Theoretical Formulation of the General Reverse Ising Problem and Fast Algorithmic Solutions

Isaac Martin and Andrew Moore

University of Texas at Austin

July 27, 2023

Circuits

Fix the following data:

- X some arbitrary index set
- $\Sigma = \{-1, 1\}$ the set of possible spin states (switch between $\{0, 1\}$ and $\{-1, 1\}$ conventions via $x \longmapsto 2x - 1$).

We define a **circuit** to be a tuple (N, M, f) where

- $N, M \subseteq X$ are arbitrary finite disjoint subsets of X. We call N the collection of *input* indices or vertices and M the collection of *output* indices/vertices.
- $f : \Sigma^N \longrightarrow \Sigma^M$ is an arbitrary function – the *logic* of the circuit.

Additionally,

- Σ^X is the collection of functions $X \longrightarrow \Sigma$. It is isomorphic to $\overbrace{\Sigma \times \ldots \times \Sigma}^{|X| \text{ times}}$; tuples valued in Σ with $|X|$ many components. We call Σ^X the **spin space** or **state space** of X.
- Σ^N is the **input spin space**.
- Σ^M is the **output spin space**.

Ising Systems

$$H(x_1, x_2) = 2x_1 + 2x_2 - 3x_1x_2$$

An **Ising system** is a pair (X, H), often referred to as simply X, where

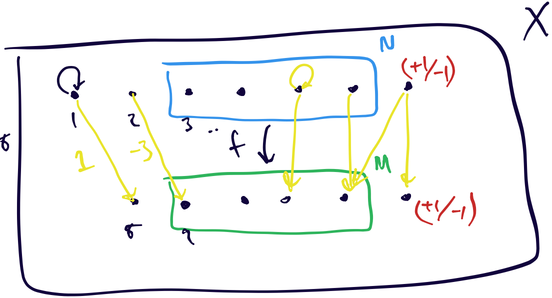
- X is the set from the previous slide, often a subset of $\mathbb{N}$
- $H \in \mathbb{R}[X]$ is a multilinear quadratic polynomial in elements of X called the *Hamiltonian* of the Ising system.

$$H : \Sigma^X \to \mathbb{R}$$

The **state space** of X is Σ^X. An Ising system X in state $\sigma \in \Sigma^X$ has energy $H(\sigma)$ given by evaluating the Hamiltonian at σ.

Picture of Circuits and Systems

$$H(\vec{x}) = x_1 + x_1x_8 - 3x_1x_9$$



(N, M, f) (X, H)

Reverse Ising Problem: Solving Circuits with Ising Systems

Given a circuit (N,M,f) we would like to design Ising systems (X,H) with the following features:

(1) A subset $N \subseteq X$ of spins whose state can be fixed

(2) A subset $M \subseteq X$ whose states vary freely with dynamics and will be read off after stabilizing

(3) For every $\sigma_N \in \Sigma^N$, the most likely spin state in $\sigma_M \in \Sigma^M$ is $f(\sigma_N)$.

Accomplishing this is called **solving the reverse Ising problem** for (N,M,f). Stated another way:

Reverse Ising Problem: Decompose X as $X = N \cup M \cup A$ so that $\Sigma^X = \Sigma^N \times \Sigma^M \times \Sigma^A$. Given an abstract circuit (N,M,f), design an Ising system (X,H) such that for every choice of input state $\sigma \in \Sigma^N$, there is some $\eta \in \Sigma^A$ such that

$$H(\sigma, f(\sigma), \eta) < H(\sigma, \omega, \eta') \text{ whenever } \omega \neq f(\sigma).$$

In this situation we say that (X,H) *solves* the circuit (N,M,f).

More terminology

- We often let $A = X \setminus (N \cup M)$ be the set of **auxiliary** spins.

- For each input spin state $\sigma \in \Sigma^N$ we call $\{\sigma\} \times \Sigma^{M \cup A}$ the *input level* of σ. This is verbally useful – an Ising system (X, H) solves a circuit (N, M, f) if the correct output $f(\sigma)$ is the *minimizer of its input level.*

