What is “epistasis”?

H. Van Hove A. Verschoren

Abstract

In this note, we compare several descriptions of “epistasis”, which have been used in the literature and apply some techniques originating from linear algebra, to present a formal approach, which encompasses all of the previous ones.

0 Introduction

Although several variants of the genetic algorithm (GA) have been introduced in the past, basically “the” GA always acts on a set Ω of binary strings, of length l say, and aims to maximize a real valued fitness function $f : \Omega \to \mathbb{R}$, through the use of a suitable set of genetic operators, acting on successive populations sampled within Ω.

Most models use crossover and mutation as their fundamental operators. Crossover basically constructs new strings by combining pieces of given ones, whereas mutation randomly changes bits of a given string (with a very low probability). The GA starts from a randomly chosen initial population $P(0) \subseteq \Omega$ (a multiset, in general). It then picks couples of strings within $P(0)$, with a probability of being selected proportional to their fitness and applies crossover and mutation to this pair, thus producing a new population $P(1)$. This process is then iterated until some initially given criterion is satisfied.

The GA is a very robust and widely applicable algorithm, due to the fact that it does not directly act on the crude data of a problem, but on the encoding of these data into binary strings. As a side effect, it should be clear that the efficiency of the GA is highly dependent on the chosen encoding scheme and, in particular, on the interdependency of different bits in the strings resulting from it.

Received by the editors September 1996.
Communicated by A. Hoogewijs.
1991 *Mathematics Subject Classification :* 68Q27, 68T05,15A09.
Key words and phrases : epistasis, generalized inverse.