

How to Model Honeybee Colonies

David J.T. Sumpter, Uppsala University, Sweden

Mathematics is a way of travelling logically from A to B , where A is a set of precisely stated rules or axioms and B is a set of predictions logically arising from these rules. Mathematical modelling is the art of mapping the real world onto these sets A and B . These are the two basic components of applied mathematics. One part is an abstract, rigorous toolkit for investigating logical truth, the other part is an attempt to align that truth with a more messy reality.

One thing that applied mathematics definitely is not is a specific set of equations. Applied maths is not just the study of partial differential equations. Nor is it just dynamical systems, or stochastic processes, differential geometry of strings or graph theory of networks. If applied mathematics involves getting from A to B , then it does not matter too much how we get ourselves between these two places. The transport method should be chosen on the basis of what works best. If one method is faster or more reliable then we should use it. And we should not expect any one part of mathematics to work best everywhere. I cycle to work, but I take the train between Manchester and London.¹ Once we have worked out what aspect of the real world we want to study, we should map it into mathematics in the best way we can.

The idea that different parts of mathematics are more appropriate in different application areas is not controversial. What is difficult, however, is choosing the correct method for any particular problem. This is where it gets messy, and this is where honeybees come in. How do we build a mathematical model of a honeybee colony? This problem is relevant because it is (probably) more complicated than asking how we build a model of planetary motion, but it is (probably) less complicated than building a model of economics, our own society or the human brain. It is an example of building a model of a complex system.

There is no single correct answer to modelling problems like this. What I am going to do here is give a variety of answers. Or more precisely: I am going to describe some of the mathematical tools that have been used to help answer questions about the social behaviour of honeybees. This turns out to be an extremely rich variety, ranging from simple inequalities and basic statistics, through systems of ordinary and partial differential equations, as well as agent-based and computational models. The modelling of honeybee colonies is a microcosm of applied mathematics.

By describing this microcosm, I hope to give some insight into how we can get the best out of mathematical modelling, and avoid the pitfalls applied mathematicians sometimes fall into.

The best of these insights I learnt from Dave Broomhead, to whom this issue of *Mathematics Today* is dedicated and who supervised my PhD on mathematical models of honeybee colonies. I will return to Dave's influence at the end of this article. For now let us learn from the bees.

1 Why do bees co-operate?

Honeybees are social insects. The queen is the only individual in the colony that successfully reproduces, and the other female workers (males do no useful work) collect food, nurse young, clean and carry out all the tasks that allow the colony to function properly. This would appear to pose an evolutionary problem: the workers get no direct benefit from all their hard work, and the queen passes her genes on to the next generation. Why is there not a mutiny? Or at the very least, why do the workers not try to reproduce?

A first model of this question can be stated in the form of an inequality, known as Hamilton's rule. Workers should help out provided that

$$rb > c$$

where r is a proportion measuring the degree of genetic relatedness between the worker and the queen, b is the benefit gained by the queen if the worker helps her, and c is the cost to the worker of helping out. Honeybees, ants, and other social insects have a genetic system that means r is higher between queen and workers than it is between pairs of workers (who are half-sisters). As a result, if workers do try to reproduce then their sister workers kill their offspring, preferring to rear the offspring of the queen.

It might be tempting to dismiss the equation above as being too simple to merit consideration as a proper model. But Hamilton's rule has led to a whole host of mathematical problems. These arise because a rule like this should be derivable from some set of first principles [1]. In terms of my transportation analogy: a starting set of rules A about animal interactions should be linked to a prediction B about which interactions are stable once evolution has acted. Hamilton's rule should be derivable in terms of interactions between individuals and genes.

You can get some feeling for the challenges of first-principles thinking if you think in terms of game theory and the famous prisoner's dilemma. Two-player games like the prisoner's dilemma can be expressed as a pay-off matrix. The rows are the strategies you (the focal individual) can adopt, which in this case are either co-operating with your partner or defecting. Likewise, the columns are the potential strategies of your partner. The pay-off matrix for the prisoner's dilemma is

	Partner co-operates	Partner defects
You co-operate	$b - c$	$-c$
You defect	b	0

Here b is the benefit you get if your partner co-operates with you and c is the cost you pay if she defects. If you both co-operate then you both get a benefit b but also pay a small cost c . This is the top left-hand entry in the matrix. Similarly, if you both defect then the pay-off is 0, since neither of you pay a cost or get a benefit. Irrespective of what your partner does, your best response is



to defect: if your partner co-operates then by defecting you get b instead of $b - c$, and if your partner also defects you get 0 instead of $-c$.

