) \to V(H)$ such that if $(v, v') \in E(G)$ then $(f(v), f(v')) \in E(H)$ (this is a *graph homomorphism*), and a labeling $\ell(w)$ of $w \in V(H)$ such that, for all $v \in V(G)$ we have $\ell(f(v)) = \ell(v)$.

We investigate a number of different encodings of the global knowledge and their effect on solving times for various SAT solvers.

1 A logical statement

We may state the problem as follows

$$\exists f \forall v, v' \in V(G)((v, v') \in E(G) \Rightarrow (f(v), f(v')) \in E(H)) \wedge (\ell(v) \neq \ell(v') \Rightarrow f(v) \neq f(v')), \tag{1}$$

where f is a function. However, in order to encode this is quantifier normal form we need to encode f as a binary relation.

So $w = f(x)$ means

$$\forall x \exists w R(x, w) \wedge \forall x, w, w'((w = w') \vee \neg R(w, x) \vee \neg R(x, w')).$$