*Journal of Mathematical Chemistry, Vol. 39, Nos. 3/4, May 2006 (© 2006)*
DOI: 10.1007/s10910-005-9051-z

# Alternative integer-linear-programming formulations of the Clar problem in hexagonal systems

Khaled Salem*,** and Hernán Abeledo

*Department of Engineering Management and Systems Engineering, The George Washington University, Washington, D.C. 20052, USA*
E-mails: khaled@gwu.edu; abeledo@gwu.edu

Received 10 March 2005 / Published online: 13 January 2006

We present two alternative objective functions for the integer-linear-programming formulation of the Clar problem in hexagonal systems proposed by Hansen and Zheng [1994, J. Math. Chem. 15, 93]. Also, we note that these formulations can be solved in polynomial-time with linear programming algorithms.

**KEY WORDS:** hexagonal, benzenoid, resonant, integer programming, linear programming

**AMS subject classification:** 90C90, 05C90

## 1. Introduction

A hexagonal system $H$ is a 2-connected subgraph of the hexagonal lattice without non-hexagonal interior faces. The nodes of $H$ are classified into two types. A node of $H$ lying on the boundary of the exterior face of $H$ is called an external node, otherwise, it is called an internal node. If a hexagonal system has no internal nodes, it said to be catacondensed, otherwise, it is pericondensed. A catacondensed hexagonal system is called a hexagonal chain if each of its hexagons is adjacent to at most two hexagons.

Let $P$ be a non-empty set of hexagons of a hexagonal system $H$. We call $P$ a resonant set of $H$ if the hexagons in $P$ are pair-wise disjoint and $H$–$P$ (the subgraph of $H$ obtained by deleting from $H$ the vertices of the hexagons in $P$) has a perfect matching or is empty [1] or, equivalently, if the hexagons in $P$ are pair-wise disjoint and there exists a perfect matching of $H$ that contains a perfect matching of each hexagon in $P$ [2]. A resonant set is maximum if its cardinality is. The cardinality of a maximum resonant set is of significance

*Corresponding author.
**Postal address: Khaled Salem, 66 Youssef Abbas Street, Apt. 71, Nasr city, Cairo, Postal Code 11371, Egypt.

in the chemistry of benzenoid hydrocarbons [3]. It is called the Clar number [4]. A hexagonal system that has a perfect matching also has a resonant set [5]. By solving the Clar problem in a hexagonal system that has a perfect matching, we mean obtaining a maximum resonant set and computing the Clar number.

Hansen and Zheng [4] showed that the solution of the so-called minimum weight cut cover problem for a hexagonal system that has a perfect matching yields an upper bound for the Clar number and conjectured that equality holds. Abeledo and Atkinson [2] used a network flow formulation of the minimum weight cut cover problem to prove this conjecture. This allowed computing the Clar number, but not a maximum resonant set, via polynomial-time (efficient) combinatorial (graph-theoretic) algorithms. As a background reference on network flow problems we suggest Cook et al. [6].

A polynomial-time combinatorial algorithm to solve the Clar problem in catacondensed hexagonal systems was given by Atkinson [7]. A version of this algorithm that computes only the Clar number was re-invented by Klavžar et al. [8]. Salem and Gutman [9] used this algorithm to obtain an algebraic expression for the Clar number of a hexagonal chain in terms of its $L$, $A$-sequence. (This sequence was introduced by Gutman [10].)

Zhang and Li [11] designed a polynomial-time combinatorial algorithm to solve the Clar problem of a restricted class of hexagonal systems. Salem [12] obtained a result that is potentially useful in designing a polynomial-time combinatorial algorithm to solve the Clar problem in an arbitrary hexagonal system that has a perfect matching.

Away from combinatorial approaches, Hansen and Zheng [13] formulated the Clar problem as an integer linear program. In this paper we present two alternative objective functions for the integer-linear-programming formulation proposed by Hansen and Zheng [13]. These will be given in Section 3, but first we give some definitions in Section 2.

## 2. Preliminaries for linear and integer programming

Given an $m \times n$ real matrix $A$ and a vector $\mathbf{b} \in R^m$, the set $\{\mathbf{x} \in R^n : A\mathbf{x} \leqslant b\}$ is called a polyhedral set (or a polyhedron). Given a vector $\mathbf{c} \in R^n$, it defines a linear function $\mathbf{c}^{\mathbf{t}}\mathbf{x}$ from $R^n$ to $R$.