Consider now what to do if you are related to your partner and any benefit she gets you also receive albeit reduced by a factor r . The new pay-off table is

	Partner co-operates	Partner defects
You co-operate	$(1 + r)b - c$	$rb - c$
You defect	b	0

Now we get Hamilton's rule: you should co-operate whenever $rb > c$. We have also established a simple but clean example of a first-principles derivation. We start by describing a one-shot evolutionary game and end up finding Hamilton's rule as an evolutionarily stable strategy (for a more complete first-principles derivation see Sumpter [2]).

But here is the twist. Why should we assume that honeybees are playing the prisoner's dilemma when they decide whether or not to help out in the hive? There are a whole range of two-player games that might effectively model honeybee interactions. Moreover, there are tens of thousands of bees in a hive, not just two. Once we start to think about how bees interact we see that honeybee co-operation has nothing to do with prisoner's dilemmas. Moreover, the prisoner's dilemma is the only game that has Hamilton's rule as its evolutionarily stable strategy! All other games have some other condition for helping behaviour, with numbers or letters representing the principles upon which the model is based, but none of these have the precise constants $rb > c$.

The problems get even more complicated when we introduce repeated and spatially local interactions. These problems are far from resolved and have become the subject of quite some controversy. One school of thought concludes that Hamilton's rule is largely meaningless [3], while the other side argues that evolutionary problems should be rearranged in order to reveal their own Hamilton's rule, making it an even more general and powerful result [4]. This whole debate has been made even more complicated (and heated) because, irrespective of its mathematical roots, many biologists have found Hamilton's rule to be extremely useful in practice [5]. If it works, why throw it away?

The reasons for honeybee co-operation almost certainly have something to do with genetic relatedness. But co-operation is also selected because of the fact that large numbers of individuals can achieve more together than they can alone. Honeybee colonies become more than the sum of their parts. And this leads us to our next problem. To understand how bees become more than the sum of their parts we need to know more about how they organise their work.

2 How do bees organise their work?

The Wisdom of the Hive by Thomas Seeley [6] is one of the most impressive scientific books ever written. The book describes Seeley's own and others' painstaking experimental work to identify how honeybees choose what tasks to do, how they decide where to look for food and resources, how they communicate with each other, and how they build their nest and regulate its temperature. What marks the book out is the way Seeley then puts together the biological detail he has discovered. The side notes of the book are filled with diagrams of feedbacks, time series and spatial organisation. Seeley takes the honeybee colony and



describes it as a regulatory system. Twenty years before our current era of big data and systems biology, Seeley saw clearly how to combine understanding of the details with the bigger picture of how the system works. This is a challenge that we are still struggling with in developmental biology, neuroscience and sociology.

The way Seeley presented his work lends itself to mathematical modelling. For example, when a honeybee finds food it does its famous waggle dance to advertise the location of food to the other bees. If the bees are dancing for two different sites, as is often the case, then they compete for the attention of other follower bees. One way of converting this competition process into a mathematical model is to think in terms of drawing balls randomly from an urn. If bees dancing for site 1 are black balls and bees for site 2 are white balls, then the follower bees can be thought of as picking black and white balls at random from an urn. If the follower picks a black ball then it starts to recruit to site 1 and thus adds a further black ball to the urn. If it picks a white ball then it adds an extra white ball and recruits to site 2. This process is known as a Pólya urn model.

By seeing bees on a dance floor as balls in an urn, the problem becomes a mathematical journey from rules to outcomes. Pólya urns provide a theoretical long-term distribution of how bees spread between the two sites (it is beta distributed). In turn, this distribution tells us something about the responsiveness of honeybees to changes in food quality. An increased intensity of dancing is equivalent to adding balls more rapidly for one of the sites, increasing the probability of others visiting that site and contributing more balls to it. As a result, the bee colony can rapidly respond to changes, while still assigning the majority of its workforce to the most profitable sites. Using the urn model allows the bees to avoid becoming locked into an unprofitable food source.

The Pólya urn captures the essence of honeybee dance competition, but it ignores many of the details found in Seeley's experiments. The challenge for mathematical modellers is deciding what level of detail is useful for describing a particular system. Honeybee foraging has been described as everything from simple urn models like that above; as differential equations representing the number of bees in various states of foraging; with agent-based models which capture even more behavioural states; to full visualisations that include the flight paths of the bees. Precisely because of the richness of Seeley's and others' experimental work, modellers have been encouraged to try out all of these tools.

