Sudoku Simulated Annealing: Duality of Non-Symmetric 9-State Potts Model

Analysis of Frustrated Lattice effects from Varying Sudoku Difficulties

Cormac Butler
Student Number: 15317164
Amateur Coder
butlerc9@tcd.ie

Tom Archer
PY3C03 Professor
Not-So-Amateur Coder
archert@tcd.ie

Abstract

In this experiment the metropolis algorithm combined with simulated annealing was used to investigate the change in energy of a 2D ising model. This code was then adapted and was directly applied to the solution of Sudoku puzzles of various difficulties. Comparisons between the 9-State Potts model are also evident from the second order phase change which occurs at a temperature of (0.75 ± 0.15)

Ising Model Diagram

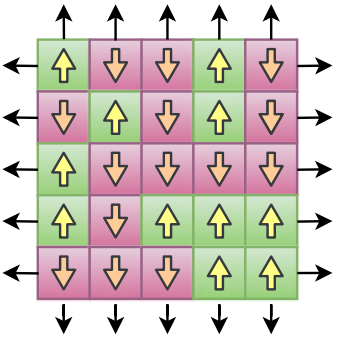

Figure 1: *Diagram of Ising Model Representation of Lattice. At each point is a particle with either a positive (green) or negative (purple) magnetic moment. Ideally this lattice would stretch off to infinity in every direction.*

0.1 Ising Model

The Ising Model (IM) is a mathematical model used to simulate the magnetic properties of ferromagnetic materials. The IM is made up of discrete variables stored in a list (1-D), matrix (2-D) or tensor ($\geq$3D) of which the entries have a value of ± 1 and represent the individual magnetic moments of particles comprising a solid lattice. Each lattice point is assigned in index i/j. In the Ising model these particles/atoms interact with every other spin in the lattice. The Hamiltonian $\hat{\mathcal{H}}$ of the lattice is given by

$$\hat{\mathcal{H}} = -J \sum_{<ij>} S_i S_j - h \sum_i S_i \qquad (1)$$

where: J = Exchange Energy
i, j = indices of the ith/jth particles in lattice
Si = Spin Configuration of the ith particle
h = External Magnetic field Strength

By examining equation 1 a few observations can be made. If J is taken to be a positive constant $>$ 1 (which it is for ferromagnetic materials) then if spins S_i & S_j have the same direction of magnetic moment the overall contribution from the first term will be negative. This negative contribution leads to a favourable decreasing energy. Thus a positive coupling constant J leads to the formation of domains of similar spin. This formation of domains is present in real magnetic materials and was first theorized in 1904[5]. The reason why entire magnetic materials don't form one large domain is nucleation. Domains of opposite spin are equally likely to form (in the presence of no external magnetic field) and so when two magnetic regions start to form in separate regions they grow larger and larger until the meet forming a boundary which is constantly morphing while both region attempt to convert the other one to their spin. **This is seen in magnetic materials such as chromium. Which is an incredibly magnetic atom but forms small domains with no net magnetization present across the entire crystal.**

This formation of magnetic domains is behaviour ex-

hibited in actual magnets. This will be a running theme across this experiment. The use of simple rule sets to model macro behaviour of more complex systems. Here the simplistic IM can exhibit behaviours such as a second order phase change ($\geq 2D$), a curie temperature and hysteresis.

0.2 Q-State Potts Model

The Q-State Potts model is a generalization of the IM. Instead of the binary 1,-1 spins present in the IM the Q-state Potts Model has a range of Q options from -1 to 1. For example a 5 state Potts model would have the spins [1, 0.5, 0, -0.5 , -1]. To be specific this is the q-vector Potts Model. The true value of this model lies in the fact that it is exactly solvable. The Hamiltonian of the Potts model is therefore given by:

$$H = -J \sum_{i,j=nm(i)} \delta_{S_i,S_j} \tag{2}$$

$$\delta_{S_i,S_j} = \begin{cases} 1, & \text{if } S_i = S_j, \\ 0, & \text{if } S_i \neq S_j. \end{cases}$$

As in the IM, A positive coupling constant J will lead to the formation of domains. These domains of equal values i.e. spin zero will be more likely to generate more spin zeros near it.

