ABSTRACT
Behaviors of the colonies of small unsophisticated agents have been analyzed in the literature with the purpose of developing efficient algorithms to solve complex, dynamic and burden problems in other societies. Among them, only a few research have been conducted in the area of swarm cognition which tries to understand the cognitive behaviors exhibited by human brain by using the cognitive behaviors demonstrated by a colony as a self-organized entity. In this aspect, the role of a neuron and a role of an insect have been equally considered as an unsophisticated agent which adjusts its actions according to the fluctuations of local environment without knowing any global information. The cognitive behavior, such as effective labor division of honeybees at food foraging process, was analyzed in this paper and has been exploited under operant conditioning. The paper has proposed a simple but effective computational model which demonstrates that, the positive reinforcement and the negative reinforcement in operant conditioning are the real factors that affect to the emergent of cognitive behaviors at swarm level when swarm is observed as a self-organized entity.

General Terms
Artificial Intelligence, Swarm Cognition

Keywords
Swarm Cognition, Operant condition, Honey Bee colonies.

1. INTRODUCTION
A swarm is a large number of unsophisticated agents with limited capabilities, interacting cooperatively and locally among themselves, and environment, with no central control. These local interactions of the agents in a swarm allow the emergent of globally interesting behaviors that is necessary for their survival. The interaction of these social insects can be direct or indirect. Visual or audio contact, such as waggle dance and tremble dance, of honeybees are examples of direct communication while stigmergy or pheromones based communication between social insects are some examples for indirect communication[1].

The collective intelligence behavior of these natural organizations such as food foraging, division of labor, nest building, etc. are emerged in these colonies when the swarm is self-organized. Self organization is a set of dynamical mechanisms whereby global representation appears from the interactions of its lower level components. The ingredients of self-organism are multiple interactions which results in positive feedback and negative feedback that allows amplification of random fluctuations and control the evolution.

This analysis of the natural self-organized systems has been further diversified into the development of efficient algorithms when it comes to the discussion of how nature-inspired self-organized societies can be used to solve and optimize the complex problems in other societies. This swarm-based algorithms are capable of providing low cost, efficient, and robust solutions to solve complex problems in other information societies. Ant Colony Optimization (ACO)[2], Artificial Bee Colony (ABC)[3], and Particle Swarm Optimization (PSO) are such significant swarm algorithms which have been introduced by observing the ants food foraging, bee food foraging and nest selection, and birds flocking respectively. The principle of these algorithms are successfully applied in image and data analysis, machine learning, operational research, and in finance and business applications[4]. Moreover many significant attempts[5,6] have been made in developing models to demonstrate the dynamicity of these societies, and to investigate how random fluctuations of the key parameters of self-organization affect to the decision making process of the swarm.

In the field of Swarm Cognition, the cooperative interactions of swarm that is necessary for the emergent of collective intelligence such as census decision making, finding the shortest path to the food sources, etc are exploited in the perspective of understating higher cognitive behaviors in human beings. Swarm cognition[7] works on the basic postulation that a neuron as a part of the brain can be expressed in similar to a social insect as a part of colony. A neuron in isolation has very limited capabilities and depends only on local interactions, however, brain demonstrates highly complex cognitive process similar to what swarm displays as a colony. Self-organization is the common mechanism that enables these simple units to display higher cognitive behaviors through the cooperative and local interactions between them. Therefore, behavior of a colony and cognitive process of a brain can be explained using simple rules of self-organization. In self-organizing communities each individual either a insect or a neuron acts according to the information it receives from its local community, without having any global representation, but by following simple individual rules.

2. SELF-ORGANIZATION IN HUMAN BRAIN
In order to understand how the key components of self-organization affect to the decision making process of human brain, we explored the swarm of honeybees food foraging process as a decision making process, which demonstrates smart division of observers in the colony into foragers and receivers to have efficient nectar take in and storing mechanism. This decision making process can be characterized by the information which is processing through
multiple local communication and the adaptive behavior of bees which adjust their responses according to their neighborhood environments[8]. This adaptation of honeybees is similar to the adaptation made by neurons when learning occurs at human brain. There a neuron adjust its internal metabolic processes according to its local environmental fluctuations by updating its synaptic strength by creating new synaptic connections or by withdrawing existing synaptic connections[9].

Two key types of learning mechanisms have been discussed in the literature which define the change of behavior according to the responses of the surrounding environments. They are: classical conditioning and operant conditioning.

