Abstract:

The concept of fitness is key in both evolutionary biology and evolutionary computation. Yet the meaning of the term differs in important ways between these fields. These differences include simple yet easily overlooked differences in terminology (eg, a more fit biological organism will have a higher fitness score, while many evolutionary computation problems are set up such that a more fit solution has a lower fitness score), mechanistic differences (eg evolutionary computation problems typically have explicit fitness functions, while biological systems rarely do), and large conceptual differences (eg, many evolutionary computation problems have a mathematically provable optimal solution and thus a maximum fitness, while current evidence in biology suggests in even simple biological systems over vast time periods fitness is unbounded). In this paper, we will summarize how both evolutionary biology and evolutionary computation define and conceptualize fitness, synthesize the commonalities in both fields' approaches, and highly important differences to keep in mind when crossing disciplinary boundaries so as to not misunderstand results.

Fitness in evolutionary biology:

Fitness is the key concept in understanding evolutionary dynamics, as fitness is the quantity that natural selection optimizes over evolutionary time. Topics that we will focus on in our review of the relevant literature in evolutionary biology include:

- *Absolute vs relative fitness*. That is, whether fitness is conceptualized as a number for an organism irrespective of others (absolute), or if it is measured in proportion to others in the population (relative).

- *Use of fitness correlates*. Ideally, we would like to know how many copies of a given genetic variant are making it into the future population gene pool. In practice, in many biological systems this is difficult to measure, so various correlates of fitness (seeds set, eggs laid, young hatched, etc) are used. We will review how different correlates of fitness result in different understandings of evolutionary dynamics.

- *w vs r.* Population genetics equations are often based on w, the ratio of the Malthusian parameters of two competing subpopulations. Many experimental biologists report fitness in terms of r, the selection rate. These two values are not interchangeable, and different conclusions can be reached depending on which version of fitness is being quantified.

Fitness in evolutionary computation:

Evolutionary computation also relies heavily on fitness in its application of evolutionary principles to computational problems. Unlike in evolutionary biology, fitness can often be calculated without measurement noise, and separately for each individual in the population. This wealth of data can be of tremendous value in understanding the problem domain and interpreting differences among solutions. Yet important differences in various interpretations of fitness remain. From the literature on evolutionary computation, topics we will focus on include:

- *Explicit vs implicit fitness.* Computational systems employing evolutionary strategies can be formulated in different ways. One important difference is whether the fitness score of a solution is determined explicitly (for a trivial example, number of 1s in a binary string) or implicitly (for example, by instituting multiple rules about how to govern replication, and determining how many offspring are produced by any given solution). We will review how the different ways of assigning fitness change how fitness can be interpreted.

- *Bounded vs unbounded fitness changes.* For some problems in EC, mathematically provably optimal solutions exist. As such, once a population converges on this solution, no further improvement in the population is possible. This is in contrast with fitness changes in biological systems, which exhibit unbounded improvements in fitness over evolutionary time. We will explore the degree to which this difference is driven by time frames (for example, that environmental changes in biological systems may happen before an optimal solution can be reached), and the degree to which this is driven complexity (for example, that negative frequency dependence in fitness is pervasive in biological systems both within and between species, while interaction between solutions to different problems are rare in EC).

- *Maximization vs minimization.* Evolutionary biologists nearly universally refer to the most fit individuals as having the highest fitness, and natural selection maximizing fitness. Many EC researchers, on the other hand, quantify fitness such that the most fit solutions have the lowest fitness, and thus speak of minimizing fitness. Although a seemingly trivial difference, this different way of conceptualizing the best solution results in words used commonly in the different subfields – hills, valleys, etc -- meaning different things to different audiences. We will highlight these differences, and offer suggestions for how to make one's assumptions more explicit, to minimize confusion across the disciplines.

Frequent but often unrecognized differences between biological and computational models

There are many commonalities between evolutionary models in biological and computational systems. Some of the differences (for example, cellular machinery) are explicitly recognized. Some of the differences, however, may go unrecognized. We will call out a few of the differences that may be underrecognized in their importance, such as:

- *Continuous vs discrete generations.* Most biological systems have overlapping, continuous generations. Most EC instantiations have synchronized, discrete generations. Confusingly, most population genetics as it is taught relies on discrete generations, potentially leading researchers in the two fields to not realize that they are making different assumptions that can lead to important differences in results. We will examine the effects of different fitness values in the two different scenarios.

- *Distribution of fitness effects.* Evolutionary dynamics are influence by a variety of fitness values, including the mean fitness in the population, the best fitness in the population, and the distribution of fitness values within the population. However, biological and computational models often vary widely in their distribution of fitness effects. We will discuss how this impacts our understanding of fitness dynamics in the two fields.

- *Mechanisms of diversity preservation.* Preserving diversity is important in EC, as solutions quickly homogenize without it. Yet this preservation of diversity is often explicit in nature, while diversity maintenance in biological systems is implicit, and driven by certain phenotypes being more advantageous when rare. We will discuss the extent to which these differences are important in our understanding of evolutionary dynamics.