Probability Collectives for Decentralized, Distributed Optimization: A Collective Intelligence Approach

Anand J. Kulkarni
School of Mechanical and Aerospace Engineering
Nanyang Technological University
Singapore
E-mail: kulk0003@ntu.edu.sg

Kang Tai
School of Mechanical and Aerospace Engineering
Nanyang Technological University
Singapore
E-mail: mktai@ntu.edu.sg

***Abstract*— The growing number of components and communication in complex systems requires them to be treated as a collective of subsystems/agents having distributed and decentralized control. The major challenge in such approach is the coordination among agents optimizing their local goals and contributing towards optimization of the global objective. This paper implements the theory of Collective Intelligence (COIN) using Probability Collectives (PC) approach to achieve the global goal. This approach works on probability distribution, directly incorporating uncertainty and has deep connections to Game theory, Statistical Physics and Optimization. In this approach, the agents select actions over a particular range and receive some rewards on the basis of the overall system objective achieved because of those actions. The approach is illustrated using the problem of segmented beam minimizing total volume. Each segment is considered as an agent competing with one another to achieve the total minimum volume. The implementation produced encouraging results.**

***Keywords*— collective intelligence, nash equilibrium, probability collectives, maxent, probability**

I. INTRODUCTION

Complex systems may have many components that not only interact but also compete with one another to deliver the best they can to reach the desired system objective. Traditionally, the complex systems were seen as centralized systems, but as the complexity grew, it became necessary to handle the systems using distributed and decentralized optimization. In a decentralized approach, the system can be divided into smaller subsystems and optimized individually to get the system level optimum.

In the approach presented in this paper, the system was divided into subsystems and each subsystem was considered as an agent. The framework of Collective Intelligence (COIN) is an approach to design a collective, which includes the computational agents working to enhance the system performance. Essentially, in COIN approach, the agents select actions over a particular range and receive some rewards on the basis of the system objective achieved because of those actions. The process iterates and reaches equilibrium when no further increase in reward is possible for the individual agent by changing the actions further. This equilibrium concept is known as Nash Equilibrium [1]. This concept can be successfully formalized and implemented through the concept of Probability Collectives (PC). This approach works on probability distribution, directly incorporating uncertainty.

The paper is organized as follows. Section II reviews COIN using PC theory and the relevant concepts. It is followed by the formulation of PC theory in Section III. The illustration of the theory using the practical problem is in Section IV. It is followed by results, discussion in Section V. Some concluding remarks and future work are at the end in Section VI.

II. COLLECTIVE INTELLIGENCE (COIN) USING PROBABILITY COLLECTIVES

A. Probability Collectives (PC)

A collective is a group of learning agents, which are self interested and work in some definite direction to optimize the local rewards or payoffs, which also optimizes the global or system objective. The system objective is also referred to as world utility and is a measure of performance of the whole system. Probability Collectives (PC) theory is a broad framework for modeling and controlling distributed systems, and it has deep connections to Game Theory, Statistical Physics, and Optimization [2] [3].

The method of PC theory is an efficient way of sampling the joint probability space, converting the problem into the convex space of probability distribution. PC considers the variables in the systems as individual agents/players of a game being played iteratively [4]. Unlike stochastic approaches such as Genetic Algorithms (GA), Swarm Optimization and Simulated Annealing (SA), rather than deciding over the agent's moves/set of actions, PC allocates probability values to each agent's moves. In each iteration, every agent independently updates its own probability distribution to select a particular action out of its strategy set having the highest probability of optimizing its own utility (private utility) which also results in optimizing the world utility or system objective [3]. This is based on the prior knowledge of the actions/strategies of all other agents.

B. Why Apply PC?

PC approach has the following advantages over the other tools that can be used in optimizing collectives.

