Collective decision making: From nest-site selection in honeybee swarms to kilobot ensembles

A computational study on complex systems

Institut de Sistemes Complexes de la Universitat de Barcelona (UBICS)

Author: David March Pons Advisor: Carmen Miguel Lopez

Complex systems are formed by the union of many individuals with a simple behaviour, but whose interactions give rise to unexpected properties that are not predictable from the single knowledge about just one these individuals. Flocks of animals such as birds or sheep moving as a collective, the coordinated action of neurons forming the nervous system, societies interacting and forming commerce routes or the spread of a pandemic in a community are examples of complex systems.

The study of complex systems is an interdisciplinary area that, starting from a physico-mathematical background and making use of computational tools, cuts across as many areas of knowledge as those systems of interacting components can be found. For instance: biology, ecology, economics, sociology or humanities, to cite some of them. One of its most interesting features is that, from the knowledge and the data that we have of the real world, we can build models that help us understand deeper the dynamics of complex systems, allowing then to make predictions on situations that can happen in the real world, for instance in a situation of a global pandemic.

In this article we will study a process that can take place in a complex system: a **collective decision-making process**. This is a situation in which a group of animals or humans must choose the best possible option for them, without the presence of a leader that speaks for the whole group but rather by debating and reaching a decision by a group majority. The latter case is what is known as a **decentralized decision** process. Think for instance of a group of friends that meets to have dinner, but first they must choose at which restaurant to. A discussion will be held about which is the best option for the group, and despite different opinions may exist a final quorum will be achieved. Another example, which is a relevant matter of study in ecology, is how colonies of insects such as ants or honeybees collectively coordinate to gather a specific food source or to build a nest. Particularly, we will study how a swarm of honeybees is able to select a site to build their nest.

The article is divided in three parts. First, a brief description of the nest site decision process in honeybee swarms is given. Then, we will explain how from the knowledge about such a complex system we can build a computational model. Finally, we will discuss the results we can obtain from the computational model and what they tell us about a real honeybee swarm.

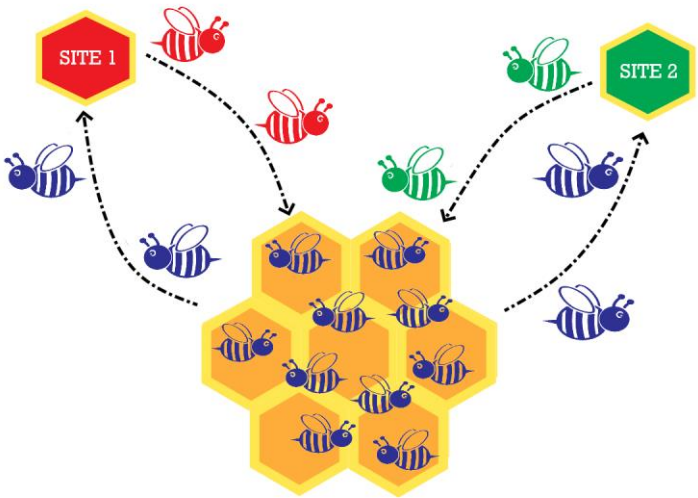

Figure 1. *Schematic view of the exploration process of honeybee swarms. A fraction of uniformed bees leaves the hive, explore and discover different possible sites. Then they will return to the nest and there they will perform the waggle dance.*

I. The nest decision process in honeybees

Every year, at the end of spring, honeybee colonies split. Around two thirds leave the nest and build a new one in some other place. From these leaving bees, a fraction of them will scout the environment in order to identify possible locations where to build a new hive (blue bees in figure 1). Once they have discovered an option, they return and communicate their findings to other bees that have remained in the hive (red and green bees in figure 1). They do so by performing the **waggle dance**. This is a pattern of movement that consists of a waggle run, in which bees agitate the rear part of their body, towards the location of the site they are announcing followed by a return phase to restart the waggle run (see figure 2). The key feature of the waggle dance is that it is **sensible to the quality of the site** that the bee has discovered. Biologists, by doing in field experiments with real honeybees, have learned that not all places where bees can build a new hive are equally fit for them: ideally, they will search for big places but also hidden from predators. Waggle dances will be made more attractive when bees estimate that the site they have discovered is more fit for the swarm, or in other words, has a higher quality. Attractive waggle dances consist of longer waggle runs and quick

