Chaotic Time Series Prediction for Rock-Paper-Scissors using Adaptive Social Behaviour Optimization (ASBO)

S.Rajarajeshwari
Dept of Computer Science
MSRIT,Bangalore,India

Raveena Kumar
Dept of Computer Science
MSRIT,Bangalore,India

Sandhya Sankaranarayanan
Dept of Computer Science
MSRIT,Bangalore,India

ABSTRACT
Time series prediction involves analyzing a set of data from past and current occurrences in order to predict the future set of data. In dynamic systems, chaotic behaviour is intrinsically observable and the resulting chaotic time series have non-linear characteristics. Nevertheless, such data can be optimized to make sense out of the chaos. Multiple algorithms exist to this end, which have various applications. In this paper, this phenomenon is illustrated by making use of the well-known game rock-paper-scissors(R-P-S) played between two agents, one real and one adaptive. It is possible to identify and predict patterns in the choices made by the real agent during the course of play by analyzing the sequence of chaotic data using the Adaptive Social Behaviour Optimization (ASBO) algorithm. This optimization method makes use of a self-adaptive mutation strategy which takes into consideration dynamic factors such as leadership, confidence and competition, which are all functions of time.

KEYWORDS
ASBO, Chaotic time series, optimization, predict patterns, Rock Paper Scissors

1. INTRODUCTION
The source of chaos in a game of R-P-S is the availability of three choices of play, namely rock, paper and scissors. Previous studies made on this game have suggested that the agents change the probability of playing each choice over 'n' rounds. The underlying concept for this calculation is coupled Ordinary Differential Equations, which forms the basis of reinforced learning of agents.

Real players, assuming the role of agents, introduce some scope for predictability into what would have otherwise been erratic behaviour. Each player will have an affinity towards a particular choice or sequence of choices of play. This affinity is associated with the subconscious mind of the player and results in the real agent exhibiting a pattern. Identifying such patterns emerging from chaos has multiple applications. For example, predictions can be made about the rises and declines in the market, in a country's annual economic output and other such phenomena.

At any given point in the game, one of the two agents involved will have an upper hand and hence can be termed as the temporary leader. Interactions with the other agent will result in building the knowledge base. Reinforced learning can be used by an agent to adapt itself based on the behaviour of the other agent. The purpose of this study is to observe such interactions and their consequences that cause the real agents to fall into their patterns.

The Adaptive Social Behaviour Optimization (ASBO) method derives from the characteristics of competition, leadership, influence and self-confidence. Although multiple factors contribute heavily to the adaptability and ultimate success of the adaptive agent, identifying and abstracting only a few suggestive characteristics facilitate in establishing the underlying social structure. This optimization algorithm has been incorporated as it includes the inherent characteristics of competition and influence.

2. OPTIMIZATION ALGORITHMS
Multiple algorithms such as the Particle Swarm Optimization algorithm, the Ant Colony Optimization algorithm and the Adaptive Social Behaviour Optimization algorithm exist to find solutions for problems with chaotic data.

2.1 Particle Swarm Optimization (PSO)
This optimization method is derived from a phenomenon of nature: a swarm. This is exhibited by various creatures such as fish, and a flock of birds. Such groups of animals do not have any specific leader in their group. When such a society or group has a goal (solution) to achieve, for example, to find food or to reach a destination for hibernation, they will do so by following a member who is visibly in the best situation and hence is closest to the potential solution. Each 'particle' in the algorithm is synonymous to a member of the group, where each particle represents a solution. Consider the example of a flock of birds flying to a destination. Each member follows a member that is in a better situation than itself. The group is only as good as its worst member. In this manner, through constant communication with each other, the best condition is reached. This would happen repeatedly until the destination is reached i.e. - the solution is obtained. In order to improve the speed of convergence and the quality of the solution obtained, a number of variations of PSO have been developed in recent times. While basic PSO is appropriate for a simple static problem, modified PSO algorithms can be applied to the more complex ones.

