Extension of Ugi’s Scheme for Model-Driven Classification of Chemical Reactions

Shyantani Maiti, Department of Computer Science and Technology, Bengal Engineering and Science University, Shibpur, Howrah, India

Sanjay Ram, Department of Computer Science and Technology, Bengal Engineering and Science University, Shibpur, Howrah, India

Somnath Pal, Department of Computer Science and Technology, Bengal Engineering and Science University, Shibpur, Howrah, India

ABSTRACT

The first step to predict the outcome of a chemical reaction is to classify existing chemical reactions, on the basis of which possible outcome of unknown reaction can be predicted. There are two approaches for classification of chemical reactions: Model-Driven and Data-Driven. In model-driven approach, chemical structures are usually stored in a computer as molecular graphs. Such graphs can also be represented as matrices. The most preferred matrix representation to store molecular graph is Bond-Electron matrix (BE-matrix). The Reaction matrix (R-matrix) of a chemical reaction can be obtained from the BE-matrices of educts and products was shown by Ugi and his co-workers. Ugi’s Scheme comprises of 30 reaction classes according to which reactions can be classified, but in spite of such reaction classes there were several reactions which could not be classified. About 4000 reactions were studied in this work from The Chemical Thesaurus (a chemical reaction database) and accordingly 24 new classes have emerged which led to the extension of Ugi’s Scheme. An efficient algorithm based on the extended Ugi’s scheme have been developed for classification of chemical reactions. Reaction matrices being symmetric, matrix implementation of extended Ugi’s scheme using conventional upper/lower tri-angular matrix is of O(n²) in terms of space complexity. Time complexity of similar matrix implementation is O(n²) in worst case. The authors’ proposed algorithm uses two fixed size look-up tables in a novel way and requires constant space complexity. Worst case time complexity of their algorithm although still O(n²) but it outperforms conventional matrix implementation when number of atoms or components in the chemical reaction is 4 or more.

Keywords: Bond-electron Matrix, Model-driven Classification of Chemical Reaction, Reaction Matrix, Space and Time Complexity, Ugi’s Scheme

DOI: 10.4018/IJCCE.2015010103
INTRODUCTION

Representation of chemical reactions is the first step in dealing with the prediction of chemical reactions in Chemoinformatics. In this paper we used Bond-Electron matrix (Dugundji & Ugi, 1973) for representing chemical reactions. Reaction matrices (Ivanciuc, 2010; Gasteiger & Engel, 2003; Kvasnickaa & Pospichala, 1991), are useful derivatives from bond electron matrices of educts and products of chemical reactions. Classification of chemical reactions is the second step for prediction of chemical reactions which utilizes reaction matrices as shown by Ugi and his co-workers (Dugundji & Ugi, 1973; Ivanciuc, 2010) in the mid 1970s. Ugi developed 30 reaction classes according to which chemical reactions could be classified which came to be known as Ugi’s Scheme (Gasteiger & Engel, 2003; Garagnani & Bart, 1977). An efficient implementation of Ugi’s scheme was proposed in (Ram & Pal, 2012). However, it has been observed that some important chemical reactions could not be classified by this scheme. So more reaction classes need to be introduced thereby requiring extension of Ugi’s Scheme. With this objective, we have searched The Chemical Thesaurus (Cousins & AmChem, 2001) containing about 4000 reactions and came up with 24 new reaction classes over and above 30 reaction classes of Ugi’s scheme. In this connection, we have developed an algorithm to obtain reaction class templates from reaction matrices. We have also modified the classification algorithm of (Ram & Pal, 2012) to an efficient algorithm to accommodate the new reaction classes.

The construction of Reaction matrices, which are derived from Bond Electron Matrices of educts and products, has been discussed in section Background. Section Background also discusses classification of chemical reactions using Ugi’s scheme. The extension of Ugi’s scheme along with above mentioned first algorithm to obtain reaction class templates from reaction matrix is discussed in the following section. The next section discusses the modified algorithm to classify chemical reactions using Extended Ugi’s scheme. Theoretical development of the time and space complexity of our algorithm has been shown in section Results and Discussion. Finally, the last section summarizes our conclusion.

BACKGROUND

B.E. Matrix for Representation of Chemical Reaction

A chemical compound can be represented by graph (Trinajstic, 1992) known as molecular graph (Beck et. al., 1969). In this type of graph, the atoms or components of chemical compounds are represented by nodes in the graph and bonds present in the chemical compound are represented by edges in the graph. Bond electron matrices are matrix representation of this molecular graph. The bond-electron matrix (BE-matrix) was introduced in the Dugundji-Ugi model (Dugundgi & Ugi, 1973). Suppose a molecular graph contains n nodes, thus the bond electron matrix will be of size n x n (Ivanciuc, 2003). The entries in the bond-electron matrices represent the type of bond present in between the several atoms or components in the chemical compound. Such bond-electron matrices (BE-matrix) have been widely used (Mandal & Pal, 2013; Leber et. al., 2009; Bharatam et. al. 2008). In this paper we have worked on Model driven classification of chemical reactions using Bond-electron matrices.

The representation of BE matrix for acetic acid in Figure 1 is shown in Table 1. The 0th row/column in the BE-matrix of Table 1 represents each atom of the chemical compound/molecule. Bond order is represented by the off-diagonal elements between the corresponding row and column atoms or components and the number of free electrons of the corresponding atom is shown in the left to right diagonal elements (e.g., O6 = 4 in Table 1).
Hydrogen-Bonded Interactions in the Systems L-Cysteine - H2SeO3 and L-Cysteine - H2SeO4: A DFT Study
www.igi-global.com/article/hydrogen-bonded-interactions-systems-cysteine/56116?camid=4v1a