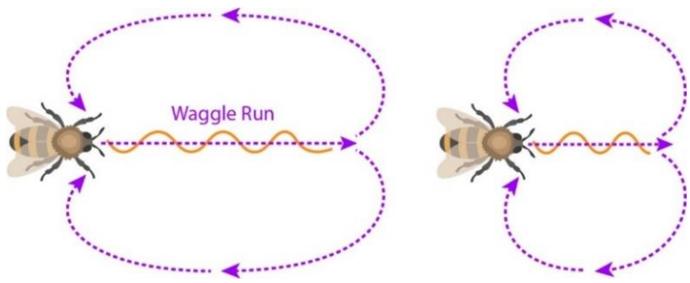

Figure 2. *Representation of the waggle dance: a waggle run forwards followed by a return.*

return phases; also, this cycle is repeated more times, increasing the duration of the waggle dance.

Bees that have not left the hive to explore may start mimicking a very attractive waggle dance that is being performed by other bees. Then we can say that they follow the opinion of the bee they imitate, or equivalently, commit their opinion to the site the dancing bee was already committed. As the waggle dances for a very good option are longer, the chance that an uncommitted bee sees an attractive waggle dance will be greater than if the waggle dance was shorter. This will increase the chances that bees copy waggle dances for the best possible options, and after some time there will be a great part of the swarm dancing for - or committed to - a good option. This is a crucial mechanism that bees have developed over the years to increase their chances of survival as a group, by selecting good location for where to build their nest or, in a similar process, by choosing profitable food gathering sources.

During the decision process - that lasts no more than the sunlight hours in the span of two days - bees will be exploring, resting or dancing for possible options, because they explore or because they followed the opinion of other bees. Combining the exploration of the environment and the waggle dance following mechanism we expect that a global opinion will be formed in the swarm for, presumably, the best possible option they have.

Nevertheless, many factors come into play in the process that may difficult the final decision for the best possible option. For example:

- The available sites may be similarly fit for the swarm, producing waggle dances of similar duration. This will make hard for bees to distinguish which is the waggle dance for the best possible option.
- Sites may be difficult to discover, especially the ones that are best in all aspects.

For instance, think of a situation in which there is a very good option, but far away from the original hive and thus difficult to discover. Will it be enough that only a few bees discover and perform the waggle dance for it? Communication between bees will play a relevant role in such a situation. If bees highly rely on the waggle-dances of other bees a large fraction of the swarm may end up mimicking the few initial dances for this good but hidden option, allowing the swarm to choose the best possible site in an unfavourable environment. Instead, if they take more independent decisions (they usually prefer to explore rather than copy other waggle dances), the easiest site to discover will probably by imposed, as bees won't pay particular attention to these few attractive waggle dances.

II. Modelling the nest site decision process in honeybees

With the knowledge we have about how a real swam of honeybees works out this process we can design a computational model that helps us understand the outcome of the problem bees face in many different situations.

We must begin by designing an algorithm that replicates this process. As these processes taking place in nature are inherently random, this is, the outcome will never be exactly the same starting from the same initial conditions, we must implement what is known as a **stochastic algorithm**. This is an algorithm that following a set of rules that defines the actual problem, use randomness to achieve the final state of the process, not following the same succession of events twice. Think for instance of *parchesi* game (or any other dice games), where the initial state and the rules are always the same, but the development of the game is never repeated twice.

To build such a stochastic algorithm we must identify which features of a real swarm decision process can be implemented as parameters of the model. Then, with these parameters, define a simple set of 'game rules' that replicate what bees do in the real world.

Parameters of the model

The parameters of the model must serve to replicate a situation or a scenario that a honeybee swarm may face. From what we learned in the previous section these parameters will be:

i. Number of sites available.
ii. Probability to discover the site: for each site, how hard is it to be discovered by a scout bee.
iii. Quality of the site: for each site, the duration of the waggle dances a bee does to announce this site.
iv. Interdependence: the fraction of waggle dances that are initiated because bees copy other bees rather than independently explore.

