Blasco Latorre, D., Iglesias Soria, A., Echeverría Ochoa, J., Pérez Pérez, F., Cetina Englada, C.: Introducing Phylogenetics in Search-based Software Engineering: Phylogenetics-aware SBSE. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/32

Introducing Phylogenetics in Search-based Software Engineering: Phylogenetics-aware SBSE

Daniel Blasco[1], Antonio Iglesias[1], Jorge Echeverría[1], Francisca Pérez[2], and Carlos Cetina[2]
dblasco@usj.es, aiglesias@usj.es, jecheverria@usj.es, franciscaperez@upv.es, cetina@upv.es

[1] SVIT Research Group, Universidad San Jorge, Zaragoza, Spain
[2] Centro de Investigación en Métodos de Producción de Software, Universitat Politècnica de València, Valencia, Spain

Keywords: Phylogenetics, Search-based Software Engineering, Evolutionary Computation, Game Software Engineering, Procedural Content Generation, Software Maintenance, Feature Location, Model-Driven Engineering
Published in: ACM Transactions on Software Engineering and Methodology (TOSEM), Vol. 34, Issue 7, No. 193, pp. 1-38, 2025
Impact Factor: JCR 6.2 - Q1 - Position: 9/129 - Area: Computer Science, Software Engineering - JCR Index 2024
DOI: https://doi.org/10.1145/3715002

Abstract. Phylogenetics studies the relationships, in terms of biological history and kinship, of a set of taxa (e.g., species). We argue that in Search-based Software Engineering (SBSE), the individuals of an evolutionary computation-driven population could be considered as taxa for which the leverage of Phylogenetic Inference might be beneficial. In this work, we present our Phylogenetics-aware SBSE approach. Our approach introduces a novel Phylogenetic Operation to promote results which are sufficiently aligned (in terms of lineage) with a certain reference given by the domain expert. Our approach is evaluated in two heterogeneous industrial case studies: Procedural Content Generation from Game Software Engineering, and Feature Location from Software Maintenance. The results are analyzed using quality-of-the-solution and acceptance-by-developers measurements. We performed a statistical analysis to determine whether the impact on the results is significant compared to baselines that do not leverage Phylogenetics. The results show that our approach significantly outperforms two baselines in both case studies. Furthermore, two focus groups confirmed the acceptance of our approach and stressed that solution acceptance may make the difference in industrial environments. Our work has the potential to motivate a new breed of research work on Phylogenetics awareness to produce better results in Software Engineering.