- PC is a distributed solution approach in which each agent independently updates its probability distribution at any time instance and can be applied to continuous, discrete or mixed variables, etc. [3], [4], [5]. The probability of the strategy set is always a vector of real numbers, regardless of the type of data under consideration. This exploits the techniques of the optimization for Euclidean vectors, such as gradient descent.
- It is robust in the sense that the cost function (Global Utility function) can be irregular or noisy i.e. can accommodate noisy and poorly modelled problems [6], [7].
- It provides sensitivity information about the problem in the sense that a variable with a peaky distribution (having highest probability value) is more important in the solution than a variable with a broad distribution i.e. peaky distribution provides the best choice of action that can optimize the global utility [2], [6].
- The minimum value of the global cost function can be found by considering the maxent Lagrangian equation for each agent (variable) [4].
- It can include Bayesian prior knowledge [5].
- It can efficiently handle the problems having large number of variables [6], [8].
- PC theory can handle constraints which are particularly useful for the aerospace systems.

These advantages of PC make it a preferred choice over other algorithms. The key concept of PC theory is that the minimum value of the world utility can be found by considering the maximum entropy (maxent) of agent/variable.

C. Maximum Entropy (Maxent) Principle

The maximum entropy principle is related to the field of statistical physics. This principle is especially useful when the system has many degrees of freedom, or has many variables which are very difficult to measure (sometimes impossible); also on the other hand, the known quantities are very few. To estimate the states of unknown variables, all their possible states are assigned uniform probabilities, [3], [9].

The discussion can be extended in terms of information theory [7], [3]. The information is measured in terms of bits. If one was told the information which he already knows, then he was not given any extra information that he can learn and hence learnt '0' (zero) bits. On the other hand if given some information which is unknown to him, then one can say that he learnt some (> zero) bits. Now, in terms of information theory, entropy gives average information in a probability distribution over a sample space. It is expressed in terms of *S* bits as

$$S = -\sum_i q_i . \log_2 \left(1/q_i \right) \tag{1}$$

Now, in the same terms of information theory, principle of Maximum Entropy (maxent) states that one should select the probability distribution which leaves the largest remaining uncertainty due to the unavailable information about the states or the values that have been occupied by each agent. The maximum entropy should be evaluated subject to the system constraints (the expected utilities/payoffs). The maxent Lagrangian is necessarily convex over the set of the product distribution over the agents' strategy set or the move space. Hence it is clear that, by iteratively searching the probability space using the methods like Gradient Descent or Newton Updating, the bottom of the convex bowl can be reached. While adding entropy makes the descent easier, it also biases the method away from extreme (hence true) solutions. That bias is gradually lowered by annealing the temperature (T).

The Principle of Maximum Entropy is used to discover the probability distribution which leads to the highest value for this uncertainty, thereby assuring that no information is inadvertently assumed [2], [4]. To deal with real world/physical systems and to avoid having a large number of bits, it is multiplied by Boltzmann's Constant (K_B) [4], [10]. In game theory, the game is generally viewed as an external observer. If the game is given information like the utility functions, information (strategy) sets and the expected utility, then the maxent principle says that the best estimate of the probability distribution is the one with maximum entropy subject to the expected values. The probabilities associated with maxent is the optimum joint strategy of the players of the game should follow [4], [9], [10].

III. COIN Formulation

The detailed implementation of COIN using PC theory is as follows, along with the algorithm flowchart represented in Figure 1.

Consider there are N agents. Each agent i is given strategy set $\mathbf{X}_i$ represented as

$$\mathbf{X}_i = \left\{ X_i^{[1]}, X_i^{[2]}, X_i^{[3]}, ..., X_i^{[m_i]} \right\} \tag{2}$$

where $i \in \{1,2,...,N\}$ and m_i is the number of strategies for each agent i.

1) Assign uniform probabilities to the strategies of Agent *i*. This is because, at the beginning, the *least* information is available (largest uncertainty) about which strategy is *favorable* and results in *minimization* of the objective function at the particular iteration under consideration of the particular values selected by agents other than i (i.e. guessed by agent i). Therefore, *'at the beginning of the game',* each agent's every strategy has probability '$1/m_i$' of being most favorable.

Therefore, probability of strategy r of Agent i is

$$q\left(X_i^{[r]}\right) = \frac{1}{m_i} \tag{3}$$

where $r = 1,2,...,m_i$.

