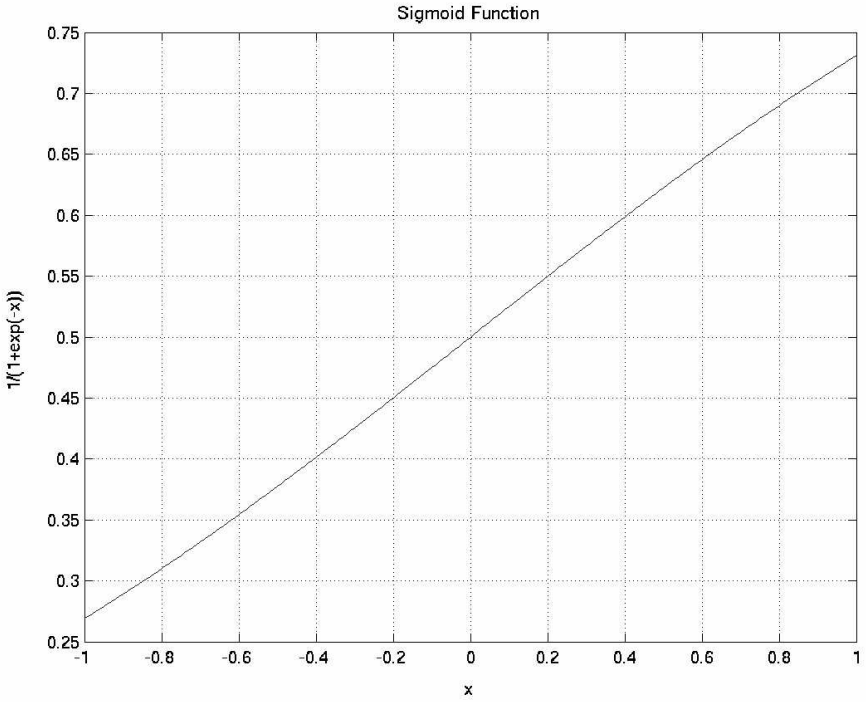
Zoomed view



2002 Exam 2 Problem 3: Genetic Algorithms (16 points)

Professor C. Ross Ovorr has become sick and tired of creating new final exam questions, so he has decided to try to use a genetic algorithm to generate exam questions from previous exam questions.

First, he decides on a question fitness function, f. Then, he figures out how he can encode questions in "genes" on which his program can perform crossover and mutation operations.

To refine his algorithm, he decides to devote a few weeks of recitation classes to trying out the questions generated by four algorithm variations. Fortunately, the students are half asleep, so they do not remember anything from week to week, so he does not have to worry about memory or any kind of learning mucking up his experiments.

Section R1 receives questions that are generated by crossing the best two questions given to R1 the week before. There is no mutation. The crossover rate is such that there is usually one crossover when two genes are combined.

Section R2 receives questions that are generated by mutating the best two questions given to R2 the week before. There is no crossover. The mutation rate is such that most of the parts of the mutated gene are different from its not-mutated source.

Section R3 receives questions that are generated by crossing and mutating selected questions given to R3 the previous week.

Candidate questions are selected for crossover and mutation using the following survival probabilities:

$P_i = f_i / \Sigma_i f_i$

where f_i is the fitness of the question.

Note that all candidates are considered for each selection, so the same candidate can be selected multiple times.

Section R4 receives questions that are generated in the same way as section R3, except that candidate questions are selected using the following survival probabilities:

$P_i = [f_i^2 + d_i^2]^{1/2} / \Sigma_i [f_i^2 + d_i^2]^{1/2}$

where f_i is the fitness of question *i* and d_i, the diversity of question *i*, is given by:

$d_i = 0$ if question *i* is the most fit question

$d_i = f_h - f_i$ otherwise, where h identifies the question with the ***next higher*** fitness

Thus, for four questions with fitnesses of 3, 0, 3, and 4, d_i would be 1, 3, 1, and 0. (d_1 is 1 because its fitness is 3 and the fitness of the next higher-fitness question is 4, d_2 is 3 because its fitness is 0 and the fitness of the next higher-fitness question is 3, d_3 is 1 by the same argument as for d_1, d_4 is 0 because it is most fit.)

Part A: Performance (4 points)

Which of the four section's methods is **least** likely to produce steady improvement? Circle the best answer:

R1 **R2** R3 R4

Part B: Diversity (12 points)

B.1 (3 points)

In week 1, the fitness scores of the questions given to R4 were: 3, 0, 3, and 4. Calculate for each question the probability of being a parent for a question in R4 in week 2.

P_1 P_3 = (/ (sqrt 10) (+ 7 (* 2 (sqrt 10)))) = 0.237

P_2 = (/ 3) (+ 7 (* 2 (sqrt 10)))) = 0.225

P_4 = (/ 4) (+ 7 (* 2 (sqrt 10)))) = 0.300

B.2 (3 points) Consider the ratio of the probability of the most-likely-to-survive question in a recitation section to the probability of the least likely to survive. Circle the best statement:

- **The ratio will likely be greater in R4 than R3. Compare numbers in B.1 with infinite ratio obtained with R3.**
- The ratio will likely be greater in R3 than R4.
- They will be the same.

B.3 (3 points) Assume that genetic diversity of the questions is the same in both R3 and R4 in week 1. From week 2 onwards, is the genetic diversity in R3 likely to be less or more than in R4? Circle the best answer:

- R3 is more diverse.
- **R4 is more diverse. Emphasis on fitness, without diversity, often drives GA up closest local maximum.**
- Both are equally likely to be more diverse.

B.4 (3 points) After many weeks, is the most fit question in the general population of questions likely to be higher in R3 or R4? Circle the best answer:

- Question in R3 is most fit.
- **Question in R4 is most fit. Use of diversity can help GA get off of local maximum.**
- R3 and R4 are equally likely to produce the most fit question.

Problem 4: Miscellaneous (14 points)

Select the single **BEST** answer for the following questions. No points can be awarded for multiple selections.

1. Alan Turing evidently believed that the computations performed by the human brain are:
   1. **Computable by a universal Turing machine.**
   2. Computable by a universal Turing machine with 10 fundamental exceptions.
   3. Beyond the reach of all conceivable Turing machines.
   4. Fundamentally different from the computations performed by other primates.
   5. None of the above.

2. Alan Turing argued that:
   1. Computers can be intelligent because they will eventually pass the Turing test.
   2. **Computers can be intelligent, by refuting contrary arguments.**
   3. Computers cannot be truly intelligent because they cannot have free will.
   4. Computers cannot be truly intelligent because they do not have bodies.
   5. None of the above.

3. According to Minsky, frames are suited for:
   1. Representing situations.
   2. Representing visual scenes.
   3. **Dealing with multiple perspectives.**
   4. **All of the above.**
   5. None of the above.