



Department of Computer Science
University of Pretoria

Artificial Intelligence
COS710

Assignment 1: Applied Particle Swarm Optimization

Due date: 3 April 2018, at 23:59

Do the assignment described below. Note that you have to submit a pdf document, containing a technical report wherein you describe what you have done, and present and discuss your results. Remember that you can not report results for only one run of an algorithm on a problem, but you have to report results as averages and standard deviations over a sufficient number of independent runs of the algorithms. I usually use 50 independent runs.

For this assignment, you will use a PSO algorithm to solve one of the problems given below. You may choose any of the choices given below.

Sequence Alignment

Given a number of sequences of characters, the objective is to find the best possible vertical alignment of these sequences. Alignment is achieved by inserting leading spaces before characters in the sequences so that as many as possible characters are vertically aligned. While this sounds simple, the problem has two objectives:

- maximize the number of vertically aligned characters, but
- minimize the number of leading spaces inserted.

As an example, consider the problem of aligning the following sequences:

```
a b c d e f
b b d h g
c a b f
```

Note that the sequences can be of different length. A possible solution to this alignment problem is

```
- a b - c d e f - -
- - b b- d - - h g
c a b - - - f - -
```

You have to develop a multi-objective PSO algorithm to solve such alignment problems. You have to very carefully think about the representation scheme. That is, how a particle will represent a solution. Then, also think carefully about the formulation of the two objectives.

In your report, you have to provide a clear description with illustrations of your representation scheme, and provide motivations for why this is a good scheme. Then also provide clear discussions with motivations for your objective functions. You also have to describe the multi-objective PSO that you have used to solve the problem. Consider appropriate performance measures and report the performance of the algorithm using these measures. You have to test your PSO algorithm for a number of different sequence alignment problems. Be careful to develop sequences that vary in difficulty.

Clustering Static Data

For this assignment, you have to develop a multi-objective PSO to cluster data, assuming static data. The data clustering problem is a multi-objective problem, with the following two objectives:

- minimize intra-cluster distances, and
- maximize inter-cluster distances.

Think carefully about the representation (how a particle will represent the solution) and the formulation of the objective functions. You have to describe and motivate in your report your choice of representation and objective functions. Start by assuming that the number of clusters is *a priori* specified (i.e known). Then, think of an approach in which the PSO can also dynamically determine the optimal number of clusters.

Evaluate the performance of your multi-objective PSO clustering algorithm using appropriate performance measures for multi-objective optimization problems. Make sure that you evaluate the performance over a number of data sets of varying complexity.

Solving Systems of Equations

For this assignment, you will develop a PSO approach to solve systems of equations. Note that the equations can be linear and/or non-linear, and you should allow for systems of equations that have more variables than equations.

Think carefully how you will define the objective function for this problem, and how solutions will be represented. Describe and motivate your implementation choices.

When you choose a PSO to solve a system of equations, note that there may be more than one solutions.

You have to evaluate the performance of your PSO algorithm on a number of different systems of equations, of varying complexity. Think carefully of the performance measures that you will use to report the performance.

Format of your Technical Report

1. Introduction:

The introduction sets the stage for the remainder of your report. You usually have very general statements here. The introduction prepares the reader for what to expect from reading your report. In general, the introduction should either contain or be a summary of your ENTIRE report.

2. Background:

A very high level discussion on the problem domain and the algorithms and/or approaches that you have used. Do not be too specific on the algorithms and approaches. This section is typically where the "base cases" of concepts that appear throughout the remainder of your report are discussed. It is also an ideal place to refer a reader to other sources containing relevant information on the topic but which is outside the scope of your assignment. It is the perfect place for pseudo code. Remember to discuss very generally. After reading this section the marker should be able to determine whether or not you know what you're talking about.

3. Implementation:

In this section you discuss how you approached, implemented and solved your assignment choice. Mention, for example, the values set for the algorithm's control parameters, how many simulations you have run and what the characteristics for candidate solutions to your problems are. After reading this section (in addition to the background) the reader should be able to duplicate your experiments to obtain similar results to those obtained by you. This is also the section where your discussion specializes on the concepts mentioned in the background section. Be very specific in your discussions in this section.

4. Research Results:

This is the section where you report your results obtained from running the experiments as discussed in the implementation section. You have to give, at least, averages and standard deviations for the experiments/simulations. Thoroughly discuss the results that you've obtained and reason about why you obtained the results that you have. Answer questions like "are these results to be expected?" and "why these results occurred?" and "would different circumstances lead to different results?".

5. Conclusion(s):

Very general conclusions about the assignment that you have done. This section "answers" the questions and issues that you've raised and investigated. This section is, in general, a summary of what you have done, what the results were and finally what you concluded from these results. This is the final section in your document so be sure that all the issues raised up until now are answered here. This is also the perfect section to discuss what you have learnt in doing this assignment.