Input: $[-1, +1]$

Input level: $[-1, +1, *, *, *, *]$

Examples

AND

$N = \{1, 2\}$

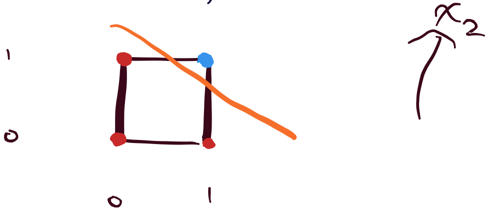
$M = \{3\}$

$f : \Sigma^N \rightarrow \Sigma^M$

$f(x_1, x_2) = 4x_1x_2 - 2x_1 - 2x_2 + 1$

$X = N \cup M$

$H(x_1, x_2, x_3) = x_3 - x_1x_3 - x_2x_3$



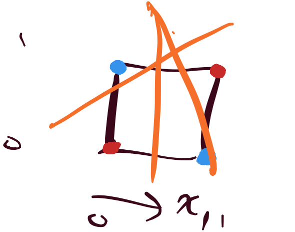
XOR

$N = \{1, 2\}$

$M = \{3\}$

$X = N \cup M \cup \{4\}$

$H(x_1, x_2, x_3, x_4)$ exists



Dynamics and our Goal

We are only interested in finding Ising systems such that the correct answer of each input minimizes the Hamiltonian on its input level. This does *not* necessarily yield good dynamics, but it is a necessary condition for good dynamics to be present.

Our Goal: Find robust methods for algorithmically finding Ising systems which solve arbitrary circuits. Work on dynamics later.

XOR is the first example of a circuit which is infeasible without auxiliaries. This is common – an Ising Hamiltonian only has access to quadratic terms and is hence not especially expressive.

What would a **higher degree Hamiltonian** look like?

$$H(\sigma) = \underline{\underline{4\sigma_1 + 3\sigma_2}} - \underline{4\sigma_1\sigma_2 + 5\sigma_2\sigma_3}$$

$$\sigma \in \Sigma^X \qquad + \underline{3\sigma_1\sigma_2\sigma_4} - 5\sigma_1\sigma_2\sigma_4\sigma_6$$

All Circuits are Solvable via Higher Degree Hamiltonians

Proposition 2.1

Any circuit (N, M, f) can be solved without auxiliaries by a multilinear polynomial H of high enough degree.

Key Fact: Any pseudo boolean function $g : \Sigma^X \longrightarrow \mathbb{R}$ can be uniquely represented by a multilinear polynomial [1].

Proof:

$$g(\sigma_N, \sigma_M) = \text{hamming}(\sigma_M, f(\sigma_N))$$

hamming distance

- has minima exactly at correct answers
- is pseudo-boolean

$\Longrightarrow$ $\exists$ some multi-lin poly H such that

$$H(\sigma) = g(\sigma)$$

Fourier/Hadamard Transform

Note: Given a pseudo-Boolean function $g : \Sigma^X \longrightarrow \mathbb{R}$, the unique multilinear polynomial representation of g is actually the *Hadamard Transform of g*, a type of generalized Fourier transform.

Rosenberg Reduction

These higher degree Hamiltonians can actually be reduced to quadratic polynomials at the cost of adding auxiliary variables using **Rosenberg reduction.**

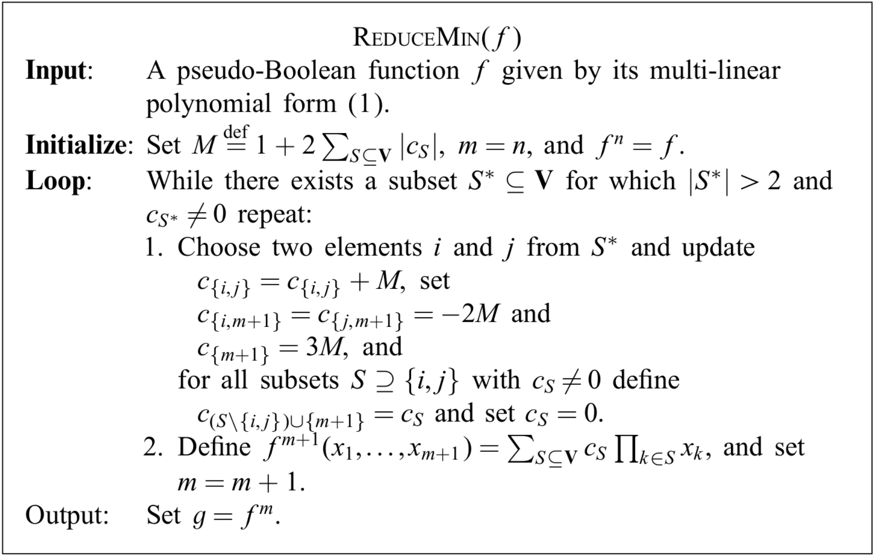
Observation: For $x, y, z \in \{0, 1\}$ the following equivalences hold:

- $xy = z$ iff $xy - 2xz - 2yz + 3z = 0$
- $xy \neq z$ iff $xy - 2xz - 2yz + 3z > 0$.

