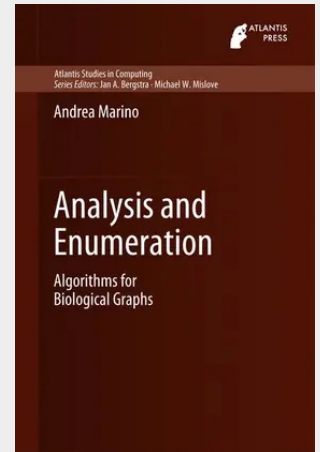


Analysis and Enumeration

Algorithms for Biological Graphs

In this work we plan to revise the main techniques for enumeration algorithms and to show four examples of enumeration algorithms that can be applied to efficiently deal with some biological problems modelled by using biological networks: enumerating central and peripheral nodes of a network, enumerating stories, enumerating paths or cycles, and enumerating bubbles. Notice that the corresponding computational problems we define are of more general interest and our results hold in the case of arbitrary graphs. Enumerating all the most and less central vertices in a network according to their eccentricity is an example of an enumeration problem whose solutions are polynomial and can be listed in polynomial time, very often in linear or almost linear time in practice. Enumerating stories, i.e. all maximal directed acyclic subgraphs of a graph G whose sources and targets belong to a predefined subset of the vertices, is on the other hand an example of an enumeration problem with an exponential number of solutions, that can be solved by using a non trivial brute-force approach. Given a metabolic network, each individual story should explain how some interesting metabolites are derived from some others through a chain of reactions, by keeping all alternative pathways between sources and targets. Enumerating cycles or paths in an undirected graph, such as a protein-protein interaction undirected network, is an example of an enumeration problem in which all the solutions can be listed through an optimal algorithm, i.e. the time required to list all the solutions is dominated by the time to read the graph plus the time required to print all of them. By extending this result to directed graphs, it would be possible to deal more efficiently with feedback loops and signed paths analysis in signed or interaction directed graphs, such as gene regulatory networks. Finally, enumerating mouths or bubbles with a source s in a directed graph, that is enumerating all the two vertex-disjoint directed paths between the source s and all the possible targets, is an example of an enumeration problem in which all the solutions can be listed through a linear delay algorithm, meaning that the delay between any two consecutive solutions is linear, by turning the problem into a constrained cycle enumeration problem. Such patterns, in a de Bruijn graph representation of the reads obtained by sequencing, are related to polymorphisms in DNA- or RNA-seq data.

In this work we plan to revise the main techniques for enumeration algorithms and to show four examples of enumeration algorithms that can be applied to efficiently deal with some biological problems modelled by using biological networks: enumerating central and peripheral nodes of a network, enumerating stories, enumerating paths or cycles, and enumerating bubbles. Notice that the corresponding computational problems we define are of more general interest and our results hold in the case of arbitrary graphs. Enumerating all the most and less central vertices in a network according to their eccentricity is an example of an enumeration problem whose solutions are polynomial and can be listed in polynomial time, very often in linear or almost linear time in practice. Enumerating stories, i.e. all maximal directed acyclic subgraphs of a graph G whose sources and targets belong to a predefined subset of the vertices, is on the other hand an example of an enumeration problem with an exponential number of solutions, that can be solved by using a non trivial brute-force approach. Given a metabolic network, each individual story should explain how some interesting metabolites are derived from some others through a chain of reactions, by keeping all alternative pathways between sources and targets. Enumerating cycles or paths in an undirected graph, such as a protein-protein interaction undirected network, is an example of an enumeration problem in which all the solutions can be listed through an optimal algorithm, i.e. the time required to list all the solutions is dominated by the time to read the graph plus the time required to print all of them. By extending this result to directed graphs, it would be possible to deal more efficiently with feedback loops and signed paths analysis in signed or interaction directed graphs, such as gene regulatory networks. Finally, enumerating mouths or bubbles with a source s in a directed graph, that is enumerating all the two vertex-disjoint directed paths



74,89 €

69,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9789462390966

Medium: Buch

ISBN: 978-94-6239-096-6

Verlag: Atlantis Press

Erscheinungstermin: 01.04.2015

Sprache(n): Englisch

Auflage: Erscheinungsjahr 2015

Serie: Atlantis Studies in Computing

Produktform: Gebunden

Gewicht: 3849 g

Seiten: 151

Format (B x H): 160 x 241 mm