Linear programming (LP) concerns the problem of maximizing (max) or minimizing (min) a linear function over a polyhedron. Examples are

$$\max\{\mathbf{c}^{\mathbf{t}}\mathbf{x} : A\mathbf{x} \leqslant \mathbf{b}\}$$
$$\max\{\mathbf{c}^{\mathbf{t}}\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant 0\}$$
$$\min\{\mathbf{c}^{\mathbf{t}}\mathbf{x} : A\mathbf{x} \geqslant \mathbf{b}\}$$
$$\min\{\mathbf{c}^{\mathbf{t}}\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant 0\}$$

The linear function to be optimized is called the objective function. The polyhedron is called the feasible region and a vector in the polyhedron is called a feasible solution. An optimal solution $\mathbf{x}^*$ of a maximization problem (respectively, a minimization problem) is a feasible solution such that $\mathbf{c}^t\mathbf{x}^* \geqslant \mathbf{c}^t\mathbf{x}$ (respectively, $\mathbf{c}^t\mathbf{x}^* \leqslant \mathbf{c}^t\mathbf{x}$) for any feasible solution $\mathbf{x}$. The real number $\mathbf{c}^t\mathbf{x}^*$ is called the optimal value.

An integer linear programming (ILP) problem is essentially an LP problem with the additional requirement that the variables take integral values. An example of an integer linear programming problem is $\max\{\mathbf{c}^t\mathbf{x}\colon A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant 0, \mathbf{x} \text{ integral}\}$.

ILP and LP problems differ greatly in their computational complexity: whereas polynomial-time algorithms exist for solving LP problems [14–16], the class of ILP problems is NP-complete [17]. This means that there is no known polynomial-time algorithm that can solve an arbitrary ILP problem and that it is generally believed that such an algorithm cannot be designed.

As background references on linear and integer programming we suggest Chvátal [18] and Schrijver [19].

## 3. The Clar problem and integer linear programming

Consider a hexagonal system that has a perfect matching. Let $Q$ be its node-edge incidence matrix and $R$ its node-hexagon incidence matrix. Consider the integer linear programming problem:

$$\max\{\mathbf{1}^t\mathbf{y} : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \text{ binary}\}.$$

This ILP was proposed by Hansen and Zheng [13] to model the Clar problem. In fact, they showed that there is a 1–1 mapping of the set of optimal solutions of this ILP problem onto the family of maximum resonant sets of the hexagonal system, i.e. the mapping is bijective. The image of an optimal solution $(\mathbf{x}, \mathbf{y})$ under this mapping is the set of hexagons corresponding to the $\mathbf{1}$'s in $\mathbf{y}$.

**Theorem 1.** Consider a hexagonal system that has a perfect matching. Let $Q$ be its node-edge incidence matrix and $R$ its node-hexagon incidence matrix. The following integer linear programming problems have the same set of optimal solutions.

(i) $\max\{\mathbf{1}^t\mathbf{y} : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \text{ binary}\}$

(ii) $\min\{\mathbf{1}^t\mathbf{x} + \mathbf{1}^t\mathbf{y} : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \text{ binary}\}$

*Proof.* Let $(\mathbf{x}, \mathbf{y})$ be an optimal solution of (i). Then $(\mathbf{x}, \mathbf{y})$ is a feasible solution of (i) and of (ii). Assume that $(\mathbf{x}, \mathbf{y})$ is not an optimal solution of (ii). Then there exists a feasible solution of (ii), $(\mathbf{x}', \mathbf{y}')$ say, such that $\mathbf{1^t x}' + \mathbf{1^t y}' < \mathbf{1^t x} + \mathbf{1^t y}$. Since $(\mathbf{x}, \mathbf{y})$ is a feasible solution of (i),

$$Q\mathbf{x} + R\mathbf{y} = 1.$$

Each edge is incident with two nodes and each hexagon is incident with six nodes and so summing the equations, we obtain

$$2(\mathbf{1^t x}) + 6(\mathbf{1^t y}) = |V|,$$

where $|V|$ is the number of nodes of the hexagonal system. Similarly,

$$2(\mathbf{1^t x}') + 6(\mathbf{1^t y}') = |V|.$$

Hence, $2(\mathbf{1^t x}) + 6(\mathbf{1^t y}) = 2(\mathbf{1^t x}') + 6(\mathbf{1^t y}')$. Dividing by 2, we obtain $\mathbf{1^t x} - \mathbf{1^t x}' = 3(\mathbf{1^t y}') - 3(\mathbf{1^t y})$. Since $\mathbf{1^t x}' + \mathbf{1^t y}' < \mathbf{1^t x} + \mathbf{1^t y}$, we have $\mathbf{1^t x} - \mathbf{1^t x}' > \mathbf{1^t y}' - \mathbf{1^t y}$. Thus, $3(\mathbf{1^t y}') - 3(\mathbf{1^t y}) > \mathbf{1^t y}' - \mathbf{1^t y}$ which implies that $\mathbf{1^t y}' > \mathbf{1^t y}$. Recall that **(x, y)** is an optimal solution of (i) and note that $(\mathbf{x}', \mathbf{y}')$ is a feasible solution of (i). Hence, $\mathbf{1^t y} \geqslant \mathbf{1^t y}'$, a contradiction. Hence, $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (ii).