Agent i selects its first strategy and samples randomly from other agents' strategies as well. This is a random guess by agent i. This forms a '*combined strategy set*' $\mathbf{Y}_i^{[r]}$ represented as

$$\mathbf{Y}_i^{[r]} = \left\{X_1^{[?]}, X_2^{[?]}, ..., X_i^{[r]}, ..., X_N^{[?]}\right\} \tag{4}$$

Superscript [?] indicates that it is a '*random guess*' and not known in advance. It is important to note that all the agents make their own '*combined strategy sets*' $\mathbf{Y}_i^{[r]}$.

From each of the '*combined strategy sets*' $\mathbf{Y}_i^{[r]}$ for agent i, compute the '*expected local utility*' for each agent i with strategy r as follows:

$$Exp\ Utility\ of\ \ Agent\ i^r = q_i^r \prod_{(i)} q(X_{(i)}^{[?]}) \cdot G(\mathbf{Y}_i^{[r]}) \tag{5}$$

where (i) represents every agent other than i and G is World Utility/System Objective.

Compute the Expected Global Utility on the basis of these *'combined strategy sets'* for every agent i. This is problem dependent. Update the probabilities of all the strategies of each agent i using Boltzmann's temperature T as follows,

$$q(X_i^r) \leftarrow q(X_i^r) - \alpha_{Step} \cdot q(X_i^r) \cdot k_{r\ update} \tag{6}$$

where

$$k_{r\ update} = \frac{(Contrib.\ of\ \ Agent\ i)}{T} + S_i(q) + \ln(q(X_i^{[r]})) \tag{7}$$

and

$$Contrib.\ of\ \ Agent\ i = Exp\ Utility\ of\ \ Agent\ i^r - Exp\ Global\ Utility \tag{8}$$

2) Using the above *second order update rule*, the strategy having maximum contribution towards the minimization of the objective is separated from the other strategies, i.e. its probability is increased iteratively.

For any agent i, if strategy r contributes more towards the minimization of the objective than other strategies, its probability increases with certain amount greater than the other strategies. This updating continues for predefined number of iterations to get clear probability distribution showing the highest probability for certain strategy value.

Compute the *'entropy'* of every agent's probability distribution as

$$S_i(q) = -\sum_{r=1}^{m_i} q(X_i^{[r]}) \cdot \ln q(X_i^{[r]}) \tag{9}$$

3) Entropy is the indication of the information availability. It increases when the information becomes clearer and reaches to maximum when information available also reaches to the maximum. When the entropy reaches the maximum, the probability distribution available clearly distinguishes every strategy's contribution towards the minimization of the Expected Global Utility. For detailed discussion on entropy refer to [7].

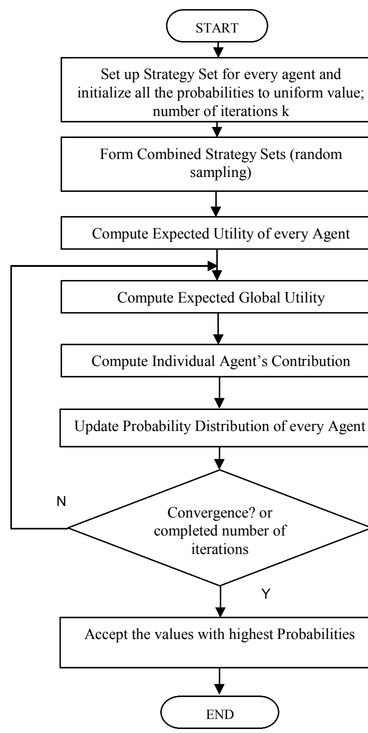

Figure 1. Algorithm of PC approach implemented

Repeat the procedure from Step 2 through Step 3 for predefined iterations (k) or until the difference between the probabilities of the successive iterations is less than the predefined value 'ε'. So, the convergence criterion is $\| \vec{q}^{\,k} - \vec{q}^{\,k+1} \| \le \varepsilon$

4. For each agent, identify the strategy, which contributes the maximum in minimizing the Global Utility and refer to it as the '*favorable strategy'*.

It is important to note that the above procedure is carried out for all the agents of the system in parallel.

Repeat the process for predefined number of times and store the *favorable strategy* values. Finally, retrieve the stored strategy values related to those maximum probability values and compute the related final objective values. The minimum of all those final objective values was accepted as the final objective. The next section implements this approach.