0.3 Sudoku

Sudoku is an old puzzle played on a 9x9 grid where numbers on each row and column must consist of the numbers 1 to 9 (though not necessarily chronologically). In addition to this the 81 tile grid is split into 9 smaller square boxes of 9 tiles each and also must contain the numbers 1 to 9.*

Similarity between Sudoku and 9-State Potts Model

Just as neighbouring lattice sites in the Ising model affect each other so too do connected Sudoku grid sites (which will now also be referred to as lattice sites). In the basic Potts Model the only contributing factor for their interaction is whether they have identical values. In this way the numbers 1-9 and the Potts model could be exchanged with the letters A,B,C... with the formation of say B domains. Moreover, Sudoku also could in principle also be played using the letters A to I (9 letters). This is of course only relevant for domain formation and ignores the physical meaning of the spins, from which observable such as magnetization are calculated.

The Q-State Potts Model also has an energy associated with its configuration as was discussed in the previous section. As with all physical systems, it will tend towards a low energy configuration until it reaches a meta-stable state. As outlined later a certain energy can also be assigned to a randomly filled Sudoku puzzle. The solution of the puzzle should be the lowest energy configuration and the entire puzzle filled in with one number should be the highest energy as this will correspond to the maximum amount of conficts withing the lattice.

In this way we could say that Sudoku acts as an anti-ferromagnetic material, trying to minimize the number of similar adjacent entries, with entries being compared to similar neighbouring spin configurations.

1 Methods

1.1 Metropolis Algorithm

The Metropolis Algorithm (MA) is a method used in statistical physics to generate random samples from where direct sampling may be difficult[5]. In this simulation the MA is used to greatly reduce the necessary calculations required to simulate the lattice.

In this experiment the variation of the MA is that to evolve the system through time random points are chosen within the lattice. This point will either have spin up spin down. **For this algorithm only directly adjacent points will be considered i.e. the particles to the left & right and up & down**. This approximation work well as the particles directly adjacent will be the most strongly interacting.

The MA was implemented by first generating a random $n \times n$ matrix for the IM. Equation 1 was used to

*A comprehensive guide to the rules and effective strategy of Sudoku can be found at www.sudokuessentials.com how_to_play_sudoku.html

calculate the initial energy of every particle and hence the energy of the system. The MA then randomly selects a lattice point i,j ($0 \leq i, j < n$). It then evaluates the energy change associated with potentially flipping the lattice site spin.

$$\Delta H = -2JS_i \sum_j S_j \tag{3}$$

If the energy will decrease then the spin will be flipped. If the energy of the lattice will increase it will only be accepted with the probability P:

$$P \equiv e^{-\frac{\Delta H}{kT}} \tag{4}$$

where: T = Temperature of System
k = Boltzmann Constant
ΔH = Energy Change of Spin Flip

This random sampling of lattice sites is iterated to evolve the system and lead to the formation of the domains seen later. The acceptance of less favourable energy configurations of the lattice is not a trivial choice but is of vital importance. If the lower energy configuration is constantly chosen, the system will likely not reach its lowest possible energy state (i.e. the global minimum). It may instead get stuck in one of the many local minima in which it cannot escape by the definition of a minimum it must increase.

A simple way to improve your chances of decreasing the energy of your configuration easily is to accept the change when then new energy is **less than or equal to** the old energy change. This addition of the equal to clause allows also for the lateral movement of the configuration across equal energies. So if a system gets stuck in a flat region such as that in figure 2 it can cross it to get to a potential lower minimum.