With these parameters we can model any situation, for instance the one brought up in the previous section: there are two sites, the first with a small quality and the second with a high quality. The first has a greater probability to be discovered than the second. How much interdependence will we need to reach a majority of bees committed to the second option?

The algorithm (the rules of the game)

But, having defined these parameters, how do we replicate the decision process? As explained earlier, during the decision process bees can be committed to no option, because they are exploring or resting (uncommitted state) or committed to one of the possible options, when they are performing the waggle dance (committed state). They will go on and forth between these states in quasi-random way. It's relevant to say quasi-random, and not just random, because the parameters of the model specify that maybe a site is easier to discover, or maybe the waggle dances for a site are longer and more visible, so the event of committing to a site is not like just throwing a dice or tossing a coin, but the setting of the experiment can favour starting waggle dances one particular site.

The algorithm will specify how bees can go from the uncommitted state to the committed state. The hallmark the algorithm will be making the event of starting a waggle dance a quasi-random event, like throwing a loaded dice. Each face of the dice represents committing to one possible option, plus an additional face for the chance to not start a dance. Each side that represents a possible site will be differently loaded. Combining the parameters of the model, there are two factors that matter when loading the dice:

1. The chance to explore and discover the site. If the swarm has a small interdependence, this will be the main factor contributing to load the dice as bees will prefer to explore rather than to trust other bees. Also, if the chance to discover the site is bigger, this factor will contribute more when loading the dice.
2. The chance to copy a waggle dance. If the swarm has a high interdependence this will be the main factor contributing to load the dice, as bees will rely a lot on other bees' waggle dances, being less prone to explore by themselves. Also, if there are many bees already dancing for this site, this factor will contribute more to load the dice as the chances that a bee sees a waggle dance for this site will be higher.

The simulation of a decision process will proceed in discrete time steps, or turns, at which the opinion of each bee will be updated. At the start all bees will be in an uncommitted state, with no information of the available sites, as a real swarm.

At the start of each turn the dice will be loaded as specified by the simulation parameters and the number of bees dancing for each option. Then it will be thrown for each uncommitted bee in order to update their state. For example, if there are two available sites the dice will have three faces. As represented in figure 3, there will be different probabilities that the dice shows a particular outcome: p_1, p_2 are the chances to start dancing for sites 1 or 2 and p_0 is the chance to remain exploring. Once a bee chooses to dance for a site, it will do so as many time steps (or turns of the game) as the quality parameter of the site. In this way, in the following time steps bees that are uncommitted may see these bees dancing for their site,

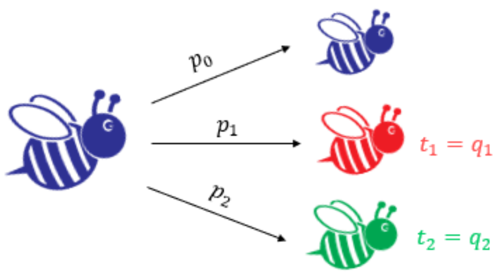

Figure 3. *An uncommitted be will have different chances to commit to different options or to remain uncommitted. Once committed to an option, they fix the number of turns they will dance afterwards.*

increasing the chances that uncommitted bees copy their waggle dance and thus commit to their option.

When bees that are committed to an option finish their waggle dance, they return to the uncommitted state, forgetting the option they had discovered earlier.

Simulations of the nest site selection process

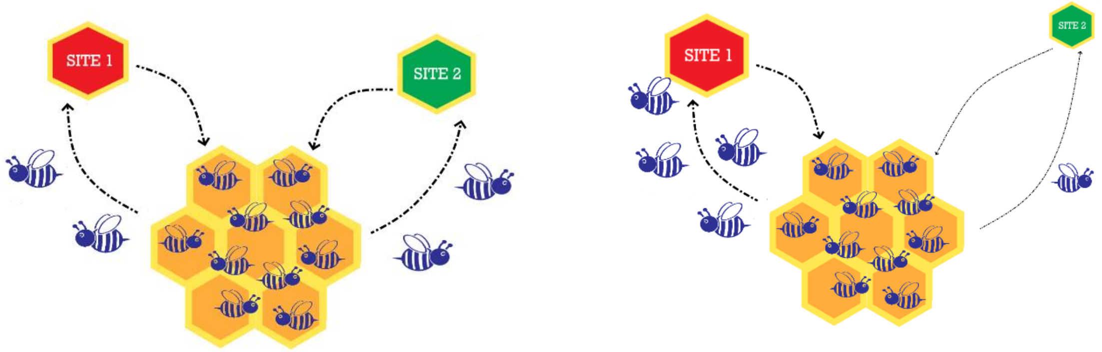

Figure 4. *Representation of the exploration process in different conditions. Left) each site is equally likely to be discovered, so the number of scout bees that will find each site is the same. Right) The first site is more likely to be discovered than the second; consequently, more bees will discover and dance for the first site than for the second.*