IV. Solution to the Segmented Beam Problem using COIN (Unconstrained)

To test the methodology presented in the previous section, the problem of segmented beam with rectangular cross-section was considered as shown in Fig.2 below. As this is the primary stage of the research in the field of COIN using PC theory, unconstrained approach was tested.

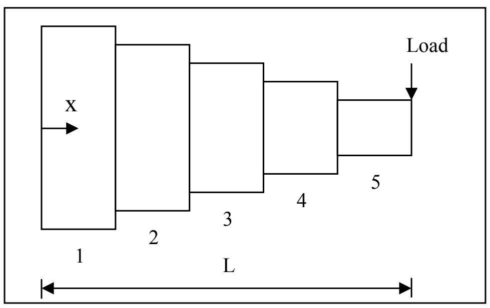

Figure 2. Idealized Geometry of tapered beam

As a general case, the individual segment of the entire beam contributes towards the whole beam volume. If the volume of any of the beam segment varies (increases/ decreases) it correspondingly varies (increases/ decreases) the volume of entire beam. In the context of COIN, the entire beam was seen as a collection of segments with each segment representing an agent. These agents competed with each other to minimize their individual volume (i.e. individual utility) and ultimately the volume of the entire beam (i.e. global utility).

The strategies, i.e. possible cross-sections (A_i) of the beam segments, were assigned randomly to each of the agents from the range of values between 1045 and 21613 mm3 for the particular iteration. The details of values and further references can be found in [5]. The range of the area values were considered same for all the agents. The global objective was to minimize volume of beam shown in the Fig.2. The global objective function was represented as

$$V = \sum_{i=1}^{N} v_i = \sum_{i=1}^{N} A_i \times l_i \tag{10}$$

The length of each segment was considered constant 100 mm. The starting probabilities assigned for each strategy was uniform. As there were 42 strategies for every agent, the probability of every strategy was 1/42. The Boltzmann temperature T was assumed 30 at the first instant. The step parameters for updating temperature and probability of the agent, were $\alpha_T = 0.9$ and $\alpha_{Step} = 0.98$ respectively.

The procedure explained in Section III was applied to reach the convergence. The procedure was terminated when the difference between two successive probability updating was less than the predefined value (i.e. in this case, $\| \vec{q}^{\,k} - \vec{q}^{\,k+1} \| \le 0.0009$).

V. Results and Discussion

The problem was coded in MATLAB 7.4.0 (R2007A). The updating of the probabilities was done until the convergence criterion was satisfied. As the samples were drawn at random, once the convergence criterion was satisfied, the system was sampled again. The procedure was repeated 10 times. The results of the entropy are plotted below in Fig.3. These results are for Trial 2, 3 and 5.

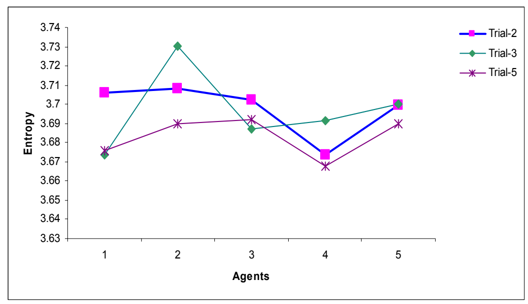

Figure 3. Entropy plot

The results of five trials are shown in Table I and II. It is evident from the results in Table I and Table II and the plot in Fig.3, that the maximum entropy is the one, which directed towards the combined strategy having highest probability of contributes towards the minimization of the beam volume.

From the results it is clear that the approach worked satisfactorily as agents selected the strategy with maximum probability, which was contributing maximum towards minimization of the volume. On the other hand, the strategy with least probability, contributed least towards minimization of the volume. This is evident when compared the corresponding rows of Table I and Table II.

