

# ABSR: An algorithm based on similarity rules assisted by bio-inspired algorithms for teaching context

## Gilberto Nerino de Souza Junior

Electrical Engineering Graduate Program - Federal University of Pará Rua Augusto Corrêa, 1 - Guamá, Belém – PA, Brazil, Postal code: 66075-110 gilbertonerinojr@gmail.com

## Márcia Fontes Pinheiro

Electrical Engineering Graduate Program - Federal University of Pará Rua Augusto Corrêa, 1 - Guamá, Belém – PA, Brazil, Postal code: 66075-110 eng.marciafontes@gmail.com

#### **Dionne Cavalcante Monteiro**

Institute of Exact and Natural Sciences - Federal University of Pará Rua Augusto Corrêa, 1 - Guamá, Belém – PA, Brazil, Postal code: 66075-110 dionne@ufpa.br

## Ádamo Lima de Santana

Electrical Engineering Graduate Program - Federal University of Pará Rua Augusto Corrêa, 1 - Guamá, Belém – PA, Brazil, Postal code: 66075-110 adamo@ufpa.br

# ABSTRACT

In the educational context, measure the similarity between words is an important component to analyze the perception and difficulty of tasks. However, the assignment of similarity can be subjective, which can cause inconsistencies of a expert to another. Furthermore, rarely uses specific rules for refinement of the rate of similarity. In another way, the search for similarity by brute force can generate a high computational cost.

This paper proposes a Algorithm Based on Similarity Rules (ABSR) capable of generating the similarity between words within an educational context. The similarity can be calculated by each rule that has an associated weight as Algorithm 1,

Algorithm 1: ABSR
<i>Input</i> : Word1, Word2, <b>Rules</b> , <i>W</i> , <i>n</i>
Output : S
Begin
For $i \leftarrow 1$ , to $n$ , do,
If $Rules[i]$ .(Word1, Word2) = True, then
$W_{r_i} \leftarrow \text{Rules}[i].\text{CalculateRuleValue}(\text{Word1}, \text{Word2}, W_i)$
EndIf
EndFor
n
$S \leftarrow \sum W_i V_{r_i}$
$\underset{i=1}{\overset{\leftarrow}{\leftarrow}}$
Fnd

where: S denotes a similarity;  $W_i$  denotes a weight for the *i*th rule;  $V_{r_i}$  denotes a value of the *i*th rule; n is the number of rules.

A experiment was conducted where the opinion of similarity from 9 experts of the teaching program called, Learning to Read and Write in Small Steps, were collected through questionnaires. A total of 25 words, were used to the collection of 1400 similarity values. The following rules were also defined: What is the Levensthein's similarity?; What is the difference between the size of the words?; First letter of the first word is the equal the first syllable of the second word?; Last syllable of the first word is equal the last syllable of the second word?; Last letter of the second word?



To perform the search of the weights of the rules was used the following bio-inspired algorithms: GA (Genetic Algorithm), PSO (Particle Swarm Optimization), EPSO (Evolutionary Particle Swarm Optimization) and DEEPSO (Differential Evolutionary Particle Swarm Optimization). A fitness function was used to find a set of 6 weights that generate similarities within the error range and close to the average of the similarities provided by the opinions of experts.

Results reveal that the GA is superior in terms of population diversity and convergence of fitness, generating the following weights to the rules cited: 0.49, 0.174, 0.064, 0.0772, 0.1702, and 0.0246. The opinions of experts were also used to calculate the RMSE (Root Mean Square Error) regarding the following similarity metrics: ABSR (0.083), Levenshtein (0.096), Dice (0.207), Jaccard (0.224) and Jaro (0.303). Therefore, the ABSR approached the opinions of experts making it feasible to identify similarities and formulation of new teaching assignments.

**KEYWORDS.** Similarities between words, Bio-inspired algorithms. Topics EDU - OP in Education, OA - Other applications in OP