

Horn Renamability and Hypergraphs

Dušan Hvalica

Faculty of Economics, University of Ljubljana, Slovenia

Satisfiability testing in the context of directed hypergraphs is discussed. A characterization of Horn-renamable formulae is given and a subclass of SAT that belongs to \mathcal{P} is described. Two algorithms for Horn renaming with linear time complexity are presented.

Keywords: satisfiability, Horn, renamable, hypergraph

1. Introduction

The *satisfiability problem*, SAT, consisting of testing the satisfiability of a propositional formula, is known to be \mathcal{NP} -complete. On the other hand, several subclasses of propositional formulae are known, so that the restriction of SAT to such a subclass is solvable in polynomial time [7, 2] — among them Horn formulae [5] and formulae that are reducible to Horn, such as Horn-renamable formulae [10, 4, 1, 3, 9]. Naturally, these classes have attracted much attention.

SAT can be translated in a natural way into the context of directed hypergraphs — testing of satisfiability translates to searching for a zero-cardinality cut [6, 8].

The aim of this paper is to show that the hypergraph approach to SAT gives rise to new concepts, advantageous in designing algorithms and in specifying classes of formulae. We demonstrate this by giving a characterization of Horn-renamable formulae in terms of directed hypergraphs (as a by-product we obtain a subclass of SAT that belongs to \mathcal{P}); we also describe two algorithms for Horn renaming with linear time complexity.

2. Definitions and Notation

An (oriented) *hypergraph* G is defined as $G = (\mathcal{V}, \mathcal{A})$, where \mathcal{V} and \mathcal{A} are the sets of *nodes* and *hyperarcs*, respectively. A hyperarc \mathcal{E} is defined as $\mathcal{E} = (T(\mathcal{E}), H(\mathcal{E}))$, where $T(\mathcal{E}), H(\mathcal{E}) \subset \mathcal{V}$; the sets $T(\mathcal{E})$ and $H(\mathcal{E})$ are called the *tail* and *head* of \mathcal{E} , respectively. A hyperarc, whose head has at most one element, is called a *B-arc* (backward (hyper)arc), a hypergraph, the hyperarcs of which are all B-arcs, is a *B-graph*.

A *subhypergraph* of a hypergraph $G = (\mathcal{V}, \mathcal{A})$ is a hypergraph $G_1 = (\mathcal{V}_1, \mathcal{A}_1)$ such that $\mathcal{V}_1 \subset \mathcal{V}$ and $\mathcal{A}_1 \subset \mathcal{A}$. When suitable, \mathcal{V}_1 will be denoted by $\mathcal{V}(G_1)$ and \mathcal{A}_1 by $\mathcal{A}(G_1)$.

For any node u its *backward star* $BS(u)$ is defined by $BS(u) = \{\mathcal{E}; u \in H(\mathcal{E})\}$, while its *forward star* $FS(u)$ is $FS(u) = \{\mathcal{E}; u \in T(\mathcal{E})\}$. A node u for which $BS(u) = \emptyset$ or $FS(u) = \emptyset$ will be called a *tip node*.

A *path* is a sequence $u_1, \mathcal{E}_1, u_2, \mathcal{E}_2, \dots, \mathcal{E}_{q-1}, u_q$, such that $u_i \in H(\mathcal{E}_{i-1})$ for $1 < i \leq q$ and $u_i \in T(\mathcal{E}_i)$ for $1 \leq i < q$. If $u_q \in T(\mathcal{E}_1)$, such a path is called a *cycle*.

Let S be any set of nodes in a B-graph G . A *B-hyperpath* or, shortly, *B-path*, based on S and ending at t is any hypergraph P , which is a minimal subhypergraph of G such that

- $t \in \mathcal{V}(P)$,
- for every $v \in \mathcal{V}(P)$ there exists in P a simple cycle-free path from some $u \in S$ to v .

When such a B-path exists, we also say that t is *B-connected* to S .

3. Satisfiability of Propositional Formulae

Let A be a set of propositional variables and B a set of clauses over A , i.e., of formulae of the form:

$$C_1 \wedge \dots \wedge C_m \Rightarrow D_1 \vee \dots \vee D_n, \quad (1)$$

where $C_1, \dots, C_m, D_1, \dots, D_n$ belong to $A \cup \{true, false\}$. We say that B is *satisfiable* if there exists a truth assignment $A \rightarrow \{true, false\}$ such that every clause in B is *true*.

