

Log-Linear Learning Over Large Iterations in Sudoku Puzzles

Vince Coghlan

April 29, 2014

Sudoku is a puzzle game played on a 9×9 board. The goal is to have a unique occurrence of each didget in each row, column, and 3×3 square. This can be modeled as a game with the following parameters:

- Players $N = \{1, 2, \dots, n\}$ Where n is the number of empty spaces
- Actions $a_i \in \mathcal{A}$ where an action is a number between 1 and 9 and the actions set $\mathcal{A} = a_1 \times a_2 \times \dots \times a_n$
- Utility Functions $u_i : \mathcal{A} \rightarrow \mathbb{R}$ Where the utility will be the number of repetitions of a player's current action in a row, column, or square.

$$u_i(a) = u_{\text{row}_i}(a) + u_{\text{column}_i}(a) + u_{\text{square}_i}(a) \quad (1)$$

where

$$u_{\text{row}_i} = \sum_{j \in i_{\text{row}}} (a_j = a_i)$$

- Welfare Function $W(a) : \mathcal{A} \rightarrow \mathbb{R}$ where the global welfare of an action represents the global desirability of that action. Let a be the joint action (a_i, a_{-i}) where $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$

We are given the ability to design the welfare function however we would like. Suppose that our objective is captured by a function $\phi : \mathcal{A} \rightarrow \mathbb{R}$ where

$$u_i(a_i, a_{-i}) - u_i(a_i^*, a_{-i}) = \phi(a_i, a_{-i}) - \phi(a_i^*, a_{-i}) \quad (2)$$

A ϕ satisfying (2) will be known as a potential function [6]. It can be shown that the game of Sudoku has a potential function when we define the utility as in (1). We begin by assuming that the potential function will take the form

$$\phi(a) = \phi_{\text{row}}(a) + \phi_{\text{column}}(a) + \phi_{\text{square}}(a)$$

We can then show that

$$\phi_{\text{row}}(a) - \phi_{\text{row}}(a^*) = u_{\text{row}_i}(a) - u_{\text{row}_i}(a^*)$$

And so on. When a player changes his number, he increases the amount of repetitions of that number by 1. By doing this, however, he has unintentionally increased the amount of repetitions of any other player performing the same action, and decreased the amount of repetitions of his old action. One would expect that the potential function is the sum of utilities of each player

$$\begin{aligned} \phi_{\text{row}}(a) &= \sum_{i \in N} u_{\text{row}_i}(a) \\ \phi_{\text{row}}(a) - \phi_{\text{row}}(a^*) &= \sum_{i \in N} u_{\text{row}_i}(a) - \sum_{i \in N} u_{\text{row}_i}(a^*) \\ &= \sum_{i \in N} \sum_{j \in i_{\text{row}}} (a_j = a_i) - \sum_{i \in N} \sum_{j \in i_{\text{row}}} (a_j = a_i^*) \\ &= \sum_{j \in i_{\text{row}}} (a_j = a_i) - \sum_{j \in i_{\text{row}}} (a_j = a_i) + \sum_{i \in N \setminus \{i\}} \sum_{j \in i_{\text{row}}} (a_j = a_i) - \sum_{i \in N \setminus \{i\}} \sum_{j \in i_{\text{row}}} (a_j = a_i^*) \end{aligned}$$

Now we will use the fact that an increase in n in the repetitions of a_i corresponds to an increase of 1 of n other players playing a_i . This means that:

$$\begin{aligned} &= \sum_{j \in i_{\text{row}}} (a_j = a_i) - \sum_{j \in i_{\text{row}}} (a_j = a_i) + \sum_{j \in i_{\text{row}}} (a_j = a_i) - \sum_{j \in i_{\text{row}}} (a_j = a_i) \\ &= 2 \left(\sum_{j \in i_{\text{row}}} (a_j = a_i) - \sum_{j \in i_{\text{row}}} (a_j = a_i) \right) \end{aligned}$$

This tells us that we have twice the potential function, meaning our potential function for the row is

$$\phi_{\text{row}}(a) = \frac{1}{2} \sum_{i \in N} u_{\text{row}_i}(a)$$

This can be easily extended to rows and columns to find:

$$\phi(a) = \frac{1}{2} \sum_{i \in N} u_i(a) \tag{3}$$

Since sudoku has a potential function, the minimization of the potential is a pure equilibrium [3]. Since we cannot have a negative potential by our definition of the utility, and a potential of 0 means that there are no repetitions in any row, column, or square, a pure Nash Equilibrium exists, and must be the solution of the puzzle.