2.2 Ant Colony Optimization (ACO)
Ants leave a trail of pheromone, a hormone, as they move. Other ants can detect the pheromone and hence know the path taken by any predecessors. When there are multiple trails, the ants chose a trail with larger amounts of pheromone. Consider the ant system illustrated in Fig 1. Initially, the ants move in a straight line, the shortest path from their nest to a source of food. When their path is obstructed by a rod, the ants cannot continue in their earlier path anymore. Hence, randomly, some go left, while the rest go right. Assuming all the ants move at the same speed and deposit the same amount of pheromone each, the ants that turned left will reach sooner. Therefore, the pheromone trails on that path will be stronger, and eventually, all the ants will converge onto this path. An artificial Ant

Colony System (ACS) is an agent-based system, which emulates the natural behaviour of ants in colonies to solve combinatorial optimization problems. ACO reduces any problem into a graph, where the cost of the edges is synonymous to the amount of pheromone in the trails. The solution can be obtained by conducting a search for the path with the minimum cost.

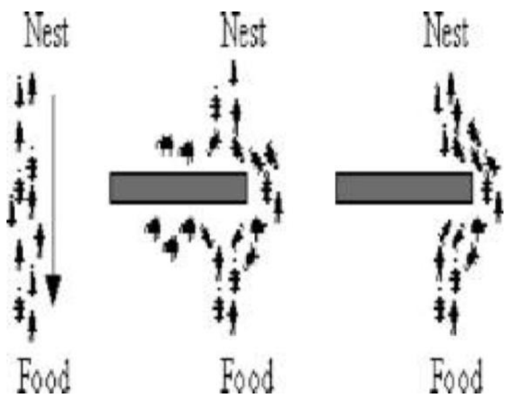

Fig 1: Illustrating the behaviour of real ant movements [3]

2.3 Drawbacks

Although both these are candidate algorithms for R-P-S pattern prediction, both the algorithms have a few drawbacks:

Drawbacks of PSO:
1) Results in sub-optimal solutions at times, which are difficult to recover from
2) Scattering and optimization problems cannot be solved

Although in ACO, probability distribution changes with iterations, which complements the R-P-S situation, it is not as good a match as we would like, as sequences of random decisions are made, which would introduce abnormalities in our pattern prediction and would result in wildly incorrect fitness values.
Therefore, we propose to use the Adaptive Social Behaviour algorithm.

3. ADAPTIVE SOCIAL BEHAVIOUR OPTIMIZATION (ASBO)

3.1 ASBO Based Learning

The ASBO algorithm proposes the concept of fitness. An objective function derives the fitness value of each involved agent. At a given point of time, the current leader is he who has the maximum fitness value. This position is temporary as the existing status is prone to change. This change in existing status because of influences is brought about by each and every member of the population using Equation 1 (1) and the location of the next status is given by Equation 2 (2).

$$\Delta X(i+1) = Cg * R1 * (Gbi - Xi) + (Cs * R2 * (Sbi - Xi) + Cn * R3 * (Nci - Xi) \quad \text{...(1)}$$

$$X(i+1) = Xi + \Delta X(i+1) \quad \text{...(2)}$$

Where,
$\Delta X(i+1)$ represents the new change in the i^{th} dimension of an individual element
Cg, Cn, Cs >=0 are adaptive progress constants
Ri for all i=1,2,3 are uniformly distributed random numbers in the range [0,1]
Gb : Global best individual at present population
(member having maximum fitness value)
Sb: Self-best for an individual
Nc: Center position of a group formed by an individual and its neighbours

For a D-dimensional problem, Gb, Sb and Nc represent vectors of D-dimension

$$Gb = [Gb1, Gb2, Gb3, \ldots GbD];$$
$$Sb = [Sb1, Sb2, Sb3, \ldots SbD];$$
$$Nc = [Nc1, Nc2, Nc3, \ldots NcD];$$

Our population consists of only two agents. Each agent represents a solution which is initially random. Also, each agent is a neighbour of the other. The initial value of self-best for both the agents is same as their initial solution. A new set of progress constants are obtained by applying self-adaptive mutation strategies.

3.2 ASBO Methodology

The ASBO method is applied to both the members independently for 'n' iterations. The fitness values and progress constants are recorded for each member. At the end of this process, the adaptive agent would have built his knowledge base due to the diversity of outcomes. Successive fitness values that are calculated at the end of each iteration contribute to the calculation of the next set of values of fitness and progress constants.

At the end of each round, one agent will seem to be in a more favourable situation than the other. This result is valid only in a limited domain and not for an extended time period as the environment is dynamic. In this manner, the solution region can also be localized effectively. The agent with the higher fitness value is the one with the more reliable knowledge base and hence can be termed as the leader. We can also assume that this agent represents the optimal solution in the local domain and in the current time period.