The *satisfiability problem*, SAT, consists of testing the satisfiability of B . As already pointed out, it is \mathcal{NP} -complete.

A clause (1) with $n \leq 1$ is known as a *Horn clause*. If every clause in B is a Horn clause, we speak of *Horn-SAT*. It is known that it is solvable in linear time.

To every instance of SAT a hypergraph can be assigned in the following way: its nodes are the propositions, while its hyperarcs correspond to clauses such that the hyperarc corresponding to clause (1) is $(\{C_1, \dots, C_m\}, \{D_1, \dots, D_n\})$. If every clause is a Horn clause, the resulting hypergraph is clearly a B -graph. Moreover, in that case the problem of satisfiability translates into the problem of verifying B -connectedness in this B -graph; in general, a formula is satisfiable if and only if there is a 0-cardinality cut in the corresponding hypergraph (for details, see [6]).

One of the polynomially solvable subclasses of SAT is the class of *Horn-renamable formulae*, i.e., CNF formulae, for which a *renaming* (replacing, for some variables $x_i, i \in I$, each occurrence of x_j and \bar{x}_j by \bar{y}_j and y_j , respectively) exists that turns the formula into a Horn formula. Clearly, every replacement of x_j and \bar{x}_j by \bar{y}_j and y_j in the formula corresponds to a switch of position of node x_j in every hyperarc of the corresponding hypergraph — if $x_j \in T(\mathcal{E})$ we move it to $H(\mathcal{E})$ and *vice versa*. Thus, a formula is Horn-renamable iff there exists a switching that turns the corresponding hypergraph into a B -graph.

4. Hypergraphs and B -graphs

Let us introduce the following new concept:

Definition 1. A *track* is a sequence

$$u_1, \mathcal{E}_1, u_2, \mathcal{E}_2, \dots, \mathcal{E}_{q-1}, u_q, \quad (2)$$

such that

- for any triple $u_i, \mathcal{E}_i, u_{i+1}$ on the track we have $u_i \neq u_{i+1}$ and $u_i, u_{i+1} \in T(\mathcal{E}_i) \cup H(\mathcal{E}_i)$,
- for any triple $\mathcal{E}_i, u_{i+1}, \mathcal{E}_{i+1}$ on the track we have $\mathcal{E}_i \neq \mathcal{E}_{i+1}$ and one of the following holds:
 - $u_{i+1} \in H(\mathcal{E}_i)$ and $u_{i+1} \in T(\mathcal{E}_{i+1})$,
 - $u_{i+1} \in T(\mathcal{E}_i)$ and $u_{i+1} \in H(\mathcal{E}_{i+1})$.

If $u_{q-1}, \mathcal{E}_{q-1}, u_q, \mathcal{E}_1, u_2$ is a track, then track (2) is called a *hypercycle* (of course, in such case either $u_q \in T(\mathcal{E}_1)$ or $u_q \in H(\mathcal{E}_1)$; often we have $u_q = u_1$).

Thus, a track consists of paths and “reversed” paths glued together at common hyperarcs.

For instance, $u_1, \mathcal{E}_1, u_4, \mathcal{E}_2, u_5, \mathcal{E}_4, u_6, \mathcal{E}_3, u_4, \mathcal{E}_5, u_3$ is a track in the hypergraph in Figure 1.

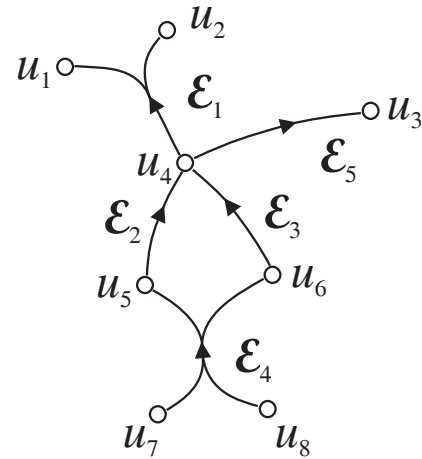


Figure 1.

If $T = u_1, \mathcal{E}_1, \dots, \mathcal{E}_{i-1}, u_i$ and $T' = u_i, \mathcal{E}'_1, \dots, \mathcal{E}'_{j-1}, u'_j$ are tracks, we shall denote

$$TT' = u_1, \mathcal{E}_1, \dots, \mathcal{E}_{i-1}, u_i, \mathcal{E}'_1, \dots, \mathcal{E}'_{j-1}, u'_j.$$

provided that TT' is also a track (which happens when \mathcal{E}_{i-1} and \mathcal{E}'_1 do not both belong to $BS(u_i)$ or to $FS(u_i)$).