Despite this wide variety of modelling approaches, the Pólya urn models and simple sets of differential equations still provide the greatest insight into the food collection of honeybees.

More detailed models tend to obscure the key processes. However, there remain important questions about honeybee organisation that require more detailed models. These questions revolve around how the colony functions as a whole. Seeley's book proposes many different regulatory feedbacks that should balance how many individuals engage in each task. But no one has yet demonstrated how these feedbacks allow the colony as a whole to function effectively.

The problem here, which appears to be a problem throughout systems biology and computational social sciences, is that we still do not have a reliable and general way of building detailed models. In most cases we cannot even agree on a formalism to work with. Agent-based simulations of a system are difficult to reproduce by other researchers, and as a result it is often impossible to build on the results of others. In Sumpter, Blanchard and Broomhead [7] we used a formal modelling tool known as process algebras to provide uniquely defined models of social insect interactions. These worked fine until we needed to incorporate space and continuous variables. Then the models became cumbersome. As I will show in the next section, once we put space into models many important emergent effects appear that are not easily captured by formal modelling tools. The challenge of finding a reliable framework for large-scale systems modelling remains largely unsolved.

In the face of this challenge it is tempting for mathematicians to focus instead on more detailed analyses of the models that we can get to grips with analytically. Sometimes we end up studying ever more obtuse aspects of, for example, variations of the Pólya urn model. Worse still, as new potential application areas arise models are reinvented in various forms, spurring on a whole new wave of analytical developments. This is a trap I think all applied mathematicians, including myself, fall into now and then. And it is exactly at this point that we should try to return to the application at hand. We should ask ourselves questions like:

How can we find out something about honeybees that would surprise and excite Tom Seeley?

If we find out things that catch the interest of our non-mathematician collaborators then we make progress. If we have nothing to say to them, or they do not understand what we are talking about, then we have failed as applied mathematicians.

3 Following the swarm

Some of the most interesting emergent collective behaviours of animals are best understood by thinking in terms of Hollywood animations. In the 1980s, Craig Reynolds proposed his 'boids' model of bird flocks and fish schools. In his model, each boid interacts with its local neighbours through local attraction, repulsion and alignment rules. Through these interactions, flocks of boids can create collective motion reminiscent of natural flocks and swarms. This technique has allowed film-makers and computer game programmers to create large realistic-looking flocks of everything from birds and humans to aliens and monsters.

Applied mathematicians, physicists and biologists have taken the Reynolds model and refined it to ask questions about real animal swarms, not least honeybees. In the spring, the queen leaves its hive together with around 10,000 worker bees and starts the process of looking for a new home. This involves sending out scouts, which find a new cavity to move to. Once a suitable cavity is found, the difficulty faced by the colony is that only a few

hundred of the scout bees have visited the new potential home. How can 2% or 3% of the bees guide the swarm?

In one modelling study, Madeleine Beekman and her colleagues used a boids-like model to test the stalker bee hypothesis [8]. The hypothesis is that the scouts fly rapidly through the swarm, the other bees then follow fast moving individuals and are thus guided in the direction of the streaking minority. In simulations this worked well, with the swarm following the fast moving leaders. However, in a more general study of leadership in collective motion, Iain Couzin and his colleagues showed that stalker bees were not a requirement for leading large swarms [9]. They proposed, what Seeley later called the subtle guide hypothesis: that small numbers of individuals can lead large groups without any of the followers knowing which bees are leaders and which are followers.

The stalker vs. subtle hypotheses set up a challenge for experimental biologists, working together with engineers, to find out what was going on inside the flying swarm. By filming a swarm from below, Schulz et al. [10] were able to show peaks in the speed distribution as the bees passed over. It appears that a small number of bees at the top of the swarm are streaking, but it still remains to be conclusively demonstrated that these are the scout bees.

The search for stalker bees is just one example of how boid-like models have inspired interdisciplinary research. As more biological species are studied, it has become increasingly clear that there is no one size fits all model for collective motion. The details of how individuals align, are attracted to each other, interact using chemicals and their environment are different and important in each case [11]. Moreover, even the simplest of model assumptions produce complex, emergent behaviour. It seems to be precisely where we lose mathematical tractability of our models, that they produce the most interesting behaviour.