4. SOCIAL STRUCTURE

The scope for application of the ASBO algorithm is extensive and can be expanded to suit a community of 'k' members. However, R-P-S requires only two agents. Hence the algorithm can be limited to complement the requirements. As mentioned before, at any point in time, one agent is in a better position than the other. As the game progresses, the situation changes according to the moves made by each agent.

We can judge the superiority of one agent based on how many rounds were won by that agent, how well he could anticipate the other agent's next move, how well he retains his position and his ability to learn from the rounds he lost. All these decisions can be inferred to some extent from the knowledge base built and through reinforced learning. The ASBO algorithm in this situation is applied to both the agents. This agent is made to play with the real agent. Leadership, computed using ASBO, is temporary; it could fluctuate between the two agents depending on the progress of the game. Leadership in the game is synonymous to better learning and predicting capabilities.

5. COMPUTATION PROCESS

Reinforced learning is imperative in a dynamic environment if an agent wants to optimize its behaviour. Although ideally, an agent only improves its decision making skills and makes accurate inferences that lead to the right actions, the fact remains that sub-optimal learning could occur. The best we can do is to ensure that the agent recovers adequately from this. A simple decision making process can be used for this. With the help of the ASBO algorithm, it would be possible to determine whether the adaptive agent is able to predict the real agent's moves with maximum precision.

Initially, the probability of each move is 1/3 for both agents, as shown in Table 1.

Table 1. Probability of each move for both the agents

AGENT / MOVE	ROCK	PAPER	SCISSORS
A	1/3	1/3	1/3
B	1/3	1/3	1/3

Let the choice made by the adaptive agent be C_1. Based on the outcome of the first round, the adaptive agent increases the probability of playing C_1 if it wins the round and decreases the probability if it loses. Say, if agent A (Adaptive agent) plays R while agent B (Real agent) plays S, then A wins. Thus A increases its probability of playing R. The table could be modified appropriately by increasing the possible moves to a number n and distributing R, P, S by appropriate factors. Two possible stages are illustrated below (Table 2 and Table 3), where n is 7. The probabilities of choices for the real agent B are unpredictable as it is influenced by personal whimsies. Hence, it is futile to compute or tabulate the probabilities of moves made by B.

Table 2. Stage 1: After Agent A plays R and Agent B plays S

AGENT / MOVE	ROCK	PAPER	SCISSORS
A	3/7	2/7	2/7

Table 3. Stage 2: After Agent A plays P and agent B plays R

AGENT / MOVE	ROCK	PAPER	SCISSORS
A	4/11	5/11	2/11

Thus, the adaptive agent updates its probability of playing a particular strategy in the future. One such method proposed by Salvetti, Patelli, Nicolo[4] is as follows:

For each strategy s_i , the probability of playing the strategy $\boldsymbol{P_{t+1}(s_i)}$ at a time t+1 is calculated as follows:

$$\boldsymbol{P(s_i)_{t+1} = P(s_i)_t + w_i(\propto(1 - P(s_i)_t))}$$

Where w_i = 1 if outcome is positive and -1 if outcome is negative.

$\propto$ is the learning rate, which is used to determine the adaptation velocity of the agent.The complementary strategy can be updated using the following formula

$$\boldsymbol{P(s_i^{'})_{t+1} = P(s_i^{'})_t + (1 - w_i) \propto (1 - P(s_i{'})_t)}$$

At the end of each update, each agent maintains

$$\boldsymbol{\Sigma P(s_i) = 1}$$

6. SUMMARY OF THE ASBO PROCESS

The process that is being suggested and the methodology to obtain an optimal knowledge base for Rock Paper Scissors is summarized below in Fig. 2.