Let $(\mathbf{x}, \mathbf{y})$ be an optimal solution of (ii). Then $(\mathbf{x}, \mathbf{y})$ is a feasible solution of (ii) and, thus, of (i). Assume that $(\mathbf{x}, \mathbf{y})$ is not an optimal solution of (i). Then there exists a feasible solution of (i), $(\mathbf{x}', \mathbf{y}')$ say, such that $\mathbf{1^t y}' > \mathbf{1^t y}$. Recall that $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (ii) and note that $(\mathbf{x}', \mathbf{y}')$ is a feasible solution of (ii). Hence, $\mathbf{1^t x} + \mathbf{1^t y} \leqslant \mathbf{1^t x}' + \mathbf{1^t y}'$ implying that $\mathbf{1^t x} - \mathbf{1^t x}' \leqslant \mathbf{1^t y}' - \mathbf{1^t y}$. Since $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ are feasible solutions of (ii), we have $2(\mathbf{1^t x}) + 6(\mathbf{1^t y}) = 2(\mathbf{1^t x}') + 6(\mathbf{1^t y}')$ and so $\mathbf{1^t x} - \mathbf{1^t x}' = 3(\mathbf{1^t y}') - 3(\mathbf{1^t y})$. Thus, $\mathbf{1^t y}' - \mathbf{1^t y} \geqslant 3(\mathbf{1^t y}') - 3(\mathbf{1^t y})$ implying that $\mathbf{1^t y} \geqslant \mathbf{1^t y}'$, a contradiction. Hence, $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (i). **Q.E.D.**

**Theorem 2.** Consider a hexagonal system that has a perfect matching. Let $Q$ be its node-edge incidence matrix and $R$ its node-hexagon incidence matrix. The following integer linear programming problems have the same set of optimal solutions.

(i) $\max\{\mathbf{1^t y} : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \text{ binary}\}$

(ii) $\min\{\mathbf{1^t x} : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \text{ binary}\}$

*Proof.* Let $(\mathbf{x}, \mathbf{y})$ be an optimal solution of (i). Then $(\mathbf{x}, \mathbf{y})$ is a feasible solution of (i) and of (ii). Assume that $(\mathbf{x}, \mathbf{y})$ is not an optimal solution of (ii). Then there exists a feasible solution of (ii), $(\mathbf{x}', \mathbf{y}')$ say, such that $\mathbf{1^t x}' < \mathbf{1^t x}$. Since $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ are feasible solutions of (ii), we have $2(\mathbf{1^t x}) + \mathbf{6}(\mathbf{1^t y}) = |V| = 2(\mathbf{1^t x}') + 6(\mathbf{1^t y}')$, where $|V|$ is the number of nodes of the hexagonal system. This implies that $\mathbf{1^t x} - \mathbf{1^t x}' = 3(\mathbf{1^t y}') - 3(\mathbf{1^t y})$. Recall that $\mathbf{1^t x}' < \mathbf{1^t x}$. Hence, $\mathbf{1^t y} < \mathbf{1^t y}'$. Note that $(\mathbf{x}', \mathbf{y}')$ is a feasible solution of (i) and recall that $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (i). Hence, $\mathbf{1^t y} \geqslant \mathbf{1^t y}'$, a contradiction. Thus, $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (ii).

Let $(\mathbf{x}, \mathbf{y})$ be an optimal solution of (ii). Then $(\mathbf{x}, \mathbf{y})$ is a feasible solution of (ii) and of (i). Assume that $(\mathbf{x}, \mathbf{y})$ is not an optimal solution of (i). Then there exists a feasible solution of (i), $(\mathbf{x}', \mathbf{y}')$ say, such that $\mathbf{1^t y}' > \mathbf{1^t y}$. Since $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ are feasible solutions of (i), we have $2(\mathbf{1^t x}) + 6(\mathbf{1^t y}) = 2(\mathbf{1^t x}') + 6(\mathbf{1^t y}')$ and so $\mathbf{1^t x} - \mathbf{1^t x}' = 3(\mathbf{1^t y}') - 3(\mathbf{1^t y})$. Recall that $\mathbf{1^t y}' > \mathbf{1^t y}$. Hence, $\mathbf{1^t x} > \mathbf{1^t x}'$. Note that $(\mathbf{x}', \mathbf{y}')$ is a feasible solution of (ii) and recall that $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (ii). Hence, $\mathbf{1^t x} \leqslant \mathbf{1^t x}'$, a contradiction. Hence, $(\mathbf{x}, \mathbf{y})$ is an optimal solution of (i). **Q.E.D.**