1.2 Simulated Annealing

Simulated annealing in its simplest context is a method of enhancing a basic hill climbing algorithm to more closely approximate the global maximum. As discussed energy changes may be accepted with a certain probability P as given in equation 4. This depends

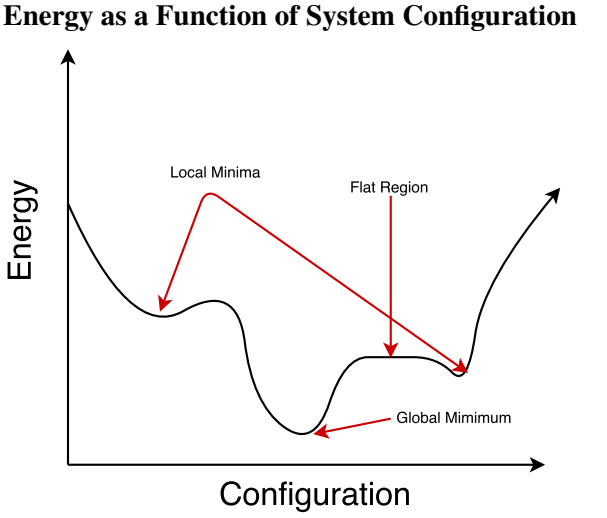

Figure 2: *Figure shows a simplified example of an energy surface with local and global mimima. A region of equal energy is also shown and labelled "flat region".*

on the temperature T of the system. A higher temperature will lead to an increase in the probability while lower temperatures will decrease the probability.

The cleverness of this system lies in the energy changes when escaping a global/close to global minimum vs a shallow local minimum. A larger ΔH will decrease the probability of a unfavourable energy change while a low ΔH will increase the likely hood of eventually coming out of this minimum. It is for this reason that when you're in a global/close to global minimum you are more unlikely to escape.

The chances of further decreasing the energy at the end of the simulation are further increased by slowly lowering the the temperature. **This is known as annealing**. The system starts at a high temperature T_0 and lowers in temperature. This is done so that at high temperatures the system can settle in the region of the global minimum and as the temperature is lowered it then can settle in the lowest part of the minimum.

The rate at which it is cooled is very important to the overall success of the simulated annealing. The rate,initial temperature and manner in which the temperature is lowered is known as the **Cooling Schedule** and is a very well studied branch of physics. In

this investigation 3 different cooling schedules were examined.

Equation 5 follows a linear cooling schedule. This schedule depends on the initial temperature and the slope m at which the temperature descends. Equations 6,7 show a very similar cooling profiles with equation 6 showing an exponential cooling schedule[†]. While equation 7 shows a logarithmic decay similar to a geometric series. In this case the value of R should be $0 < R < 1$ so that the temperature does converge to zero.

$$T(t) = mt + T_0 \tag{5}$$
$$T(t) = T_0 e^{-rt} \tag{6}$$
$$T(t) = T_0 (R)^t \tag{7}$$

There are entire papers devoted to the efficient choosing of these variables in the above equations. In this experiment equation 7 was used as the other two equations seemed to be more difficult to select the right parameters.

1.3 Evaluating Sudoku Energy and Lattice Evolution

The Metropolis Algorithm and Simulated Annealing can also be applied to solve a Sudoku Puzzle which in turn will give us additional insights into Ising Model. **The main obstacle in implementing this technique is to assign the Sudoku Puzzle an associated energy.** This must be done in a clever way to minimize the number of possible permutations of the Sudoku Puzzle.

There should also be a ground state of the Sudoku Puzzle which will be set arbitrarily at an energy of 0[‡]. This was done this by first inputting a given Sudoku Puzzle with some entries filled in and the empty entries entered as zeros.

The program then randomly fills in the remaining zeros with the numbers from 1-9. However, this is done on a box by box basis. So for example if the first box contains the numbers 1 and 9 already filled in, the program will then fill in the remaining empty squares with the numbers 2 to 8 in a random order. This ensures the initialized puzzle has each box containing the numbers 1 to 9.