Similarly, for any tracks $T = u_1, \mathcal{E}_1, \dots, \mathcal{E}_{i-1}, u_i$ and $T' = u'_i, \mathcal{E}'_1, \dots, \mathcal{E}'_{j-1}, u'_j$ and hyperarc \mathcal{E} , we shall denote

$$T\mathcal{E}T' = u_1, \mathcal{E}_1, \dots, \mathcal{E}_{i-1}, u_i, \mathcal{E}, u'_i, \mathcal{E}'_1, \dots, \mathcal{E}'_{j-1}, u'_j$$

provided that $T\mathcal{E}T'$ is a track as well.

For any track $T = u_1, \mathcal{E}_1, \dots, \mathcal{E}_{i-1}, u_i$ its *reverse track* is $T^{-1} = u_i, \mathcal{E}_{i-1}, \dots, \mathcal{E}_1, u_1$. Clearly we have $(T_1T_2)^{-1} = T_2^{-1}T_1^{-1}$.

If tracks T_1 and T_2 have a common node v , we shall say that they *lead in the same direction* at v if $T'_1(T'_2)^{-1}$, where T'_1 and T'_2 denote the subtracks of T_1 and T_2 , respectively, ending at v , is not a track.

Any track $x, \mathcal{E}_1, \dots, \mathcal{E}_2, x$ such that either $\mathcal{E}_1, \mathcal{E}_2 \in \text{BS}(x)$ or $\mathcal{E}_1, \mathcal{E}_2 \in \text{FS}(x)$ will be called a *return twist*. Thus, if T is a return twist and $T'T$ is a track, then TT'^{-1} is also a track. If T_1 is a return twist and $T = T'T_1$ is a track, we shall say that T *ends with a return twist*. If T_1, T_2 are return twists and $T = T_1T'T_2$ is a track, we shall say that T *has return twists at both ends*.

A track is *simple* if every node and hyperarc appear only once in the sequence. A hypercycle is *simple* if it is a simple track or its first and last node are the only elements in the sequence that appear more than once. A return twist is *simple* if its first and last node and hyperarc are the only elements in the sequence that may appear more than once.

For any $t \in \mathcal{V}(G)$, by G_t^- (resp. G_t^+) we shall denote the subhypergraph, consisting of the nodes and hyperarcs lying on any track $T = t, \mathcal{E}, \dots$ such that $\mathcal{E} \in \text{BS}(t)$ (resp. $\mathcal{E} \in \text{FS}(t)$) and that T is simple or of the form $T = T_1T_2$, where T_1 is simple and T_2 is a simple return twist or hypercycle.

Consider the hypergraph in Figure 2.

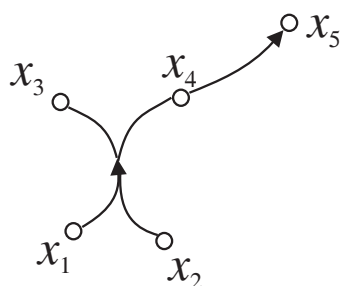


Figure 2.

Clearly, if the position of x_1 must be switched, then to get a B -graph the position of x_2 must not be switched, while the position of x_3 and x_4 must be switched. If the position of x_3 must not

be switched, then to get a B -graph the position of x_4 must be switched, while the position of x_1 and x_2 must not be switched.

It follows that, if a hypergraph is to be transformed into a B -graph, the need that a node is (resp. is not) switched, propagates along tracks: the need to be switched propagates along subtracks in the direction of hyperarcs, while the need not to be switched propagates along subtracks in the opposite direction of hyperarcs. Hence, the following applies:

Lemma 2. If every hyperarc on the track $T = x_1, \mathcal{E}_1, \dots, \mathcal{E}_2, x_2$ is to be turned into a B -arc, then if any of the following conditions is satisfied:

- $\mathcal{E}_1 \in \text{FS}(x_1)$ and the position of x_1 is switched,
- $\mathcal{E}_1 \in \text{BS}(x_1)$ and the position of x_1 is not switched,

then the following applies:

- if $\mathcal{E}_2 \in \text{BS}(x_2)$, then x_2 must be switched,
- if $\mathcal{E}_2 \in \text{FS}(x_2)$, then x_2 must not be switched.