## Remarks

The ILP formulation given by Hansen and Zheng [13] allowed computing the Clar problem for an arbitrary hexagonal system via ILP algorithms. However, this result was not satisfactory from a computational complexity perspective since as mentioned earlier, the ILP problem is NP-complete. This serious drawback was overcome when Abeledo and Atkinson [2, 7] noted that the formulation given by Hansen and Zheng [13] could be solved efficiently as an LP. In particular, they proved that the matrix $[QR]$ in the ILP formulation is unimodular (see Schrijver [19] for the definition of unimodular matrices). Thus, a result by Truemper [20] was applicable which assured that an LP problem over the polyhedron

$$\{(\mathbf{x}, \mathbf{y}) : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \geqslant 0\}$$

has an integral optimal solution. In particular, the LP problem

$$\max\{\mathbf{1^t y} : Q\mathbf{x} + R\mathbf{y} = 1; \mathbf{x}, \mathbf{y} \geqslant 0\}$$

has an integral optimal solution (which is also optimal for the ILP formulation of the Clar problem). This result was originally conjectured by Hansen and Zheng [13]. In particular, LP algorithms can be used to obtain this integral optimal solution in polynomial-time [7, 19]. The same analysis applies to the alternative ILP formulations presented in this paper since they are defined over the same feasible region.

## References

[1] I. Gutman, Wiss. Z. Thechn. Hochsch. Ilmenau 29 (1983) 57.
[2] H. Abeledo and G. Atkinson, in: *Discrete Mathematical Chemistry,* eds. P. Hansen, P. Fowler and M. Zheng (Am. Math. Soc., Providence, RI, 2000), pp 1–8.
[3] E. Clar, *The Aromatic Sextet* (Wiley, London, 1972).

[4] P. Hansen and M.L. Zheng, J. Chem. Soc., Faraday Trans. 88 (1992) 1621.
[5] I. Gutman, *Proceedings of the Fourth Yugoslav Seminar on Graph Theory* (Novi Sad: Institute of Mathematics, University of Novi Sad, 1983) 151.
[6] W.J. Cook, W.H. Cunnigham, W.R. Pulleyblank and A. Schrijver, *Combinatorial Optimization* (John Wiley & Sons, Inc., 1998).
[7] G. Atkinson, DSc dissertation, The George Washington University, USA (1998).
[8] S. Klavžar, P. Žigert and I. Gutman, J. Mol. Struct. (Theochem) 586 (2002) 235.
[9] K. Salem and I. Gutman, Chem. Phys. Lett. 394 (2004) 283.
[10] I. Gutman, Theor. Chim. Acta 45 (1977) 309.
[11] F.J. Zhang and X.L. Li, MATCH- Commun. Math. Comput. Chem. 24 (1989) 333.
[12] K. Salem, MATCH – Commun. Math. Comput. Chem. 53 (2005) 419.
[13] P. Hansen and M.L. Zheng, J. Math. Chem. 15 (1994) 93.
[14] L.G. Khachiyan, Doklady Akademii Nauk SSSR 244 (1979) 1093 [in Russian, English translation: Soviet Mathematics Doklady 20 (1979) 191].
[15] L.G. Khachiyan, Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki 20 (1980) 51 [in Russian, English translation: USSR Computational Mathematics and Mathematical Physics 20 (1980) 53].
[16] N. Karmarkar, *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing* (Washington, 1984), ACM, New York, 1984, pp. 302–311 [also: Combinatorica 4 (1984) 373].
[17] S.A. Cook, *Proceedings of Third Annual ACM Symposium on Theory of Computing* (Shaker Heights, Ohio, 1971), ACM, New York, 1971, pp. 151–158.
[18] V. Chvátal, *Linear Programming* (Freeman, New York, 1983).
[19] A. Schrijver, *Theory of Linear and Integer Programming* (John Wiley & Sons Ltd, Chichester, 1986).
[20] K. Truemper, SIAM J. Appl. Math. 35 (1978) 328.