The initial energy of the puzzle is calculated by examining the the number of non-unique entries in a given row and column. The boxes already contain the numbers 1-9 and so do not need to be factored in. The energy of each row and row and column is then summed to generate the energy. The Hamiltonian of the Sudoku puzzle is therefore given by

$$H = 162 - \sum_{i,j} \delta_{Si,Sj} \tag{8}$$

where i and j sum over each row and column. The maximum energy of the Puzzle is achieved every single lattice site is filled with the same number and hence all entries in all columns and rows will be non-unique. This corresponds to a score a 9 for each row and column leading to a maximum contribution from the second term in equation 8 of 162. The first term of 162 is hence added so that if there are no conflicts and all row and column entries are unique the energy will then be zero and puzzle will be solved.

To evolve the model one of the 9 sub-boxes is chosen. Then two random lattice points in this box (which are not fixed by the initial puzzle input) are chosen and swapped. Just like in the IM MA if the energy of the swap is less than the current energy the swap is made. If it is not it is accepted with the probability outlined in equation 4. A flow chart of the MA for the IM and Sudoku can be found in the appendix.

2 Results

For the IM implementation of lattice size of approximately 128 was used as this felt like a good compromise between the performance of the algorithm and the sample size. **A live animated plot of the Ising model along with live updating observable is available on GitHub (Link to the GitHub can be found in the appendix).**

The energy of the lattice was first initialized as some energy E_0. This was done by first counting the energy

[†]This Cooling schedule was actually directly taken from Stefan Hutzler's assignment on the ODE solution of Newton's law of cooling [2]

[‡]All Sudoku Puzzles published in all major sources actually contain only one unique solution.

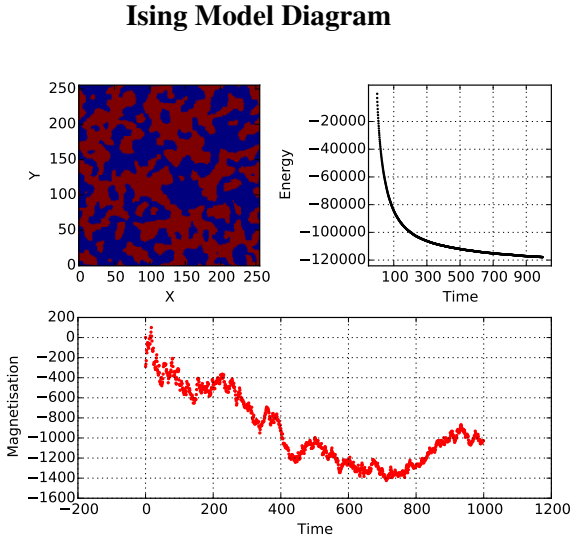

Figure 3: *Subplot figure showing the effect of the Metropolis Algorithm after 50,000 iterations on a 256x256 square Ising Model Lattice. Varying the temperature gave noticeably different results. The temperature in this plot is set to 0.1. Coupling constant J is set to 1.*

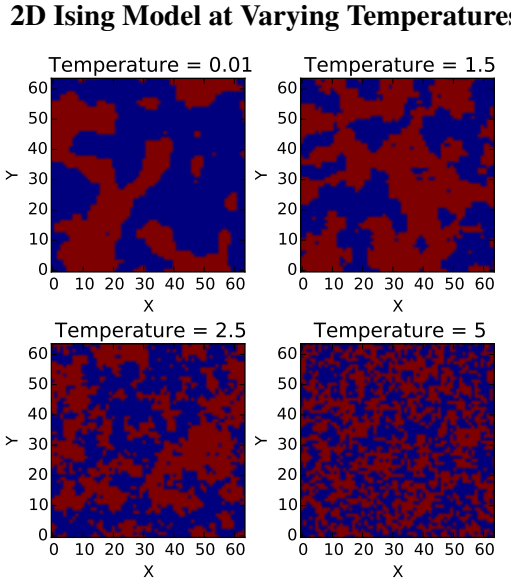

Figure 4: *Subplot figure showing the effect of the Metropolis Algorithm after 50,000 iterations on a 64x64 square Ising Model Lattice. Varying the temperature gave noticeably different results.*

for each given lattice site. However for subsequent iterations over the metropolis algorithm only the energy change was calculated. This led to a very significant boost in the performance of the animation and overall run time. Also to improve the overall run time of the program the observables were collected in intervals. That is the metropolis algorithm would run and select s different lattice points and then the observables would be calculated. **This value s is known as the sweep size and significantly increases the performance**. This is because the most resource intensive process of the program is the calculating of the macroscopic properties.