2.1 Classical Conditioning
Classical conditioning [10] explains the learning as: it occurs through the associations between neutral stimulus and environmental stimulus. While in operant conditioning[11] learning occurs as changes in behaviors of an individual that are results of the individual's responses to the events that occur in the environment. For example, assume a dog receives tasty foods, soon after it hears a bell. The dog is very happy to receive the food and he starts dancing by wagging his tail. Repeating this process many times, the dog starts dancing when he hears the bell. Pavlov[10] explained this change of behavior of the dog under classical conditioning. He described the sound of the bell as the neutral stimulus, presentation of foods to the dog as environmental stimulus. The dance of the dog is called naturally occurring response. So by associating this neutral stimulus with the environmental stimulus, the sound of the bell could alone can make the dog dance. Thus, by building this association, the neutral stimulus becomes conditioned stimulus while the dance of the dog is the conditioned response. The conditioned response is the learned response to the neutral stimulus. By pairing and un-pairing this association between the neutral stimulus and environmental stimulus, the dog can be made to learn or unlearn the conditional behavior.

2.2 Operant Conditioning
On the other hand, in operant conditioning, learning is described as a change of behavior that is resulted by causes of actions and its consequences [11]–[13]. According to Skinner, three types of operants can be identified based on the type of responses of the environment that change the probability of the behavior being repeated: neutral operant (if the responses do not change the probability of the behavior being repeated), reinforces (if the responses increase the probability of the behavior being repeated and these reinforces can either be positive or negative), and punishers (if the responses decrease the probability of the behavior being repeated. The negative reinforce increase the probability when behavior being repeated is withdrawn). The key differences between the two learning approaches are: in operant conditioning the learner is actively and voluntarily participated while in classical conditioning the learner is passively involuntary participated. In fact many studies in the literature have thoroughly discussed the behavior of honeybees when they operate under classical conditioning[14] and rewarding mechanism[15]. In our research, we focused on an extensive study about the adapting behavior of honeybees under operant conditioning than classical conditioning because the role played by honeybees in a colony is active and voluntary.

This paper postulates that cognitive behaviors that are emergent from a swarm as a self-organized system can be explained using operant conditioning. And thereby the swarm behaviors can be used to explain the emergent of cognitive behavior at human brain. By doing so, we argue that positive reinforcement and negative reinforcement in operant conditioning are the key factors that control the self-organization of any dynamic system and the emergent of cognitive behavior.

The paper is organized as follows: section 3 describes the nectar foraging process of honeybee colonies, section 4 mathematically models the process described in section 3. And section 5 presents the results of the simple computational model that align to the model presented in section 4. Section 5 discusses and concludes our findings.

3. THE FUNCTION OF NECTAR FORAGING OF HONEYBEES
The complexity and dynamicity of the nectar foraging of honeybees was analyzed in the means of identifying key parameters that are necessary for efficient nectar gathering process. It is well known that honeybees communicate through various communication channels such as using pheromones, or tactile dancing with or without some vibrating sounds. Among these techniques, waggle dance is the key communication channel that has been used by honeybees to inform new rewarding flowering sites to the colony and to recruit new foragers to the newly found flower patches[16]. The process of waggle dance with its dependent parameters in general ecology of nectar foraging of honeybees can be briefly summarized as follows:

A honeybee colony is generally composed of a queen, workers (all workers are female) and male or drones. The worker bees are almost completely responsible for caring hives, such as cleaning the hive, caring the larvae and youngs, feeding the queen and the drones, making honey, and gathering and storing nectar, water, etc. Therefore around 90% of the bees in the colony are workers, from 10% of them are scouts who find flower patches by searching[17]. A scout keep searching until its energy level is depleted or it finds flower-patches. If it found a flower patch, it comes to the comb and unloads the nectar sources to receivers at the comb. The receiver takes it to the storage area of the hive.

The returned scout can either be a scout again, or a forager, or an unemployed bee. If the returned scout feels that the flower patch from where it brought the nectar is in high quality and quantity, it performs waggle dance to recruit more foragers. If it is in considerable quality it simply returns and brings the nectar from the flower patch. Otherwise, a scout can forget the visited flower patch and settle as an unemployed bee until it is recruited, or it can become a scout again searching for new flower patches.

