

EXPECTATION-MAXIMIZATION ALGORITHMS FOR FUZZY ASSIGNMENT OF GENES TO CELLULAR PATHWAYS

Liviu Popescu
Department of Computer Science, Cornell University,
Ithaca, NY, 14853, US

Golan Yona
Computer Science Department, Technion - Israel Institute of Technology

Cellular pathways are composed of multiple reactions and interactions mediated by genes. Many of these reactions are common to multiple pathways, and each reaction might be potentially mediated by multiple genes in the same genome. Existing pathway reconstruction procedures assign a gene to all pathways in which it might catalyze a reaction, leading to a many-to-many mapping of genes to pathways. However, it is unlikely that all genes that are capable of mediating a certain reaction are involved in all the pathways that contain it. Rather, it is more likely that each gene is optimized to function in specific pathway(s). Hence, existing procedures for pathway construction produce assignments that are ambiguous. Here we present a probabilistic algorithm for the assignment of genes to pathways that addresses this problem and reduces this ambiguity. Our algorithm uses expression data, database annotations and similarity data to infer the most likely assignments, and estimate the affinity of each gene with the known cellular pathways. We apply the algorithm to metabolic pathways in Yeast and compare the results to assignments that were experimentally verified.