Figure 3 shows the evolution of a 256x256 lattice with the total magnetization and energy for 50,000 iterations with the plot being updated every 50 iterations (this is plotted as time on the x axis). This figure contains a lot of important information so it is worth analyzing it thoroughly.

As expected domains can be seen forming of similar spin. This occurs when by chance a small region of like spins will be created. Any spin within this region is highly unlikely to change from the domain direction as all neighbouring sites will be of that spin. Also spins neighbouring lattice points will be more likely to become changed to that spin if they are surrounded on 3 or more sides. Also small enclaved regions tend to eventually disappear and become engulfed by the surrounding region. This is because on average every boundary particle is surrounded by just above 2 particles of opposite spin. Leading to a slightly below zero Hamiltonian for the spin flips leads to the slow but inevitable decay of enclave area.

The graph on the top right of figure 3 shows the energy of the lattice as a function of time. As expected the initial energy begins at a high value of initial energy. This then can be observed to quickly drop off and then converge to a value dependent on the lattice size. The net magnetization can be seen at the bottom of the figure. This is obtained by summing all of the spins in the lattice. For a fully magnetized IM with all spins aligned leading to a net magnetization **M** of the region of a value n^2.

We can see the effect of applying different temperatures on the system shown in figure 4. At higher temperature values such as 5 there is no discernible pattern. At very low temperatures such as T = 0.01

Ising Model Phase Change

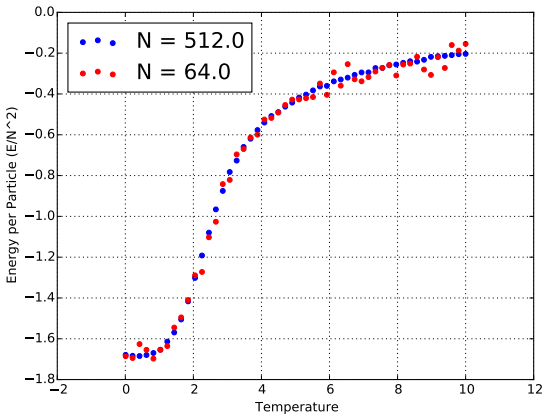

Figure 5: *Graph illustrating the Energy Temperature relationship after 50,000 iterations on a 2D square Ising Model Lattice.*

Sudoku Phase Change

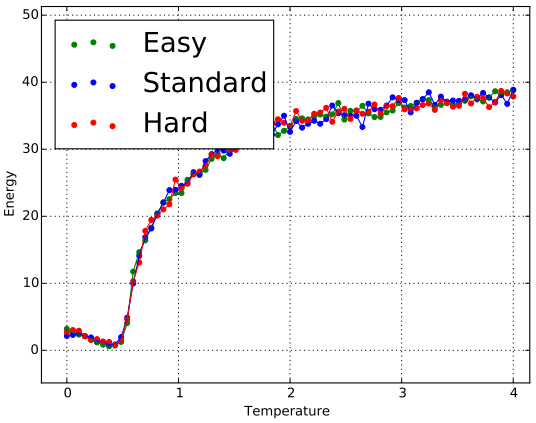

Figure 6: *Energy Plot of average of 40 Puzzles for 50,000 iterations each. 75 total points per difficulty with Easy,Medium and Hard providing the 3 reference points. No annealing was implemented. As the grid size is only ever 81 the energy was not adjusted per lattice site as in 5*

distinct domains can be seen. At T = 1.5 domains can be seen but there are other smaller domains or single points present which are not seen in the T = 0.01 plot. This is caused by the large difference in probabilities of accepting less favourable energy configurations in open areas. With the temperature at 0.01 the probability of accepting the new configuration is essentially zero. While at T = 1.5 the probability of accepting a spin flip in the centre of a different region is 0.4% which is far more likely remembering that this can "snowball" generating larger (yet still small and often short lived) enclosed regions (enclaves).