Sometimes forager bees perform tremble dances to get unemployed bees to engage in nectar-receiving task [18, 19]. Key message that a forager wants to convey through this tremble dance is that it has found more-rewarding nectar source and no enough receivers to unload them efficiently. Meanwhile by performing this dance the forager tries to inform other mate-foragers not to recruit additional foragers to their nectar sources. Therefore, the tremble dances of honeybees help to the colony to keep the balance between nectar storing and nectar take-in. Once foragers unloaded their nectar they may start to perform waggle dance to recruit unemployed foragers to visit their explored flower patches. If
a forager bee had to wait too long to unload the nectar (when it was unable to find a receiver), then the forager bee does not perform the waggle dance to recruit additional foragers bees, because it does not have enough receivers to handle the unloading process. The receiver bee may get delayed for searching further away through the hive when the hive has not enough vacant storage cells, and it is almost full.

Waggle dance[6,16,20-22] is a communication behavior which conveys the information about location, and quality (high concentration, distance, easy to collect, etc) of the food sources that have been found. A dancing bee runs forward and performs the waggle dance as shown in the figure 1, while she is on the run, she vibrates her abdomen laterally and then comes back to her starting point. According to the sources in the literature, a distance to the food source is proportional to the length of this waggle run and the angle of the run to the sun represents the direction of the food sources. The higher the quality of nectar source, the higher the number of waggle runs per dancing bout and that increases the number of recruiters. Furthermore, the number of recruiters increases in proportion to the probability of returning forager would dance and the number of waggle runs made by her per visit. Through this mechanism honeybee colonies gather nectar more efficiently by sending their foragers to the better flower patches by abandoning less reward flower-patches, recruiting to more-rewarding patches, and searching for new patches.

Fig 1: Waggle dance of a honeybee

According to Seeley[21], honeybees measure the profitability of located nectar source by sensing the energetic efficiency of their foraging. Therefore, the number of waggle runs per dance is not directly a linear function of closeness of the located nectar source to the hive but the energy expenditure per foraging trip. Moreover, through its waggle dance, a dancing bee reports on the current level of energy profitability of her forage site. Therefore, the number of waggle runs per bout is high when the nectar source is abundant, and it is low when it is scarce. The dancing bee does not perform her dance in one place but distributes it over the dance floor. Therefore when number of forage sites are being reported on the dancing floor, foragers can easily take a random sample of the dance information. This allows the colony to allocate foragers to more-rewarding food sources.

4. MODELING FOOD FORAGING FUNCTION OF HONEYBEES

The emergent of cognitive behavior at swarm level, a decision making process which divides the labors at the colony efficiently to streamline the food foraging process was analyzed in our research. This analysis was carried out to identify the key factors or reinforcements that enable the swarm as a self-organized system to demonstrate this highly cognitive task. Here, we have restricted our analysis by considering only the impacts of direct communication of honeybees on this foraging task. Further, we have assumed these bees can remember one flower patch at a time and can harvest only one flower per trip from the comb as in [6].

Let mathematically model the world of swarm as square-grid of length m (m > 0) and odd integer in meters) in which flower patches are placed on the environment according to a Gaussian distribution $N(\mu, \sigma^2)$ where $\sigma$ is $0 < \sigma < m/2$ when the comb is placed at the middle of the square. Let the maximum distance a scout can fly without returning the comb is $l > 0$. Let quality of the nectar concentration at flower patches ($n_f$) are uniformly distributed in three states: low (1 unit), moderate (2 units) and high (3 units). Moreover, the available quantity of the nectar in a flower patch is measured in terms of the number of bloomed flowers it has. Let the number of bloomed flowers are randomly distributed using a discrete uniform distribution from 1 to $N$ ($n_f \sim U(1,N)$).

Then, the overall quality of nectar a flower patch has, is defined by $n_f = n_i \times n_t$. Honeybees, in a colony of size colony_size, are initially distributed to scouts, foragers, receivers, and observers with probabilities $P_t$, $P_f$, $P_r$, and $P_t$ respectively.

We adapted the equation from[22] which measures the quality of the food sources in terms of distance $d$ and $d_{max}$ where $d$ is the distance from the comb to a flower patch and $d_{max}$ is the maximum direct distance from the comb to the edge of the square through the flower patch and, $n_f$ - the overall quality of the nectar at a flower patch. We can define the quality of a flower patch as in eq. (1). As given in the equation, the higher quality flower patches which locate near to the comb gets higher $Q_f$. A scout honey bee starts Levy flight to find these flower patches.