Once we have designed an algorithm that is capable, from a few simple but crucial insights about the real world, to replicate the decision process of honeybees, we can make computational simulations of this process at different initial conditions. In this article we study the situation in which there are two possible sites, one of higher quality. Different initial conditions will be different degrees of interdependence and different chances to discover each site, for instance when they are equally likely to be found or when one site is easier to be discovered than the other (figure 4).

When we do simulations of this process, starting from all bees uncommitted, it arrives a moment that the number of bees committed to every option remains practically constant in time. This is a stationary state. As in any collective decision process, such as elections, we want to know which option wins and by how far it wins. Concretely we want to know if the swarm chooses the best possible option between the two. We study if

i. The swarm has reached a simple majority for the best option: weak consensus.
ii. The swarm has reached a stronger majority, such a two thirds majority for the best option: strong consensus.

The fact that we distinguish between majorities, tagging a simple majority as a weak consensus is that in this case the colony would practically be split in half. A more beneficial situation for the swarm is to reach a decision where more than half of the swarm is committed to the best option, such a two thirds majority.

Figure 5. *Representations of a stationary state. In the left scenario only a simple majority is achieved for the good (green) option. Instead, on the right scenario a two thirds majority is achieved for the good option.*

III. What do we learn of the decision process with a computational model?

When bees rely more on communication rather than independently exploring the environment, what means that many bees are committing to an option by imitating other bees more long and attractive waggle dances, the decision is always for the better option, be it a strong or weak consensus. Thus, a high degree of **interdependence** or communication is always **beneficial**. This is caused by what is known as a **positive feedback modulation**: starting from a few attractive dances some bees are going to copy them, increasing the number of bees that dance for this option and consequently increasing the chance that other bees might see and follow this option, imposing finally a consensus for the best option. For this purpose, is vital that bees tune the durations of the waggle dances as a function of the quality of the site. Furthermore, relying on the opinions of other bees proves to be crucial in unfavourable situations:

- When the chance to discover the bad option is much higher than the chance to discover the good option, we will need a high degree of communication in order to spread the knowledge about the good option from only a few discoveries. The higher the interdependence is the larger can be the chance to discover the bad option and still impose a strong consensus for the better option. If the interdependence is low instead, the swarm ends up choosing the bad option because bees are not paying attention to the few waggle dances for the good option.
- When the chance to discover either of the two sites is high, what in terms of a real swarm can be related to the fact that there are many of scout bees, many waggle dances are initiated due to independent discoveries. Consequently, the swarm will need a high degree of communication to override the initial opinion that a part of the swarm will have for the bad option and impose a strong consensus. Nevertheless, a weak consensus can always be achieved thanks to the longer waggle dances for the good option.

With these results one could conclude that the most advantageous strategy for the swarm would be to fully rely on waggle-dance imitation to discriminate between the two options and impose a majority for the best possible option. This would mean having only a few scout bees (small chance to discover sites) and trusting their waggle dances to build an opinion for the good site. Nevertheless, we must consider another factor besides how strong the consensus is: how long does it take to reach the consensus. Real swarms don't take very long periods of time to wait for a very good option to appear. Then, if we are simulating this process, we must mind also how long does it take to reach the final decision in our model. And why is the duration of the process important? There are changing factors that may restrain the time window in which certain options are available and the swarm can move to one of them. For instance, changing meteorological conditions or the presence of predators. Then we must highlight that the strategy of having a small number of scout bees will produce very long decision processes. In any case, whatever are the chances to discover both sites we find two clearly distinct situations:

i. When the difference in qualities is notable, $q_1 \ll q_2$, what means that there is a clearly distinguishable good option, the more the swarm relies on communication, the quicker the decision process is. Besides, more bees will be committed to the best option at the end of the process, reaching always a strong consensus.
ii. When the quality of the bad option is close to the quality of the good option, $q_1 \rightarrow q_2$, the swarm will need a high degree of communication to reach a strong consensus. The downside is that the time to reach this stationary decision will also be larger. Here a **speed-accuracy trade off** takes place:
    - Accuracy: reach a very wide majority for the good option by highly relying on communication, but with a long decision process.
    - Speed: reach a weaker consensus by relying less on the waggle dances of other bees but in a much shorter time.

