Newly designed CONOP program helps tackle the stratigraphic correlation problem of late Paleogene foraminifera

Zhengbo Lu*[1] and Junxuan Fan[1,2]

[1]School of Earth Sciences and Engineering and Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing – China

[2]State Key Laboratory for Mineral Deposits Research, Nanjing University, Nanjing – China

Abstract

Constrained optimization (CONOP), which uses a simulated annealing approach to iteratively find optimized matches to integrate sequences of stratigraphic datums via random perturbations (Sadler et al., 2009), is regarded as an effective method for handling global stratigraphic correlation problems, one of the nondeterministic polynomial time problems in computational complexity theory. However, the original CONOP program did not support parallel computing or high-performance computing (HPC), which may hamper the search for a global optimal sequence in the CONOP process. Therefore, Fan et al. (2020) developed a new program - CONOP.SAGA by combining simulated annealing and a genetic algorithm to overcome these limitations. Although this program provides the HPC function and works well with the China Palaeozoic dataset (Fan et al., 2020), it still followed the same procedure of CONOP to generate the initializing sequence and mutation, i.e., randomly placing the first and last appearance datums to the composite sequence and stochastically re-arranging these datums at the mutation stage, leaving opportunities for further improvement. Here, we employed a new method of generic population-based metaheuristic optimization algorithm, called CONOP.EA, which combines the original CONOP program with an evolutionary algorithm (Back, 1996). It is inspired by Darwin's theory of evolution to iteratively improve a set (i.e., population) of datum sequences via procedures of resembling or mimicking, recombination, mutation, and natural selection to generate a global sequence of stratigraphic datums of late Paleogene foraminifera. We construct a 28-Myr-long species richness history of foraminifera with an average temporal resolution of ~26,000 years. This composite sequence uncovered the coupled foraminiferal diversity and environmental change pattern across the Eocene-Oligocene transition, a turning point in Earth's climate from "warmhouse" into "icehouse". The computation of this dataset was ~17 times faster with CONOP.EA compared to CONOP.SAGA, and produced a better result with 520 less in penalty. Moreover, CONOP.EA has a new ability to evolve from previous results, and the evolved offspring sequences are proven to be better than their parent sequences by an average of 112 fewer penalties. This means that CONOP.EA is able to dig further to find a more globally optimized sequence than ever before.

Sadler, P. M., Cooper, R. A., and Melchin, M. 2009. High-resolution, early Paleozoic (Ordovician-Silurian) time scales. Geological Society of America Bulletin, 121(5-6): 887-906.

*Speaker

Fan, J. X., Shen, S. Z., Erwin, D. H., Sadler, P. M., MacLeod, N., Cheng, Q. M., Hou, X. D., Yang, J., Wang, X. D., Wang, Y., Zhang, H., Chen, X., Li, G. X., Zhang, Y. C., Shi, Y. K., Yuan, D. X., Chen, Q., Zhang, L. N., Li, C., and Zhao, Y. Y. 2020. A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity. Science, 367(6475): 272-277.

Back, T. 1996. Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. Oxford university press.

Keywords: CONOP, Evolutionary algorithm, Quantitative stratigraphic method, foraminifera