Corollary 3. If there exists a return twist $T = x, \mathcal{E}_1, \dots, \mathcal{E}_2, x$, then — if all hyperarcs on T are to be turned into B -arcs — the position of x

- must be switched when $\mathcal{E}_1, \mathcal{E}_2 \in \text{BS}(x)$,
- must not be switched when $\mathcal{E}_1, \mathcal{E}_2 \in \text{FS}(x)$.

Proof. If $\mathcal{E}_1, \mathcal{E}_2 \in \text{BS}(x)$, then if x must not be switched, by Lemma 2, it must be switched. Thus it is impossible that x is not switched. The other assertion can be proven analogously. \square

Corollary 4. If every hyperarc on the track $T = x, \mathcal{E}, \dots$ which ends with a return twist is to be turned into a B -arc, then

- if $\mathcal{E} \in \text{BS}(x)$, then x must be switched,
- if $\mathcal{E} \in \text{FS}(x)$, then x must not be switched.

Proof. Let $T = T_1R$, where R is the return twist and let $T_1 = \dots, \mathcal{E}', x'$ and $R = x', \mathcal{E}_1, \dots, \mathcal{E}_2, x'$. If $\mathcal{E}_1, \mathcal{E}_2 \in \text{BS}(x')$, then by Corollary 4 x' must be switched, so that, as then $\mathcal{E}' \in \text{FS}(x')$, by Lemma 2 applied to T_1^{-1} , x must be switched as well. If $\mathcal{E}_1, \mathcal{E}_2 \in \text{FS}(x')$, then x' must not be switched, so that, as then $\mathcal{E}' \in \text{BS}(x')$, x must be switched again. The other assertion can be proved analogously. \square

Suppose now that in a hypergraph G there exists a track T with return twists at both ends — let

$T = R_1 T' R_2$, where R_1 and R_2 are return twists. Let v be a node on T' and denote by T_1 and T_2 the subtracks of T ending resp. beginning at v . Then T_1^{-1} and T_2 are tracks that end with a return twist and since their first hyperarcs belong one to $\text{FS}(v)$ and the other to $\text{BS}(v)$, by Corollary 4, if all hyperarcs on T are to be turned into B -arcs, v must be switched and must not be switched — which is a contradiction. Hence, if in a hypergraph there exists a track with return twists at both ends, by switching the position of its nodes such hypergraph cannot be turned into a B -graph.

Thus, the condition that the hypergraph does not contain tracks with return twists at both ends is a necessary condition for turning the hypergraph into a B -graph by switching the position of its nodes; we shall show that this condition is also sufficient.

Suppose that for some $u \in \mathcal{V}(G)$ there exists some $t \in \mathcal{V}(G)$ and tracks T_1, T_2 from u to t that do not lead in the same direction at t . Then, clearly, $T_1 T_2^{-1}$ is a track which is a return twist.

Thus, if in G_u^+ there are no tracks starting at u and ending with a return twist, for every $t \in \mathcal{V}(G_u^+)$ all tracks through t lead in the same direction. It follows that there is a well defined unary relation $S \subset \mathcal{V}(G_u^+)$: $t \in S$ if and only if in G_u^+ there exists a track u, \dots, \mathcal{E}, t , where $\mathcal{E} \in \text{BS}(t)$. Analogously, if in G_u^- there are no tracks starting at u and ending with a return twist, there is a well defined unary relation $S \subset \mathcal{V}(G_u^-)$: $t \in S$ if and only if in G_u^- there exists a track u, \dots, \mathcal{E}, t , where $\mathcal{E} \in \text{BS}(t)$.

Suppose now that in G there are no tracks with return twists at both ends. Let $R = x, \mathcal{E}_1, \dots, \mathcal{E}_2, x$ be a return twist and $\mathcal{E}_1, \mathcal{E}_2 \in \text{BS}(x)$. Then in G_x^+ there are no tracks starting at x and ending with a return twist so that there is a well defined unary relation $S \subset \mathcal{V}(G_x^+)$, $t \in S$ if and only if in G_x^+ there exists a track x, \dots, \mathcal{E}, t , where $\mathcal{E} \in \text{BS}(t)$. If G is to be turned into a B -graph, by Corollary 3, x must be switched, but then, by Lemma 2, every $t \in S$ must be switched. Suppose now that x and every $t \in S$ are indeed switched and let \mathcal{E} be any hyperarc in G_x^+ , let moreover $T = x, \dots, \mathcal{E}', v', \mathcal{E}, v$ be a track such that besides v and v' no other node in $H(\mathcal{E}) \cup T(\mathcal{E})$ lies on T . If $v' \in T(\mathcal{E})$, then $v' \in H(\mathcal{E}')$ so that $v' \in S$. Of course, for every $v \in T(\mathcal{E}) \setminus \{v'\}$ we have $v \notin S$ and for