Figure 5 shows the dependence of the energy on the temperature of the system. Here two distinct regions can be seen. A higher energy region and a lower energy region. The lower energy region seems to occur at lower temperature while the higher energy at higher temperatures. This is directly analogous to a phase change of a material. More specifically this is known as a **Second Order Phase Transition or a Continuous Phase Transition**. These described regions are separated by a distinct barrier present at close to a temperature of (2.5 ± 0.3). This temperature at which the phase change occurs and is known as the critical temperature T_C. This effect of spontaneous magnetization can be seen in real magnets also when they become super cooled. Conversely it was also shown that if a magnet which is already magnetized has the temperature of the system raised it will then loose its magnetization caused by the large increase in the kinetic energy of the particles and thus increases the entropy of the system.

If our attention is now instead switched to Sudoku an almost identical phase change can be seen. Figure 6 shows us the same characteristic energy profile as figure 5. This confirms our assumption that the Sudoku puzzle will imitates some of the properties of the 9-State Potts model[§]. The region of low temperature has an apparent minimum at a value of T = 0.5. The puzzle then seems to rise quickly and then converge to an energy of approximately 40. This low energy low temperature phase change occurs at a value of (0.75 ± 0.15). Something to notice is that this temperature is lower than that for the IM and it may prove insightful to ask why.

When generating random puzzles the energy of a random puzzle seems to initialize and fluctuate at approximately an energy of 40 if a high enough temperature is implemented to simulate randomness. The

[§]These Puzzles were found on an online website for Sudoku puzzles. The quoted difficulties were obtained from this sites ranking system. Link to the puzzles given in the appendix.

Sudoku Puzzle has a unique ground state that being the solved puzzle solution. Once the Sudoku puzzle reaches its ground state it is unlikely to move. This could explain the small dip seen at T = 0.5. If the temperature is too low, say at 0 the puzzle is likely to get stuck at one of the many local minima.

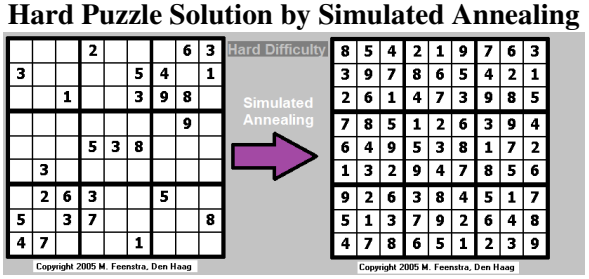

Figure 7: *Subplot figure showing the effect of the Metropolis Algorithm after 50,000 iterations on a 64x64 square Ising Model Lattice. Varying the temperature gave noticeably different results.*

This Sudoku puzzle shows a the properties of a paramagnetic phase transition. This phase transition of this puzzle is unique and provides insight for a number of reasons. Firstly, there is a unique ground state. If some lattice sites of a standard Ising Model become fixed in a simulation it is no longer the case where the ground state energy is simply all comprised of colinear spins. This make solving the system analytically challenging and it is possible that there are multiple different configurations which all have the lowest energy possible by the system. Not to mention that some of these may just be rotations and reflections of the same lattice. Also changing the coupling constants may provide us with more interesting variations on the usual ferromagnetic ground state of all co-linear magnetic moments. **Sudoku gives us a single unique global minimum so the effectiveness of things like cooling schedules in simulated annealing can be reliably tested** while being sure the system has indeed reached a global minimum.

A logarithmic cooling schedule was implemented as in 7 with an R value of 0.999. This closely approximates a cooling body to its surrounding temperature¶. T_0 was set to 0.5 and the puzzle was run of 50,000 iterations. The temperature was reassigned every 1000

¶Which admittedly is unrealistic as the surrounding temperature in this case is absolute zero.