We have half of our problem laid out, the next part is to define a learning algorithm that will converge to this equilibrium. Current strategies being studied currently include fictitious play [5], log-linear learning [1], and the cournot best reply process [2] among others. Many of these learning algorithms converge to a Nash equilibrium [4]. It has been shown [7] that

Log-linear learning provides us the fact that, in any exact potential game, log-linear learning will induce a Markov chain with a unique stationary distribution $\pi \in \Delta(\mathcal{A})$ of the form

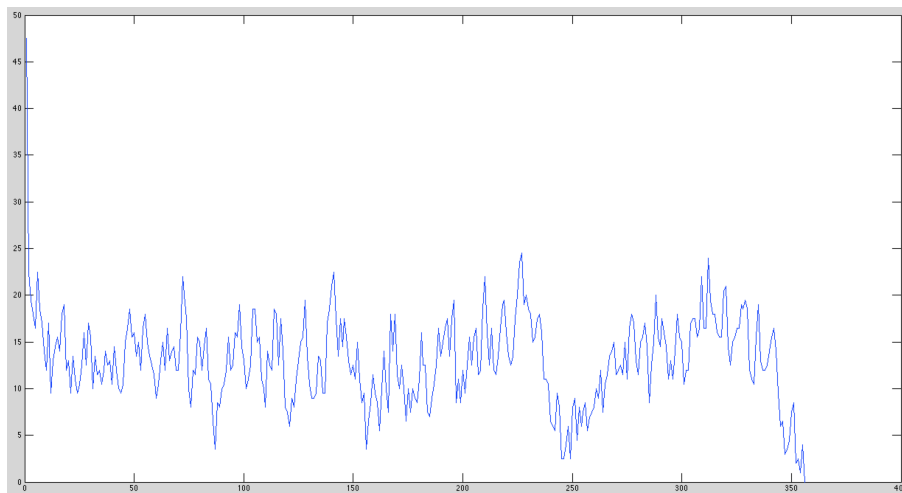
$$\pi_a = \frac{e^{\frac{1}{T}\phi(a)}}{\sum_{\tilde{a} \in \mathcal{A}} e^{\frac{1}{T}\phi(\tilde{a})}} \quad (4)$$

Since ϕ is zero at the solutions, we know that the distributions of the system as $T \rightarrow 0$ will be distributed on solutions to the sudoku puzzle. We can see this behavior for some simple sudoku puzzles, for example, this puzzle was solved in 28836 iterations, with a temperature of $T = 0.5$.

7		3		1	4	6		
			7		9		8	
		8	3					1
5		9					4	
				2		9	1	6
2			9		8			7
	8	7				4	5	
9	1			3		2		
3								9

7	9	3	8	1	4	6	2	5
1	2	6	7	5	9	3	8	4
4	5	8	3	6	2	7	9	1
5	3	9	1	7	6	8	4	2
8	7	4	5	2	3	9	1	6
2	6	1	9	4	8	5	3	7
6	8	7	2	9	1	4	5	3
9	1	5	4	3	7	2	6	8
3	4	2	6	8	5	1	7	9

The potential can be seen



You can see that once the potential is 0, we are done, and the puzzle is solved.

References

- [1] L. Blume, The statistical mechanics of strategic interaction. *Games and Economic Behavior*, pp. 387–424, 1993.
- [2] A. Cournot, *Recherches sur les principes mathématiques de la théorie des richesses*, Hachette, 1838.
- [3] A. Neyman, Correlated Equilibrium and Potential Games. *International Journal of Game Theory*, pp. 223–227, 1997.
- [4] J. Marden and J. Shamma, Revisiting Log-Linear Learning: Asynchrony, Completeness and Payoff-Based Implementation *Games and Economic Behavior*, Volume 75, Issue 4, pp. 788–808, 2012.
- [5] D. Monderer and L. Shapley, Fictitious Play Property for Games with Identical Interests. *Journal of Economic Theory*, Volume 68, pp. 258–265, 1996
- [6] D. Monderer and L. Shapley, Potential Games. *Games and Economic Behavior*, Article No. 0044 pp. 124–143, 1994.
- [7] H.P. Young. The Evolution of Conventions. *Econometrica*, pp. 57–84, 1993.