Managing Genetic Algorithm Parameters to Improve SegGen - A Thematic Segmentation Algorithm

Submitted by Bernard Levrat on Tue, 07/14/2015 - 19:27

Titre	Managing Genetic Algorithm Parameters to Improve SegGen - A Thematic Segmentation Algorithm
Type de publication	Communication
Type	Communication avec actes dans un congrès
Année	2013
Langue	Anglais
Date du colloque	26-30 Aug. 2013
Titre du colloque	24th International Workshop on Database and Expert Systems Applications (DEXA)
Titre des actes ou de la revue	Database and Expert Systems Applications
Pagination	58-62
Auteur	Saygili, Sirin [1], Acarman, Tankut [2], Amghar, Tassadit [3], Levrat, Bernard [4]
Pays	Etats-Unis
Editeur	IEEE
Ville	Los Alamitos
ISBN	978-0-7695-5070-1
Mots-clés	genetic algorithm [5], hematic segmentation [6], multi-objective optimization problem [7]
Résumé en anglais	SegGen [1] is a linear thematic segmentation algorithm grounded on a variant of the Strength Pareto Evolutionary Algorithm [2] and aims at optimizing the two criteria of the Salton's [3] definition of segments: a segment is a part of text whose internal cohesion and dissimilarity with its adjacent segments are maximal. This paper describes improvements that have been implemented in the approach taken by SegGen by tuning the genetic algorithm parameters according with the evolution of the quality of the generated populations. Two kinds of reasons originate the tuning of the parameters and have been implemented here. First as it could be measured by the values of global criteria of the population quality, the global quality of the generated populations increases as the process goes and it seems reasonable to set values to parameters and define new operators, which favor intensification and diminish diversification factors in the search process. Second since individuals in the populations are plausible segmentations it seems reasonable to weight sentences in the current segmentation depending on their distance to the boundaries of the segment they belong to for the calculus of similarities between sentences implied in the two criteria to be optimized. Although this tuning of the parameters of the algorithm currently rests on estimations based on experiments, first results are promising.
URL de la notice	http://okina.univ-angers.fr/publications/ua13593 [8]

Lien vers le document en ligne	http://dx.doi.org/10.1109/DEXA.2013.15 [9]

Liens

[1] http://okina.univ-angers.fr/publications?f[author]=23458
[2] http://okina.univ-angers.fr/publications?f[author]=23459
[3] http://okina.univ-angers.fr/t.amghar/publications
[4] http://okina.univ-angers.fr/bernard.levrat/publications
[5] http://okina.univ-angers.fr/publications?f[keyword]=6813
[6] http://okina.univ-angers.fr/publications?f[keyword]=19838
[7] http://okina.univ-angers.fr/publications?f[keyword]=19839
[8] http://okina.univ-angers.fr/publications/ua13593
[9] http://dx.doi.org/10.1109/DEXA.2013.15

Publié sur *Okina* (http://okina.univ-angers.fr)