every $v \in H(\mathcal{E})$ we have $v \in S$. Thus, after the switch, we have $H(\mathcal{E}) = \{v'\}$, so that \mathcal{E} becomes a B -arc. If $v' \in H(\mathcal{E})$, then $v' \in T(\mathcal{E}')$ so that $v' \notin S$ and $v \notin S$ for every $v \in T(\mathcal{E})$ and $v \in S$ for every $v \in H(\mathcal{E}) \setminus \{v'\}$, so that after the switch we also have $H(\mathcal{E}) = \{v'\}$ and \mathcal{E} becomes a B -arc again.

Similarly, if $R = x, \mathcal{E}_1, \dots, \mathcal{E}_2, x$ is a return twist and $\mathcal{E}_1, \mathcal{E}_2 \in \text{FS}(x)$, there is a well defined unary relation $S \subset \mathcal{V}(G_x^-)$, $t \in S$ if and only if in G_x^- there exists a track x, \dots, \mathcal{E}, t , where $\mathcal{E} \in \text{BS}(t)$, and if the position of the nodes in S is switched, every hyperarc in G_x^- is turned into a B -arc. For any return twist $R = x, \dots, x$ we shall denote by G_x^\pm the appropriate subhypergraph of the pair G_x^+, G_x^- (i.e., such that if $T = x, \dots$ is a track in G_x^\pm then RT is also a track).

When for return twists $R_1 = x_1, \dots, x_1$ and $R_2 = x_2, \dots, x_2$ hypergraphs $G_{x_1}^\pm$ and $G_{x_2}^\pm$ overlap, relations S_{x_1} and S_{x_2} agree on the intersection: if there exist tracks $T_1 = x_1, \dots, t$ in $G_{x_1}^\pm$ and $T_2 = x_2, \dots, t$ in $G_{x_2}^\pm$, they lead in the same direction at t (otherwise $T_1 T_2^{-1}$ would be a track and $R_1 T_1 T_2^{-1} R_2$ a track with return twists at both ends). Thus, relation S is defined on $\sum_{x \in R} G_x^\pm$, where R denotes the set of all nodes which are the starting nodes of a return twist in G .

Suppose now that $\mathcal{V}(\sum_{x \in R} G_x^\pm) \neq \mathcal{V}(G)$ and take any $t \in \mathcal{V}(G) \setminus \mathcal{V}(\sum_{x \in R} G_x^\pm)$. Since, for any $x \in R$, in G_x^\pm there is no track from x to t , in G_t^+ there are no tracks starting at t and ending with a return twist, so that S is well defined on G_t^+ . Moreover, for any $x \in R$, it agrees with S defined on G_x^\pm (any tracks $T_1 = x, \dots, v$ in G_x^\pm and $T_2 = t, \dots, v$ in G_t^+ lead in the same direction at v , for otherwise $T_1 T_2^{-1}$ would be a track from x to t , but by our assumption, $t \notin \mathcal{V}(G_x^\pm)$). If $\mathcal{V}(G_t^+ + \sum_{x \in R} G_x^\pm) \neq \mathcal{V}(G)$, the argument can be repeated, until finally $\mathcal{V}(\sum_x G_x^\pm) = \mathcal{V}(G)$ so that S is well defined on G . We know that when the position of the nodes in S is switched, every hyperarc in $\mathcal{A}(\sum_x G_x^\pm)$ turns into a B -arc. But this applies to other hyperarcs (if any) as well: if $\mathcal{E} \in \mathcal{A}(G) \setminus \mathcal{A}(\sum_x G_x^\pm)$, then we must have $H(\mathcal{E}) \subset S$ and $T(\mathcal{E}) \cap S = \emptyset$ (for any $v \in H(\mathcal{E})$ there exists in some G_x^\pm a track $T = x, \dots, \mathcal{E}_1, v$ and since $\mathcal{E}_1 \in \text{FS}(v)$ would imply $\mathcal{E} \in \mathcal{A}(G_x^\pm)$, we must have $\mathcal{E}_1 \in \text{BS}(v)$ and therefore $v \in S$; similarly one verifies that

