What computers won't tell you about ecological and evolutionary dynamics

December 9 2015

In a paper published this week in the *Proceedings of the National Academy of Sciences*, Rasmus Ibsen-Jensen, postdoc at the Institute of Science and Technology Austria (IST Austria) together with IST Austria Professor Krishnendu Chatterjee and Professor Martin A. Nowak from Harvard University discovered surprising connections between computer science and biology, two disciplines that study how information proliferates in time and space.

The authors specifically took a look at complexity theory which is part of theoretical [computer science](#) as well as mathematics. It classifies algorithms that can solve certain categories of computational problems according to their inherent difficulty.

They applied these well-defined classes of complexity to some fundamental questions in biology, namely to the ecological and evolutionary dynamics within structured populations. These research questions investigate how population structures affect the outcome of the evolutionary process.

Consider for example the problem of finding the probability that a genetic mutation establishes itself in a resident population, or an invasive species occupies an ecological niche. Though these problems are well studied, an understanding of the computational complexity of even such simple problems was missing.

The investigators found the rather unexpected proof that these

fundamental questions in ecology and evolution can be precisely characterized by specific classes of complexity theory, as though these evolutionary processes would mimic aspects of computation. By defining the exact complexity class, they formally proved that these questions cannot be solved with a simple formula.

However, the authors also demonstrated that two classic problems are indeed efficiently solvable: One is the molecular clock—the rate at which neutral mutations accumulate over time—and the other is the exact fixation probability for a genetic variation to take over in the case of a well-mixed population structure.

So how could the researchers tell for certain that some questions are solvable and that an algorithm must exist? And how is it possible to claim that for other specific questions a computational solution is not possible?

The authors used the established methods of computational [complexity theory](#) and applied them to the defined evolutionary scenarios in evolutionary game theory and evolutionary graph theory. As a result, they were able to derive the exact complexity class for each of these research questions. The specific complexity class in turn can tell us if an efficient algorithm exists. How so?

For instance, given a gigantic jigsaw puzzle and a guidance telling us where each piece should go, it is easy to assemble the puzzle and check that it indeed gives the picture on the box. Complexity theory calls this type of problems P as the problem can be solved in polynomial time. Among other things, P contains all problems that can be solved with a simple formula. However, many problems cannot be solved in polynomial time as the necessary computational resource would increase exponentially with the growing size of the input. Thus, for such problems, no simple formula exists.

Another important complexity class in the context of this study is nondeterministic polynomial time (NP) which categorizes the problems to which solutions can be verified in polynomial time: If we imagine the giant jigsaw puzzle without the guidance, then it might be hard to decide where some of the pieces should go. However, as mentioned above, once a solution (=guidance) is found it is easy to check that it is indeed a solution.

Incidentally, the question whether all problems in NP belong to P or not is one of the famous six yet unsolved problems of mathematics, for which a Millenium Prize of 1 million dollars was awarded in 2000 by the Clay Mathematics Institute. It is the widely believed albeit yet unproven that P does not equal NP, and hence, the hardest problems in NP cannot be solved with a simple formula.

The findings of the investigators are a first step to establish a connection between the two disciplines of computer science and biology, and they also suggest that research on certain questions in ecological and evolutionary dynamics need to focus on special aspects that can be solved with a simple formula.