Rosenberg reduction works by replacing products with new auxiliary variables and penalizing “incorrect” values of the new variables.

Example: Let $f(x_1, x_2, x_3) = x_1x_2x_3$. This has minimum value at $x_1 = x_2 = x_3 = 0$.

Full Rosenberg Algorithm



```
REDUCEMIN($f$)
Input:      A pseudo-Boolean function $f$ given by its multi-linear
            polynomial form (1).
Initialize: Set $M \overset{\text{def}}{=} 1 + 2\sum_{S\subseteq \mathbf{V}} |c_S|$, $m = n$, and $f^n = f$.
Loop:       While there exists a subset $S^* \subseteq \mathbf{V}$ for which $|S^*| > 2$ and
            $c_{S^*} \neq 0$ repeat:
            1. Choose two elements $i$ and $j$ from $S^*$ and update
                 $c_{\{i,j\}} = c_{\{i,j\}} + M$, set
                 $c_{\{i,m+1\}} = c_{\{j,m+1\}} = -2M$ and
                 $c_{\{m+1\}} = 3M$, and
               for all subsets $S \supseteq \{i,j\}$ with $c_S \neq 0$ define
                 $c_{(S\setminus\{i,j\})\cup\{m+1\}} = c_S$ and set $c_S = 0$.
            2. Define $f^{m+1}(x_1,\ldots,x_{m+1}) = \sum_{S\subseteq \mathbf{V}} c_S \prod_{k\in S} x_k$, and set
               $m = m + 1$.
Output:     Set $g = f^m$.
```

All circuits solvable with auxiliaries

Proposition 2.2

Any circuit (N, M, f) can be solved by an Ising system (X, H) with finitely many auxiliaries.

Notice that this lemma says nothing about the *number* of auxiliary spins needed to solve a circuit; in general, it can be quite large.

Proof. Take the hamming objective function from before:

$$g(\sigma_N, \sigma_M) = \text{hamming}(\sigma_M, f(\sigma_N)).$$

From this we obtain a higher degree Hamiltonian P via the Hadamard transform. By applying the Rosenberg reduction algorithm, we can obtain a quadratic multilinear polynomial H in finitely many more variables than P which shares the same global minim as P. □

Observations

Rosenberg reduction of higher-degree Hamiltonian provides an algorithm for solving arbitrary Ising circuits. However, it is horribly inefficient in the number of auxiliary spins added.

Observations

Rosenberg reduction of higher-degree Hamiltonian provides an algorithm for solving arbitrary Ising circuits. However, it is horribly inefficient in the number of auxiliary spins added.

- The use of the objective function is unnecessary, a multilinear polynomial can be fit directly

Observations

Rosenberg reduction of higher-degree Hamiltonian provides an algorithm for solving arbitrary Ising circuits. However, it is horribly inefficient in the number of auxiliary spins added.

- The use of the objective function is unnecessary, a multilinear polynomial can be fit directly
- Refitting polynomial after addition of auxiliary variables often yields a quadratic Hamiltonian before full reduction is complete – only a subset of reductions are necessary

Observations

Rosenberg reduction of higher-degree Hamiltonian provides an algorithm for solving arbitrary Ising circuits. However, it is horribly inefficient in the number of auxiliary spins added.

- The use of the objective function is unnecessary, a multilinear polynomial can be fit directly
- Refitting polynomial after addition of auxiliary variables often yields a quadratic Hamiltonian before full reduction is complete – only a subset of reductions are necessary
- All auxiliaries added by Rosenberg reductions are AND gates and the penalty term is a valid Ising Hamiltonian for the AND gate

Observations

Rosenberg reduction of higher-degree Hamiltonian provides an algorithm for solving arbitrary Ising circuits. However, it is horribly inefficient in the number of auxiliary spins added.

- The use of the objective function is unnecessary, a multilinear polynomial can be fit directly
- Refitting polynomial after addition of auxiliary variables often yields a quadratic Hamiltonian before full reduction is complete – only a subset of reductions are necessary
- All auxiliaries added by Rosenberg reductions are AND gates and the penalty term is a valid Ising Hamiltonian for the AND gate
- Starting with a multilinear polynomial is unnecessary – one can simply do a greedy search through all possible AND gates between existing spins to solve the circuit.

Observations

Rosenberg reduction of higher-degree Hamiltonian provides an algorithm for solving arbitrary Ising circuits. However, it is horribly inefficient in the number of auxiliary spins added.