TABLE I. VOLUME WITH MAXIMUM PROBABILITY

Agents (Segments)	Volume Selected with Maximum Probability ($\times 10^7$ mm^3)				
	Trial 1	***Trial 2***	***Trial 3***	***Trial 4***	***Trial 5***
Agent 1	0.3726	0.3201	0.5234	0.3127	0.2721
Agent 2	0.3939	0.2633	0.3741	0.5526	0.3939
Agent 3	0.5553	0.4145	0.4087	0.3213	0.3454
Agent 4	0.3313	0.6810	0.2020	0.2603	0.5481
Agent 5	0.3046	0.1910	0.4501	0.4511	0.3611
Total Volume	1.9577	1.8699	1.9586	1.8980	1.9205

TABLE II. VOLUME WITH MINIMUM PROBABILITY

Agents (Segments)	Volume Selected with Minimum Probability ($\times 10^7$ mm^3)				
	Trial 1	***Trial 2***	***Trial 3***	***Trial 4***	***Trial 5***
Agent 1	0.8109	0.8193	0.8202	0.8797	0.7906
Agent 2	0.8389	0.6449	0.8163	0.7690	0.7744
Agent 3	0.9027	0.7675	0.8235	0.9371	0.7455
Agent 4	0.8196	0.7721	0.8129	0.8030	0.9315
Agent 5	0.7706	0.9921	0.9900	0.7434	0.8366
Total Volume	4.1427	3.9966	4.2630	4.1323	4.0786

The little inconsistency in the final volume was due to the randomness in the initial selection of the strategies. Also, it is evident that there is need for including constraints in the strategy selection process.

VI. Conclusions and Future Work

The methodology successfully demonstrated the application of COIN using PC theory. In the theory of COIN, individual agents competed to minimize their individual/local contribution to global output. This minimized the global output correspondingly. The application results are encouraging and further investigations will be carried out with coupled variables as well as a constrained approach to problems such as the Traveling Salesman Problem and path planning of Multiple Unmanned Aerial Vehicles (MUAVs). It was also evident that the growing complexity can be handled by dividing the system into smaller subsystems or agents.

VII. Acknowledgements

The authors would like to acknowledge Dr. Rodney Teo and Ye Chuan Yeo (DSO National Laboratories, Singapore) for their useful discussions on this topic, as well as David H. Wolpert (NASA Ames Research Center, Moffett Field, CA, USA) and Dr. Stefan R. Bieniawski (Boeing Phantom Works, Seal Beach, CA, USA) for their clarification of some of the Probability Collectives concepts.

References

[1] T. Basar, G.J. Olsder, Dynamic Noncooperative Game Theory, ISBN: 0 89871-429-X, 2nd Edition, 1995.

[2] K. Tumer and D. H. Wolpert, "Collective intelligence and braess' paradox", AAAI/IAAI, pp. 104-109, Austin, TX, 2000.

[3] E. T. Jaynes, "Information theory and statistical mechanics," Physical Review, vol. 106, p. 620, 1957.

[4] D.H. Wolpert, N.E. Antoine and I.R. Kroo, "Discrete, Continuous and Constrained Optimization Using Collectives", 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 30 Aug-1 Sept, Albany, New York, 2004.

[5] N. Turkkan, "Discrete optimization of structures using a floating Point Genetic Algorithm", Annual Conference of the Canadian Society for Civil Engineering, Moncton, Nouveau-Brunswick, Canada, 4-7 June 2003.

[6] S.R. Bieniawski, Distributed Optimization and Flight Control Using Collectives. PhD Dissertation, Stanford University, Oct, 2005.

[7] D.H. Wolpert, "Information Theory ― The Bridge Connecting Bounded Rational Game Theory and Statistical Physics", Complex Engineering Systems, pp. 262-290, 2006.

[8] D.H. Wolpert, K. Tumer and J. Frank, "Collective intelligence to route internet traffic", Advances in Neural Information Processing Systems-11, MIT Press, pp. 952-958, 1999.

[9] D.H. Wolpert and K. Tumer, "Optimal payoff functions for members of collectives", Advances in Complex Systems, 4(2/3):265–279, 2001.

[10] D.H. Wolpert, N.E. Antoine, S.R. Bieniawski and I.R. Kroo, "Fleet Assignment Using Collective Intelligence", 42nd AIAA Aerospace Science Meeting Exhibit, Reno, Nevada, USA, 2004.