Notation:

$H \implies C$ where H and C are expressions means $\neg H \lor C$.

$\langle x_1, x_2, \ldots, x_n \rangle$ referrs to the vector created by concatentating the literals x_1 through x_n.

$\langle X_1; X_2; \ldots; X_n \rangle$ referrs to the vector created by concatentating the vectors X_1 through X_n.

$(\lor X)$ where X is a the vector $\langle x_1, x_2, \ldots, x_n \rangle$ refers to $x_1 \lor x_2 \lor \cdots \lor x_n$.

$(\land X)$ and $(= X)$ are defined similarly.

$P \lor (= X)$ where P is an expression and X is a vector of literals means $(P \lor k_i \lor (\lor X)) \land (P \lor \neg k_i \lor (\lor X))$ where k_i is some new unique variable.

X^T is the transpose of X

$P \lor (= X)$ where P is an expression and X^T is vector of n vectors of literals means $(P \lor (= X_1^T)) \land (P \lor (= X_2^T)) \land \cdots \land (P \lor (= X_n^T))$.

$P \lor$

$[a, b]$ means b disjunctions with a elements each.

$[1, -b]$ means the intruduction of b free variables.

There are the following variables, some of which are held constant, in addition to helpers defined for equality [and possibly other purposes]:

The turing machine is stored in M. For each state q and tape charachter a (the tape characters are $True$ and $False$), $M_{q,a}$ is a vector containing the zero-padded binary representation of the state te machine will transition to, the character the machine will write, and the direction it will move ($True$ for increasing index, and $False$ for decreasing index) when presented with the state q and tape cell a.

The execution of the turing machine on each example e is stored in T_e, P_e, Q_e, R_e, W_e, and V_e representing the tape states, machine positions, machine states, read values, machine values, and movement directions, respectively. There is some redundancy in this list to increase clarity and computational efficiency.

The machine's start state is 0 and its accept state is 1. In the accept state, the machine writes the same tape value it reads, moves left (decreasing index), and remains in the same state.

So far, we have

$$4|Q|(\lceil \log_2 |Q| \rceil + 1) + \sum_{e \in \text{Examples}} (e_{\text{Time}} + 1)(2e_{\text{Memory}} + \lceil \log_2 |Q| \rceil)$$

variables, where $|Q|$ is the number of states.

We now introduce some constant constraints. $\ldots [1, ?]$

We now introduce a set of constraints for all $e \in$ Examples:

We now introduce the following constraints on proper tape values. For all $t \in \{1, 2, \ldots, e_{\text{Time}}\}$, and $x \in \{1, 2, \ldots, e_{\text{Memory}}\}$,

$$(P_{e,t,x} \implies T_{e,t,x} = R_{e,t} = W_{e,t-1}) \wedge (\neg P_{e,t,x} \implies T_{e,t,x} = T_{e,t-1,x})$$

This produces $(e_{\text{Time}} e_{\text{Memory}}) \cdot ([1, -2] + [4, 2] + [5, 2])$ terms.

We now introduce the following constraint on proper machine opperation. For all $t \in \{0, 1, \ldots, e_{\text{Time}} - 1\}$, $q \in B(Q)$, and $a \in \{True, False\}$, where $B(Q)$ is the set of zero padded binary representations of the numbers 0 through the number of states minus 1,

$$(\langle Q_{e,t}; \langle R_{e,t} \rangle\rangle = \langle q; \langle a \rangle\rangle) \implies (M_{q,a} = \langle Q_{e,t+1}; \langle W_{e,t}, V_{e,t} \rangle\rangle)$$

This produces $(2 e_{\text{Time}} |Q| \lceil \log_2 |Q| \rceil) \cdot ([\Theta(1), \Theta(1)])$ terms.

We now introduce the following constraints on proper machine movement. For all $t \in \{0, 1, \ldots, e_{\text{Time}} - 1\}$, and $x \in \{1, 2, \ldots, e_{\text{Memory}} - 1\}$,

$$(P_{e,t,x} \wedge V_{e,t} \implies P_{e,t+1,x+1}) \wedge (P_{e,t+1,x} \wedge \neg V_{e,t} \implies P_{e,t,x+1})$$

This produces $(e_{\text{Time}} e_{\text{Memory}}) \cdot ([\Theta(1), \Theta(1)])$ terms.

We now introduce the following constraint on unique machine position. For all $t \in \{0, 1, \ldots, e_{\text{Time}} - 1\}$, $x \in \{1, 2, \ldots, e_{\text{Memory}}\}$, and $y \in \{1, 2, \ldots, e_{\text{Memory}}\}$ such that $y \neq x$,

$$\neg P_{e,t,y} \vee \neg P_{e,t,y}$$

This produces $(\sum_{e \in \text{Examples}} e_{\text{Time}} \binom{e_{\text{Memory}}}{2}) \cdot [2, 1]$ terms.

In all we end up with

$$\Theta\left(\sum_{e \in \text{Examples}} e_{\text{Time}} \cdot e_{\text{Memory}}\right)$$

variables, constrained by a disjunction of

$$\Theta\left(\sum_{e \in \mathrm{Examples}} (e_{\mathrm{Time}} \cdot (e_{\mathrm{Memory}} + |Q| \log |Q|)\right)$$

three literal conjunctions, and

$$\Theta\left(\sum_{e \in \mathrm{Examples}} (e_{\mathrm{Time}} \cdot (e_{\mathrm{Memory}})^2\right)$$

two literal conjunctions. (possibly all in horn form, not sure yet.)