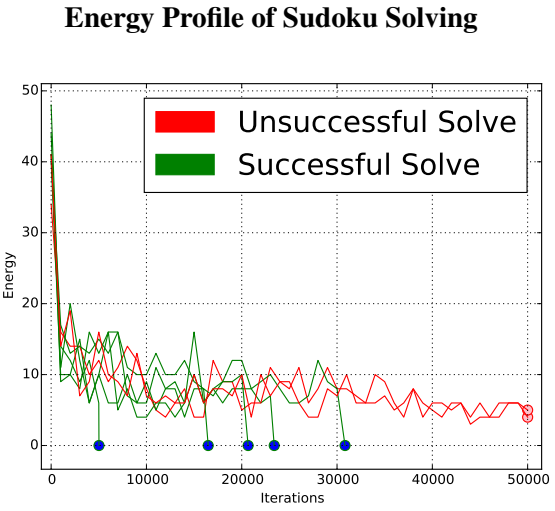

Figure 8: *Graph showing 7 solving attempts of the standard difficulty puzzle. Red lines/circles indicate an unsolved puzzle at 50,000 iterations while green indicate a successful attempt.*

iterations;

```
T *= 0.999
```

from figure 8, attempts at solving the standard puzzle were made. As the attempts were made the energy progression was recorded and plotted for 7 attempts. Things to notice include the sharp decrease at the start of the puzzle. This occurs because initially the puzzle will likely have many non-unique entries across many of the rows/columns. As lower energy states will be highly favoured at this stage the energy will rapidly drop. A more gradual decline is then seen as the energy of particles fluctuates. The energy can be seen to rise as well as fall given that it is possible for unfavourable energy configurations to be chosen. Also the graph appears to become less volatile as the number of iterations is increased. This is direct evidence of the annealing in action. Less energetically unfavourable changes occur as the puzzle attempted to descend to the global maximum with its lowest energy state.

Upon inspection of puzzles which were eventually left unsolved a recurring theme emerged. Certain boxes of the puzzle in the corners (say box 1 and 9) would be slightly incorrect, maybe by a cycle of 3

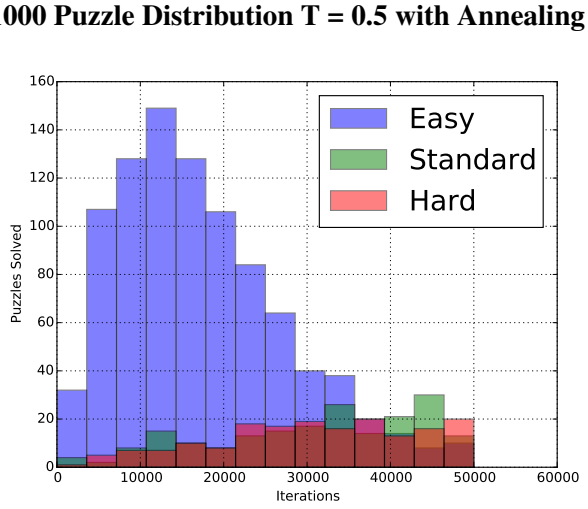

Figure 9: *Histogram plot showing the distribution of solved puzzles. Every Puzzle was run for 50,000 iterations unless solved prior to that. Unsolved puzzles were excluded. The initial temperature was set to 0.5 and cooled slowly.*

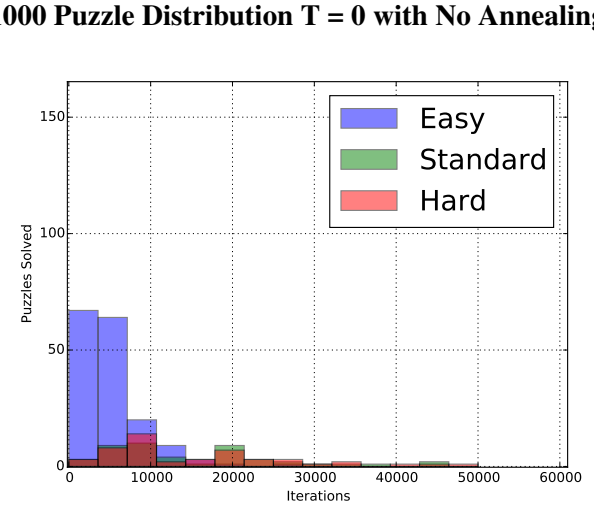

Figure 10: *Histogram plot showing the distribution of solved puzzles. Every Puzzle was run for 50,000 iterations unless solved prior to that. Unsolved puzzles were excluded. Temperature was set to 0 and the system was not cooled.*