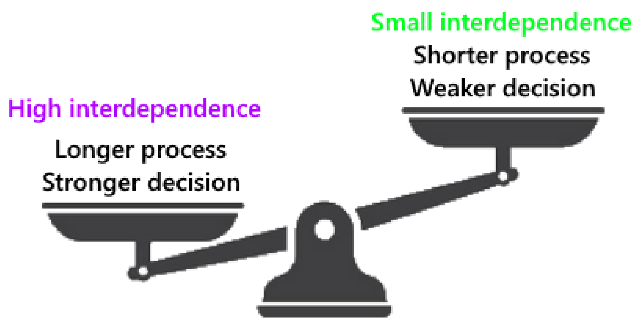

Figure 6. *Speed accuracy trade off. The swarm must balance how much they rely on waggle dances of other bees.*

Finally, another feature we have learned from this study is that the decision process follows a property of large groups know as **wisdom of the crowds**. This property tells that when a group of simple individuals is taking a collective decision, averaging the opinions of all the members of the group, the decision will be more accurate (with less statistical error) when the group is larger. This means that when repeating the stochastic decision simulation many times, the swarm will get the correct decision - a strong consensus for the good option - more times if the size of the swarm is larger. In a real swarm, the larger it is the more accurate or strong the decisions it will take for a given site to build the nest.

IV. Collective decisions in kilobot ensembles

In the recent years this kind of experiments performed via computational simulations has gone a step further. Using a swarm of mini robots, the kilobots, the decision process can be replicated with the same algorithm but with robots representing the bees, in a physical set up.

In these experiments each robot will have to look what are the states of other bees in their proximity (uncommitted, dancing for a site…). With this information then they will compute the probabilities (the loads of the dice) with which they will choose to what option commit, or to not dance for any site. These means that now each robot will have its own dice, as each one sees a different situation in their surroundings.

It is interesting to perform experiments with these robots because they can replicate the imperfections a real swarm of honeybees can have. For instance, as bees are not all equal, robots won't bee either (for instance because of differences in hardware). Also, we can see how a group of robots may form a separate cluster that functions independently of the rest of the swarm.

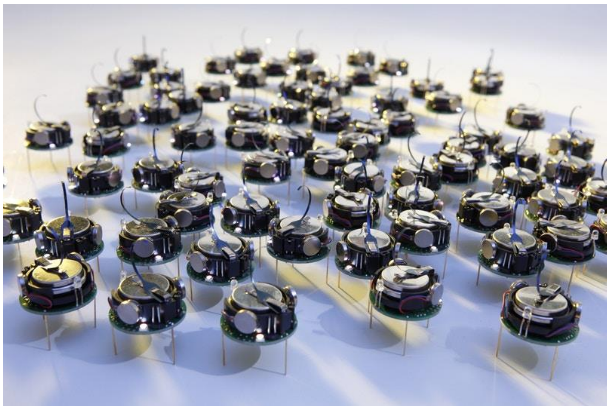

Figure 6. *Kilobots. These are robots of a few centimetres of size, capable of executing the algorithms initially designed for computer simulations.*

Conclusions

We have seen how from the knowledge of how a real complex system works we can build a computational model that allows to study it. The collective decision process in honeybees is a paradigmatic collective decision process that takes place in our world. Building a simple computational model enables us to understand deeper the basics of this mechanism. Once understood, we can complicate our model to represent more complex situations, for instance to study collective decisions in groups of humans.

Furthermore, a collective decision algorithm could have very interesting applications. For example, implementing it in a swarm of robots that communicate with each other to explore the remains of a natural disaster in search of survivors, or to explore regions of other planets that we cannot reach.