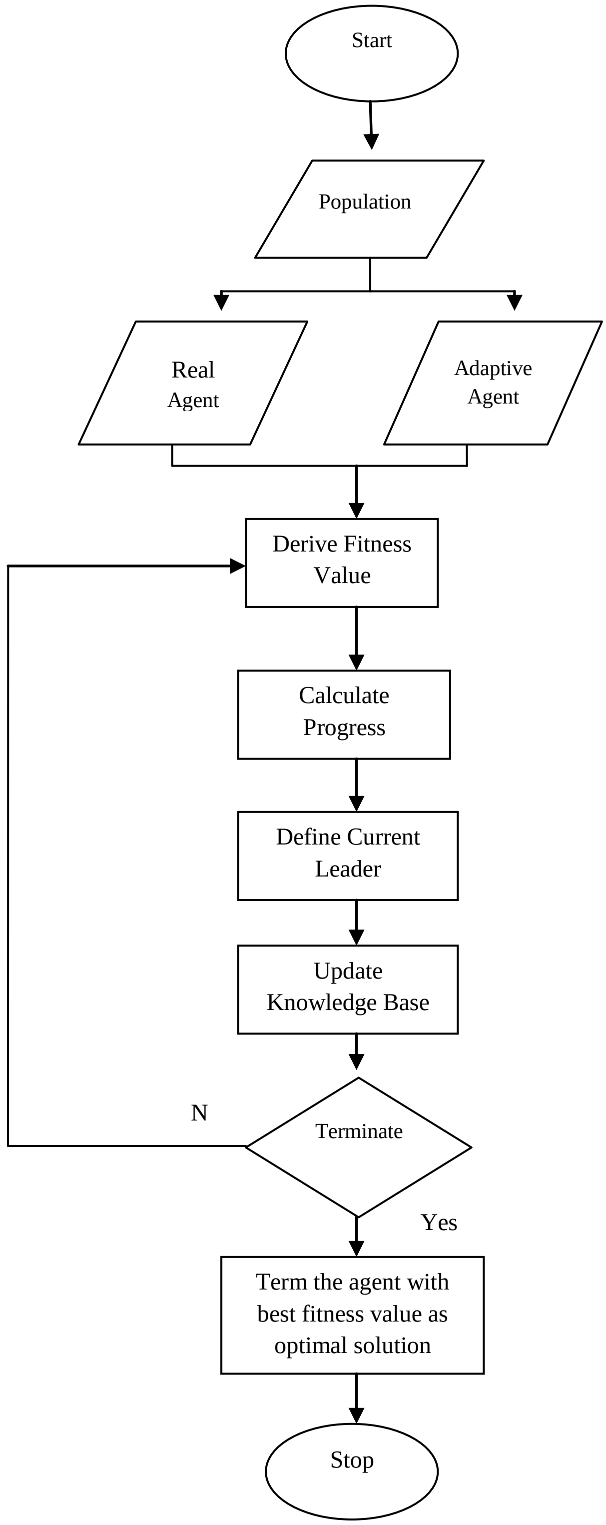

Fig 2: Summary of Method Proposed

7. CONCLUSION

The application of the ASBO to Rock Paper Scissors allows one to extract patterns exhibited, despite the chaotic nature of the game. A move is picked by the adaptive agent based on its computations of probability, taking characteristics such as past experience, competition and influence into account as all these factors are pivotal in a dynamic environment. Identifying patterns, assessing the fitness values and building a reliable knowledge base produces an optimal solution, thereby allowing the adaptive agent to become the leader of the game in the future.

8. ACKNOWLEDGEMENTS

Our thanks to Mrs S.Rajarajeshwari for her guidance and support, without which this would not have been possible.

9. REFERENCES

[1] Singh Kr. Manoj.: “A New Optimization Method Based On Adaptive Social Behaviour: ASBO”, ASIC 174, PP.823-831, Springer 2012

[2] Franco Salvetti, Paolo Patelli, Simone Nicolo.: “Chaotic Time Series Prediction for the game, Rock-Paper-Scissors”, ASOC-228

[3] Dian Palupi Rini, Siti Mariyam Shamsuddin, Siti Sophiati Yuhaniz: “Particle Swarm Optimization: Technique, System and Challenges”, International Journal of Computer Applications (0975 – 8887), Volume 14– No.1

[4] V.Selvi, Dr.R.Umarani: “Comparative Analysis of Ant Colony and Particle Swarm Optimization Techniques”, International Journal of Computer Applications (0975 – 8887), Volume 5– No.4

[5] J.V. Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, 1944.

[6] S. Nicolo. Chaos and Game Theory, a study on Rock-Scissors-Paper. Technical Report Chaotic Dynamics-CSCI 6446, University of Colorado at Boulder, 2004.

[7] World RPS. Advanced RPS. Technical report,World Rock Paper Scissors Society, 2004.

[8] Y. Sato, E. Akiyama, J.D. Farmer, Chaos in learning a simple two person game, Proc. Natl. Acad. Sci. USA 99 (2002) 4748–4751.

[9] A.S. Weigend, N.A. Gershenfeld, Time Series Prediction, Addison Wesley, 1992.

[10] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, The MIT Press, 1998