An algorithm to this end can be based on Lemma 2 and Corollary 4, specifically, it can be designed as follows: it starts by tentatively labelling (while extending tracks in a depth-first manner) the nodes of the hypergraph according to Lemma 2 — label ‘*to be switched*’ is propagated in the direction of hyperarcs, while label ‘*not to be switched*’ is propagated in the opposite direction. This can proceed until a return twist is found; when this occurs, final labels are set (again while extending tracks in a depth-first manner) according to Corollary 4 — label ‘*must be switched*’ is propagated along the hyperarcs while label ‘*must not be switched*’ is propagated in the opposite direction. If another return twist is found in this phase the algorithm stops with the output ‘*a track with return twists at both ends*’, otherwise the set of the nodes labelled ‘*must be switched*’ or ‘*to be switched*’ is the desired S . As the label of any node can change only twice (when a tentative label is replaced by a fixed one), the propagation of labels through any node can occur only twice. Consequently, the time complexity of this algorithm is also $O(|G|)$. This algorithm is quite similar to the algorithm RHSat in [11], except that, of course, it operates in the corresponding hypergraph.

6. Conclusion

We have introduced the concept of a track and, based on it, found a necessary and sufficient condition that a hypergraph can be turned into a B -graph. This yields a characterization of Horn-renamable formulae. Furthermore, we have found a new class of formulae such that the restriction of SAT to it is solvable in polynomial time.

We also present two algorithms for testing whether a given hypergraph can be turned into a B -graph and (if so) for performing this change. As these algorithms have linear time complexity (which is comparable to the best known results), it seems that the hypergraph approach to SAT is worthy of further study.

References

- [1] E. BOROS, P. L. HAMMER, X. SUN, Recognition of q -Horn formulae in linear time. *Discrete Applied Mathematics*, 55, 1–13, 1994.
- [2] O. CEPEK, P. KUCERA, Known and new classes of generalized Horn formulae with polynomial recognition and SAT testing. *Discrete Applied Mathematics*, 149, 14–52, 2005.
- [3] V. CHANDRU, C. R. COULLARD, P. L. HAMMER, M. MONTANEZ, X. SUN, On renamable Horn and generalized Horn functions. *Annals of Math. and Art. Intel.*, 1, 33–47, 1990.
- [4] V. CHANDRU, J. N. HOOKER, Extended Horn sets in propositional logic. *Journal of the Association for Computing Machinery*, 38, 205–221, 1991.
- [5] W. F. DOWLING, J. H. GALLIER, Linear time algorithms for testing the satisfiability of Horn formulae. *J. Logic Programming*, 1, 207–284, 1984.
- [6] G. GALLO, G. LONGO, S. PALLOTTINO, S. NGUYEN, Directed Hypergraphs and Applications. *Discrete Applied Mathematics*, 42, 177–201, 1993.
- [7] G. GALLO, M. G. SCUTELLA, Polynomially solvable satisfiability problems. *Information Processing Letters*, 29, 221–227, 1988.
- [8] G. GALLO, C. GENTILE, D. PRETOLANI, G. RAGO, Max Horn SAT and the Minimum Cut Problem in Directed Hypergraphs. *Mathematical Programming*, 80, 213–237, 1998.
- [9] J. J. HEBRARD, A linear algorithm for renaming a set of clauses as a Horn set. *Theoretical Computer Science*, 124, 343–350, 1994.
- [10] H. R. LEWIS, Renaming a set of clauses as a Horn set. *Journal of the ACM*, 25, 134–135, 1978.
- [11] A. DEL VAL, On 2-SAT and Renamable Horn. *Proceedings of AAAI/IAAI 2000*, 279–284, 2000.

Received: June, 2008
Accepted: February, 2009

Contact address:

Dušan Hvalica
Faculty of Economics
University of Ljubljana
Kardeljeva ploščad 17
Si – 1000 Ljubljana
Slovenia

e-mail: dusan.hvalica@ef.uni-lj.si

DUŠAN HVALICA received his Ph. D. in mathematics in 1985. He is an Associate professor at the faculty of Economics, University of Ljubljana.
