An Integrative Approach for Measuring Semantic Similarities using Gene Ontology

Jiajie Peng 1,2, Hongxiang Li 1, Qinghua Jiang 3, Yadong Wang 1 and Jin Chen 2,4

1 School of Computer Science and Technology, Harbin Institute of Technology, China

2 MSU-DOE Plant Research Laboratory, Michigan State University, USA

3 School of Life Science and Technology, Harbin Institute of Technology, China

4 Department of Computer Science and Engineering, Michigan State University, USA

Abstract

Background

Gene Ontology (GO) provides rich information and a convenient way to study gene functional similarity, which has been successfully used in various applications. However, the existing GO based similarity measurements have limited functions for only a subset of GO information is considered in each measure. An appropriate integration of the existing measures to take into account more information in GO is demanding.

Results

We propose a novel integrative measure called InteGO2 to automatically select appropriate seed measures and then to integrate them using a metaheuristic search method. The experiment results show that InteGO2 signi_cantly improves the performance of gene similarity in human, Arabidopsis and yeast on both molecular function and biological process GO categories.

Conclusions

InteGO2 computes gene-to-gene similarities more accurately than tested existing measures and has high robustness. The supplementary document and software are available at http://mlg.hit.edu.cn:8082/