digits each. However the fact that they were in boxes in opposite corners meant that they could not interact with each other and could only interact via other boxes. This formed a sort of pseudo-boundary of these boxes each in low energy configurations/regions themselves but these boundary boxes would be constantly changed.

Puzzles with fixed values that were inconsistent with the puzzles solution were also entered, this is directly analogous to a frustrated lattice. In such situations the puzzle would often attain energies as low as 2-4, without completing the puzzle. In these situations the system can no longer reach its optimum ground state and now instead must instead find another suitable configuration. Frustration appears in magnetic systems such as in a ferromagnetic triangular lattice. In which every spin tries to be anti-parallel but instead

Figures 9 and 10 show histogram plots of the number of iterations taken to solve the 3 puzzles of varying difficulties. If a puzzle was not solve it was not added to histogram. The value of 0.5 was chosen as this value corresponded to the lowest average energy of the system from plot 6. The solved puzzles in the histogram form a clear distribution.

The first obvious things by examining figure 9 is the heights/intensities of the graphs. The easy puzzle is by far the most numerously solved with a 924 puzzles of 1000 being solved in 50,000 iterations. The easy puzzles solved form a right-skewed distribution with the majority of lattices reaching their ground state before 20,000 iterations. The standard and hard puzzles were far less numerous with portions of this distribution cut off by the 50,000 iteration limit. As expected the less puzzle clues provided, the more time required to solve the puzzle on average. Also notice that almost no hard puzzles are solved within the first 10,000 iterations while easy puzzles have a moderately high success rate in this region.

The following graph (Figure 10) really highlights the critical importance of the Boltzmann distribution selection rule (temperature dependence). **The chances of reaching the ground state of a system fall dramatically without the annealing process.** This slow cooling of the system allows the lattice to enter into these low energy configurations. Named after its real life counterpart annealing is also seen in real life materials. In igneous rocks that cool from magma/lava, rocks that cool quickly such as basalt are much lighter and more brittle. This is because just as in the Ising

model and Sudoku puzzle the particles did not have time to order themselves efficiently into a super low energy configuration and so breaking the bonds in the lattice requires less energy. Whereas rocks such as granite cool over thousands of years. Leaving behind a very hard and structurally stable rock.

2.1 Conclusion

Sudoku has proved to be an insightful tool to help us explore more unique aspects of magnetic behaviour. Second order phase transitions were observed in both the Ising model and Sudoku puzzles. However the critical temperature for Sudoku was only calculated to be (0.75 ± 0.15) whereas the Ising model had a temperature of (2.5 ± 0.3) with a theoretical critical temperature of (≈ 2.27)[8]

3 Appendix

3.1 GitHub Link

https://github.com/butlerc9/IM-Simulation

3.2 Sudoku Puzzle Links

https://www.sudoku.ws/easy-1.htm
https://www.sudoku.ws/standard-1.htm
https://www.sudoku.ws/hard.htm

4 References

[1] Tom Archer Ising Model Assignment 2017/18
[2] Stefan Hutzler Numerical Methods Notes 2017/18
[3] Tom Archer Computational Notes 2017
[4] Gould & Tobochnik, An Introduction of Computer Simulation Methods (1996)
[5] Hastings, Monte Carlo Sampling Methods Using Markov Chains and Their Applications (1970)
[6] Sudoku Annealing Simulation URL: github.com/erichowens/SudokuSolver accessed 22-Dec-2017
[7] Williams A., Paramagnetic and Glass Transitions in Sudoku
[8] Baker George A., Ising-Model Critical Indices below the Critical Temperature, 1967