4 Perspective

I have used honeybees to give an overall impression of the ideas and problems that face mathematics applied to complex systems. Deriving rules from first principles, doing modelling at the correct level of detail and understanding emergent phenomena remain the key challenges in applied mathematics. These are the problems I loved to discuss with Dave Broomhead in Manchester pubs as a PhD student, and many of the ideas presented here came from these discussions. I am not sure Dave would have fully agreed with everything I have written. He had more time for mathematical diversions than I do. But we both agreed that it was the duty of an applied mathematician to keep on talking to biologists and other scientists and let the mathematics be driven by the applications.

The single most useful piece of advice I got from Dave is something that should be taught to every applied mathematician. It is this:

You should never be afraid to ask a stupid question.

Dave reasoned that applied mathematicians always had a way to save face. If you ask a stupid question to a biologist, it is because you are a mathematician. If you ask a stupid question to a pure mathematician, it is because you are a mathematical biologist. Being an applied mathematician is your excuse for behaving like an idiot. Dave was an expert at asking idiotic questions. After seminars, he would say:



© Tuthelens | Dreamstime.com

I'm being a bit thick, but there is something I don't understand about your assumption here. . . .

Five minutes later the speaker would be admitting his or her own confusion about the assumption. And at the same time he or she would have found, in Dave, a new ally to work with in solving the problem and tightening up the thinking.

Asking idiotic questions is the starting point for a great deal of good quality research. The examples I have given here – of doubting an intuitively obvious equation like Hamilton's law, trying to build a formalised model of honeybee organisation, or deciding to turn Hollywood computer animations into serious biological models – all start with an idea or a question that might seem idiotic. Unfortunately, there is no known method for deciding in advance if an idea will ultimately turn out to be genius or just as idiotic as it sounded in the first place. Even for the three examples I have given, I still do not know for sure which approaches are ultimately genius and which are ultimately idiocy. But, as a researcher, whether you are right or wrong in the end, you should never be afraid of putting in the work to find out the answer. You should never be scared of looking like an idiot.

Notes

1. Although I have never cycled from Manchester to London, I have cycled over other equally long distances. I have found that this gives both great enjoyment and a whole new perspective on the journey. I want to point this out, because there are always good and beneficial reasons for long diversions, both in travel and in mathematics.

REFERENCES

- 1 Sumpter, D.J.T. and Broomhead, D.S. (2001) Relating individual behaviour to population dynamics, *Proc. Roy. Soc. London B: Biolog. Sci.*, vol. 268, no. 1470, pp. 925–932.
- 2 Sumpter, D.J.T. (2010) *Collective Animal Behavior*, Princeton University Press.
- 3 Nowak, M.A., Tarnita, C.E. and Wilson, E.O. (2010) The evolution of eusociality, *Nature*, vol. 466, no. 7310, pp. 1057–1062.
- 4 Lehmann, L. and Keller, L. (2006) The evolution of cooperation and altruism – a general framework and a classification of models, *J. Evol. Biol.*, vol. 19, no. 5, pp. 1365–1376.
- 5 Abbot, P., Abe, J., Alcock, J., Alizon, S., Alpedrinha, J.A., et al. (2011) Inclusive fitness theory and eusociality, *Nature*, vol. 471, no. 7339, pp. E1–E4.
- 6 Seeley, T.D. (1995) *The Wisdom of the Hive: The Social Physiology of Honeybee Colonies*, Harvard University Press.
- 7 Sumpter, D.J.T., Blanchard, G.B. and Broomhead, D.S. (2001) Ants and agents: a process algebra approach to modelling ant colony behaviour, *Bull. Math. Biol.*, vol. 63, no. 5, pp. 951–980.
- 8 Janson, S., Middendorff, M. and Beekman, M. (2005) Honeybee swarms: how do scouts guide a swarm of uninformed bees?, *Animal Behaviour*, vol. 70, no. 2, pp. 349–358.
- 9 Couzin, I.D., Krause, J., Franks, N.R. and Levin, S.A. (2005) Effective leadership and decision-making in animal groups on the move, *Nature*, vol. 433, no. 7025, pp. 513–516.
- 10 Schultz, K.M., Passino, K.M. and Seeley, T.D. (2008) The mechanism of flight guidance in honeybee swarms: subtle guides or streaker bees?, *J. Exp. Biol.*, vol. 211, no. 20, pp. 3287–3295.
- 11 Sumpter, D.J.T. and Broomhead, D.S. (2000) Shape and dynamics of thermoregulating honeybee clusters, *J. Theor. Biol.*, vol. 204, no. 1, pp. 1–14.