- The use of the objective function is unnecessary, a multilinear polynomial can be fit directly
- Refitting polynomial after addition of auxiliary variables often yields a quadratic Hamiltonian before full reduction is complete – only a subset of reductions are necessary
- All auxiliaries added by Rosenberg reductions are AND gates and the penalty term is a valid Ising Hamiltonian for the AND gate
- Starting with a multilinear polynomial is unnecessary – one can simply do a greedy search through all possible AND gates between existing spins to solve the circuit.

Rosenberg reduction is equivalent to "gluing" AND circuits onto the existing circuit until enough auxiliaries are present to solve the system.

Question: Can other circuits besides AND be used in a solution of this style?

Theorem 3.1 (I. Martin, A. Moore)

*Let (N, M, f) again be an abstract circuit. There exists an Ising system which solves this circuit if and only if there is some function $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$ (called an **auxiliary function**) such that both*

(a) *the new circuit $(N \cup M, A, F)$ is solvable by an Ising system with Hamiltonian R with the following additional property:*

$$R(\sigma_N, \sigma_M, F(\sigma_N, \sigma_M)) \geq R(\sigma_N, f(\sigma_N), F(\sigma_N, f(\sigma_N))) \tag{†}$$

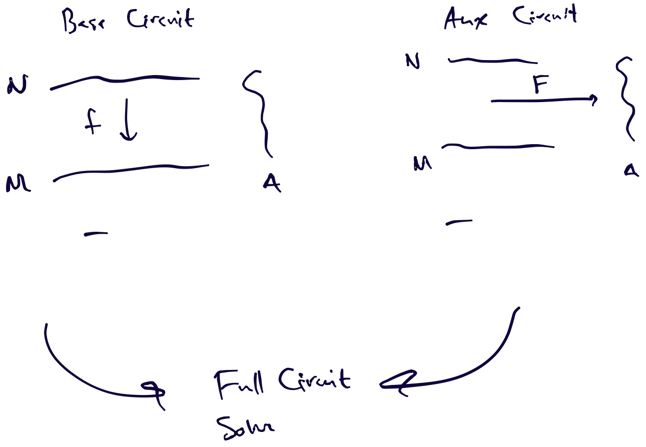
*for all σ_N and σ_M. We call this the **weak neutralizability condition.** (If the inequality is instead an equality, we call this the **strong neutralizability condition.** The system (X, R) is the **auxiliary system** and the circuit $(N \cup M, A, F)$ is the **auxiliary circuit**.*

(b) *there is an Ising system (X, S) which satisfies F-augmented constraints:*

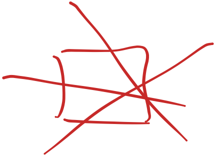
$$S(\sigma_N, \sigma_M, F(\sigma_N, \sigma_M)) > H(\sigma_N, f(\sigma_N), F(\sigma_N, f(\sigma_N))).$$

*We call (X, S) the **base system** and the circuit (N, M, f) the **base circuit**.*

Question: What are simple examples of neutralizable auxiliary functions F?


Base Circuit
N
f
M
A
—
Aux Circuit
N
F
M
A
—
Full Circuit
Solve

Linear Separability, SVM and Threshold Functions



The following are equivalent:

- Ising circuit with one output bit
- Linear SVM on input patterns
- Hyperplane separation of hypercube vertices
- Threshold function

Threshold functions depending only on input spins are always strongly neutralizable.
The condition is not well understood in general.
But we can generate libraries of strongly neutralizable threshold functions with fixed numbers of inputs.

Observation: The theorem guarantees that picking auxiliaries this way gives LP problem difficulty $\mathcal{O}(A^2 2^N)$ instead of $\mathcal{O}(2^{N+A})$.

Results and Stuff

Our practical method:

- We have described a function class from which to sample auxiliary spins
- Set heuristic function on this class, use any search algorithm (greedy etc.)
- Current results: $\boxed{\text{4x4 multiplication with 12 auxiliaries}}$ (computes in under an hour on a desktop machine)

Previously largest is 3x4 w/ 6 aux
best 4x4 had 25 aux

Next steps:

- Get better and faster heuristic functions
- Find domain optimized search algorithms
- Find theory to deal with more complex auxiliary function classes

References

[1] Peter L. Hammer Endre Boros. "Pseudo-Boolean optimization". In: Discrete Applied Mathematics 123 (2002), pp. 155–225. doi: 10.1016/S0166-218X(01)00341-9.