$$Q_f = \frac{d_{max} - d_f}{d_{max}} \times n_f$$

The number of waggle runs ($w$) performed by a honeybee is based on $Q_f$ and it is in the form of $w \propto Q_f$. A returned forager waits time $t_{wait\_to\_unload}$ to unload nectar that have been brought-in to a receiver in a colony;

$$t_{wait\_to\_unload} = \frac{1}{1 + \exp(-1/\text{number\_of\_receivers})}$$

where number_of_receivers is the number of receivers in the colony at time step $t$. The expected time duration a honeybee has to wait to unload nectar $t_{wait}$ is measured in terms of the quality of the flower patches that have been reported to the colony so far. i.e. $t_{wait} = 0.5 - (Q_f - k)$. If the foraging honeybee has to wait $t_{wait} < t_{wait\_to\_unload}$, the forager bee tends to perform tremble dance. The duration of the tremble dance is $t_{wait} - t_{wait\_to\_unload}$. The parameter $k$ is determined by a honeybee by randomly sampling its neighborhood, i.e. the roughly taking the quality of the flower patches that have been reported to the comb; so $k = \frac{1}{n} \sum_{i=1}^{n} Q_i$. 

If \( t_{\text{wait}} - t_{\text{wait_to_upload}} > \theta \), \( \theta \) is a parameter, the forager stops foraging the current site. A given forager bee can turn either to a scout, or an observer. A forager performs a waggle dance mainly based on \( Q_f \), the duration of the waggle bout is restricted by if any tremble dances performed at the colony. If forager bee encountered \( trem \) number of tremble dances that have been currently reported in the comb, then \( w = Q_f - \frac{\langle trem \rangle}{c} \), where \( n \) is the size of the random sample and \( c \) is a constant.

A observer bee can turn to a forager, or a receiver or remain as a observer based on the dominant requirement that is exhibited by the random sample that it has taken from the colony. A receiver bee can change its role to a observer if it has to wait too long to receive nectar, \( t_{\text{wait_to_receive}} > \theta \); where \( \theta \) is a parameter.

Under this mathematical model, the energetic efficiency of the nectar source of flower patches foraged by honeybees influence the entire nectar foraging cycle of the colony. Furthermore, this energetic efficiency, is the key factor that determines the duration of waggle runs performed and in fact it affects to the proper division of labors in the colony.

Therefore, we can postulate that, energetic efficiency of the nectar source is the positive reinforcement that a honeybee receives to perform waggle dance and to recruit more foragers, and the arises of cognitive function which makes the decision on dividing the labors at the colony appropriately. On contrast, the time duration, a forager bee has to wait to unload its nectar source to a receiver is the negative reinforcement that a honeybee receives from the environment to discourage the foraging of low quality flower patches compared to what have been reported to the colony.

So we can postulate that these two forms of reinforcements from the environment allow the entire colony to keep its balance between nectar take-in and storage process by efficiently dividing its labors to appropriate task according to the dynamic environmental fluctuations (according to the distribution of the quality of the nectar sources that have been reported) and according to the information communicated by its surrounding environment to the honeybees.

Therefore, a cognitive process, such as decision making which is exhibited by a self-organized system, is mainly arisen as learning adjustments which are made by each individual as responses to environmental positive and negative reinforcements.

5. RESULTS
A simple computer model was developed to simulate the effect of the positive \( (Q_f) \) and negative \( (t_{\text{wait_to_upload}}) \) reinforcements on the labor division at swarm food foraging process.

The total quality of a flower, i.e. \( Q_f \) (referred as qlyFlower in figures), the mean of the quality of flower patches that have been really reported to the colony, i.e. \( k \) (referred as meanQlyFlower in figures), the time duration a honeybee waits to unload the nectar, i.e. \( t_{\text{wait_to_upload}} \) (referred as wTime in figures) and the expected time duration a honeybee must wait to unload the nectar, i.e. \( t_{\text{wait}} \) (referred as mwTime in figures) were monitored in this process in addition to the recording of the number of receivers, the number of foragers and the number of observers in the colony.

The computer model had a colony of 150 foragers, 130 observers, and 120 receivers. The overall quality of a flower patch, i.e. \( Q_f - (0,1) \) was randomly generated. Initially the value of \( k \) was set to 0.3. If \( Q_f > k \), a observer is turned to a forager, otherwise a forager is turned to an observer. If \( t_{\text{wait}} < t_{\text{wait_to_unload}} \) we increase the number of receivers by one and decrease the number of by one. Otherwise we decrease the number of receivers by one and increase the number of observers by one. With this simple model we were able to demonstrate the effect of positive and negative reinforcements on each individual's action and thereby the emergent of cognitive behavior at swarm level.

