Information Sciences 340–341 (2016) 279–304

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Simple graphs in granular computing

Giampiero Chiaselotti [a], Davide Ciucci [b,*], Tommaso Gentile [a,c]

[a] Dipartimento di Matematica e Informatica, Università della Calabria, Via Pietro Bucci, Cubo 30B, Arcavacata di Rende (CS) 87036, Italy
[b] Dipartimento di Informatica, Sistemistica e Comunicazione, Università di Milano-Bicocca, Via Bicocca degli Arcimboldi, 8 – 20126 Milano, Italy
[c] Istituto Nazionale di Fisica Nucleare, Gruppo collegato di Cosenza, Università della Calabria, Via Pietro Bucci, Cubo 31C, Arcavacata di Rende (CS) 87036, Italy

ARTICLE INFO

Article history:
Received 2 April 2015
Revised 18 November 2015
Accepted 28 December 2015
Available online 12 January 2016

Keywords:
Granular computing
Information tables
Rough set theory
Reduct hypergraph
Simple graphs
Petersen graph

ABSTRACT

Given a graph, we interpret its adjacency matrix as an information table. We study this correspondence in two directions. Firstly, on the side of graphs by applying to it standard techniques from granular computing. In this way, we are able to connect automorphisms on graphs to the so-called indiscernibility relation and a particular hypergraph built from the starting graph to core and reducts. On the other hand, new concepts are introduced on graphs that have an interesting correspondence on information tables. In particular, some new topological interpretations of the graph and the concept of extended core are given.

1. Introduction

Granular computing (briefly GrC) deals with representing and processing information in the form of some type of aggregates. These aggregates are generally called *information granules* or simply granules and they arise in the process of data abstraction and knowledge extraction from data. Generally speaking, information granules are collections of entities arranged together due to their similarity, functional or physical adjacency, indistinguishability, coherency, and so on. The scope of GrC covers various fields of study related to knowledge representation and extraction. In 1979 the concept of *information granularity* was introduced by Zadeh [48] and it was related to the research on fuzzy sets. Then, the term *granular computing* was used again by Zadeh in 1997 [49] with the following words: “a subset of computing with words is granular computing”. Since 1979, granular computing has become a very developed area of research in the scope of both applied and theoretical information sciences [28,30]. Today GrC has emerged as a standalone research area that intersects and finds application in several fields related to knowledge management: interval analysis [21], machine learning [47], formal concept analysis [20,42], data mining [19,23,24,39,43], database theory [18,31], rough set theory [26,27,44], interactive computing [33,34] and fuzzy set theory [29,49].

In this paper we are mainly interested in the part of GrC related to information tables and their management by rough-set theory tools. An information table, from now on denoted as $\mathcal{I}$, is simply a bi-dimensional table associating to any *object* (rows of the table) a *value* (content of the table) for each *attribute* under investigation (columns). If all values are 0 or 1, the

* Corresponding author. Tel.: +390264487902; fax: +390264487839.
E-mail addresses: giampiero.chiaselotti@unical.it (G. Chiaselotti), ciucci@disco.unimb.it (D. Ciucci), gentile@mat.unical.it (T. Gentile).

http://dx.doi.org/10.1016/j.ins.2015.12.042
0020-0255/

information table is said *Boolean*. An information table is obviously a very common structure in various fields of study, both of qualitative and quantitative type. In his seminal works concerning rough set (RS) theory [25,26], Pawlak introduced several investigation tools in order to analyze a generic information table and reduce its complexity. The fundamental concept that allows us to connect the Pawlak theory to the GrC paradigm is the *indiscernibility relation* I_A, where A is any attribute subset. Very simply, two objects are *A-indiscernible* if they assume the same value for all attributes $a \in A$. Now, since I_A is an equivalence relation, we can partition the object set in equivalence classes, i.e., in granules, and therefore we can interpret a great part of Pawlak's theory as a particular type of GrC, from now on GrC-RS (for a detailed analysis concerning the links between GrC and rough sets see [46]).

There are at least four concepts in GrC-RS that deserve a special consideration: the *indiscernibility relation*, the *core*, the *reduct family* and the *discernibility matrix*. The core [26] can be intuitively described as the most important part of the attribute set that characterizes an information table. A reduct [26] is a subset of attributes that provides sufficient information to fully characterize the entire table. Finally, the discernibility matrix [32] is a square matrix having in the place (i, j) the set of attributes on which the objects u_i and u_j differ.

With the term *granular computing on graphs and hypergraphs* (abbreviated GrC-GH), we mean all the studies that link graph and hypergraph theory with GrC. We notice that GrC-GH is not a research sub-field of GrC-RS since one can study several types of granularity outside of rough set theory [28,30], and we will see in Section 1.3 that this has indeed been done. So, here we are interested in the intersection of the two fields GrC-GH and GrC-RS. More in detail, we will interpret the adjacency matrix of a (finite) simple undirected graph G as a Boolean information table $\mathcal{I}[G]$, where the object set and the attribute set coincide. In this way, we can efficiently use the theoretical framework developed in GrC-RS to find new properties concerning the graph G. From this point of view, the advantage is twofold: both in direction from GrC-RS towards discrete mathematics (briefly GrC-RS → DM) and also from discrete mathematics towards GrC-RS (briefly DM → GrC-RS). We will now illustrate this double advantage by providing more details about two results obtained in this paper.

1.1. An example of GrC-RS → DM

In the Boolean information table $\mathcal{I}[G]$, for a fixed vertex subset A, the indiscernibility relation I_A is an equivalence relation between vertices of G that can be characterized in geometric terms. Specifically, two vertices are I_A-indiscernible if and only if they are in a type of "symmetrical" position with respect to all the vertices of A. This means that we can interpret A as if it were a "symmetry block" for the graph G (see Theorem 3.3). On the basis of this result it is natural to ask what do the core, the reducts and the granular lattice for common families of graphs become. In this paper we treat the complete graphs, the complete bipartite graphs, the Petersen graph and the paths on a fixed number of vertices cases.

Let us consider for example the classical Petersen graph (briefly *Pet*) in Fig. 1.

For its symmetry properties (the Petersen graph is both vertex-transitive and edge-transitive) the core is empty (see Corollary 6.3). However the key point, a priori not obvious, turns out to be that the reducts of the Petersen graph can be characterized by a geometric point of view. In fact, we will provide a purely combinatorial and geometric proof of the following classification result (a part of Theorem 6.6): *a vertex subset A of Pet is a reduct of the Boolean information table* $\mathcal{I}[Pet]$ *if and only if A has cardinality* 5 *and the subgraph of Pet generated by A is isomorphic to one of the graphs in* Fig. 2.

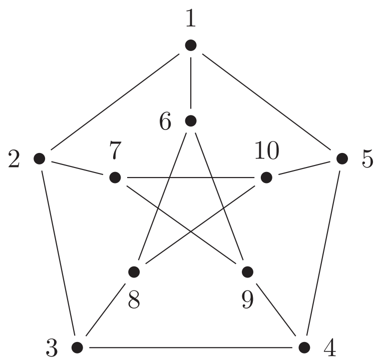

Fig. 1. The Petersen graph.

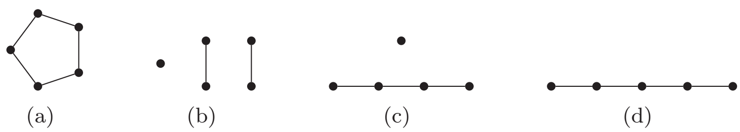

Fig. 2. Reducts of the Petersen graph.