From figure 2 to figure 5, each figure consists of three subfigures, namely, (a), (b) and (c). Subfigure (a) demonstrates the distribution of honeybees in the colony in terms of the number of receivers, the number of foragers and the number of observers when all the four parameters (qlyFlower, meanQlyFlower, mwTime, and wTime) were calculated as defined in the mathematical model. The subfigures (b) and (c) show the fluctuations of qlyFlower, and meanQlyFlower, and meanmwTime and wTime related to the scenario depicted in subfigure (a) respectively.

Figure 2 shows the situation when we let the system to run around 100 steps without enforcing any other constraints. As shown in figure2, the increases of the number of foragers have decreased the number of observers and has increased the number of receivers.

As shown in figure3. we artificially set the value of meanQlyFlower to 0.8. So that we artificially informed the colony that the colony has been reported by high quality flower patches which are in the range of 0.8. As shown in figure 3 (b), since the quality of the flower patches that have been really reported by the honeybees are less than this artificial value, wTime of foragers have been increased compared to the expected waiting time, i.e. mwTime.

In order to observe behaviors of the colony for lower quality flower patches, we artificially set meanQlyFlower to 0.4. So that we artificially informed the colony that it has been reported by low quality flower patches and it is in the range of 0.4. As shown in figure 4(b), since the quality of the flower patches that have been really reported by the honeybees are higher than this artificial value, honeybees were not requested to wait too long to unload the nectar and in fact the time they really waited is less than the expected time duration, i.e. wTime < meanwTime. This adjustment of the system has decreased both the number of observers and the number of foragers and consecutively it has increased the number of receivers.

Finally, we evaluated the behavior of the system by artificially setting mwTime to 0.7 as shown in figure 5. As shown in figure 5(c) when wTime is higher than the expected waiting time, i.e mwTime, the system has increased the number of observers, and decreased the number of receivers. When wTime was less than the mwTime, the system has increased the number of receivers and decreased the number of observers.
Fig 2: Distribution of Honeybees in the Colony

Fig 3: Distribution of Honeybees in the Colony when meanQtyFlower was artificially set to 0.8

Fig 4: Distribution of Honeybees in the Colony when meanQtyFlower was artificially set to 0.4
6. DISCUSSION
As shown in figure 2 when the colony was reported by higher quality flower patches compared to what have been reported to the colony so far, the system has increased the number of foragers, and the number of receivers while the number of observers have been decreased. This was an accepted behavior by the system which has encouraged the foragers who have reported higher quality flower patches by decreasing their real waiting time than the expected waiting time.

The qltyFlower that has been reported to the colony has affected to the system as a positive reinforcement, as given figure 3. By increasing the number of observers in the colony, system has assumed that, it allows the foragers who have reported quality flower patches in 0.8 range, to broadcast their foraging sites effectively to the colony. So the foragers who have reported poor quality flower patches have been discouraged by increasing their real waiting time, i.e. wTime, than the expected waiting time, i.e. mWTime. Therefore, in this scenario, wTime has been applied on the system as a negative reinforcement to discourage foraging of poor quality flower patches.

When quality of the flower patches that have been reported is set to 0.4, the system has encouraged foragers who have reported quality flower patches (figure 4(b)) by decreasing their real waiting time than the expected waiting time. So it has increased the number of receivers and the number of foragers and decreased the number of observers. Therefore, as expected, the system has considered qltyFlower as a positive reinforcement.

Finally when we set wTime to 0.7 as in figure 5. It has negatively impacted on the system behavior. That is, when wTime is higher than the expected, it has further decreased the number of receivers, and increased the number of observers, by assuming that, the system is being reported by low quality flower patches. And accordingly, when wTime is less than the expected waiting time, it has increased the number of receivers and the number of foragers, while it has decreased the number of observers, by assuming that the colony has been reported by higher quality flower patches.

Through this simple but effective simulation, we could show that the operant conditioning, which explains the responses made by a simple individual to the environmental positive and negative reinforcements as learning adjustments, is the underline phenomena of self-organization of dynamic system and the exhibition of cognitive behavior, such as division of colony labors effectively at the nectar foraging process.

Therefore, similar to the emergent of cognitive behavior at swarm level, the cognitive behavior emergent at human brain can also be explained using operant conditioning by taking the brain as a self-organized system. As the future work, the deeper analysis into the emergent of a particular cognitive behavior at human brain is needed to be exploited under operant conditioning as a means of discovering the underlying positive reinforcement and negative reinforcement. Doing so, the adjustment that happen at individual neuronal level to these reinforcements can be taken as the process that resulted in emergent of cognitive behavior.

7. REFERENCES


