The Complexity of Homomorphisms and Renamings for Minimal Unsatisfiable Formulas

Hans Kleine Büning and Daoyun Xu

Department of Computer Science, Universität Paderborn

33095 Paderborn (Germany), E-mail: kbcsl @upb.de

Abstract

We investigate the complexity of deciding whether for propositional CNF formulas F and H there exists a variable renaming, a literal renaming or a homomorphism ϕ such that $\phi(F) = H$. A variable renaming is a permutation of variables. A literal renaming is a permutation of variables which additionally replaces some of the variables by its complements. A homomorphism can be considered as a literal renaming which can map different literals to one literal.

1 Introduction

In [6], Krishnamurthy illustrated the power of symmetry for propositional proof systems. He added to the resolution calculus the rule of symmetry and gave short proofs for some hard formulas. For example, the pigeon hole formulas have in this extended calculus a proof of polynomial size. The rule of symmetry allows the following inference: If a clause f has been derived from a set of clauses F and φ is a permutation over the set of variables occurring in F, then the clause $\varphi(f)$ can be inferred as the next step in the derivation. Further interesting results can be found in Urquhart's paper [9]. We call a permutation of variables a variable renaming. Instead of a permutation of variables we can make use of a more general renaming, namely a so called literal renaming. That means we have a permutation of variables and additionally variables can be simultaneously replaced by its complements. More formally, for formulas F and H a *variable renaming* ϕ is a one–to–one mapping $\phi : var(F) \rightarrow var(H)$ and a *literal renaming* ϕ is a one–to–one mapping $\phi : lit(F) \rightarrow lit(H)$, where $\phi(\neg x) = \neg\phi(x)$ for any variable x. $var(F)$ resp. $lit(F)$ is the set of variables resp. literals of F.

The different renamings can be used for example in Davis–Putnam–Loveland–Logeman–algorithms (DPLL), too. Suppose in the splitting tree of the DPLL–algorithm two formulas are generated at some nodes for which one of the formulas can be mapped to the other one by a renaming. Then we can replace one of the formulas by the empty clause and continue with the remaining formula. It can easily be seen that DPLL-algorithm extended by such a symmetry rule have short proofs for the pigeon hole formulas. We need not more than a quadratic number of nodes.

Szeider [8] introduced an interesting tool for proving the unsatisfiability of formulas, namely homomorphisms. A mapping $\phi : lit(F) \rightarrow lit(H)$ is a *homomorphism* from F to H if $\phi(\neg x) = \neg\phi(x)$ for every variable x, and $\phi(F) = \{h | \exists f \in F : h = \phi(f)\} \subseteq H$, where $\phi(f) = \{\phi(L_1), \cdots, \phi(L_m)\}$ for a clause $f = \{L_1, \cdots, L_m\}$. Besides some relationships to tree resolution refutations, he investigated the core of formulas. It has been shown that the cores of a formulas are mutually isomorphic and the recognition of cores is *co*-NP-complete.

Let $MU(k)$ for $k \geq 1$ be the set of minimal unsatisfiable formulas with *deficiency* k. That means, for a formula the difference between the number of clauses and the number of variables is k. Whether or not a formula belongs to $MU(k)$ can be decided in polynomial time [2]. It has been proved in [4] that for any $k, t \geq 1$ and any formula $F \in MU(t)$, there exists a formula H in $MU(k)$ and a homomorphism ϕ from H to F such that $\phi(H) = F$. Moreover for fixed $k, t \geq 1$ the formula H and the homomorphism can be constructed in polynomial time. For a class $\mathcal{C}$ of CNF formulas we consider the problems:

Problem: $\mathcal{VR}$–$\mathcal{C}$, ($\mathcal{LR}$–$\mathcal{C}$, $\mathcal{HOM}$–$\mathcal{C}$)

Instance: $F_1, F_2 \in \mathcal{C}$

Query: $\exists\ \phi$ variable renaming (literal renaming, homomorphism): $\phi(F_1) = F_2$.

We call the problem $\mathcal{VR}$–$\mathcal{C}$ resp. $\mathcal{LR}$–$\mathcal{C}$ and $\mathcal{HOM}$–$\mathcal{C}$ the variable renaming problem resp. literal renaming problem and homomorphism problem for $\mathcal{C}$.

We investigate the above mentioned problems for the class of minimal unsatisfiable formulas and various natural subclasses. The classes considered first are minimal unsatisfiable Horn formulas $Horn{-}MU$ and MU formulas with fixed deficiency k. Additionally, for the homomorphism problem we consider the class $MU(k,t)$. The class $MU(k,t)$ is the set of pairs of formulas (F_1, F_2) where $F_1 \in MU(k)$ and $F_2 \in MU(t)$. Since a renaming preserves the number of clauses, these problems are not of interest for $MU(k,t)$. For fixed k and t the problem $\mathcal{HOM}{-}M(k,t)$ has as instances pairs of formulas $F_1 \in MU(k)$ and $F_2 \in MU(t)$. The question is whether there is a homomorphism ϕ such that $\phi(F_1) = F_2$.

GI is the graph isomorphism problem for undirected graphs which is known to be in NP. But it is an open problem whether GI is NP–complete [5]. NPC is an abbreviation for NP–complete. P stands for solvable in polynomial time.

In order to see whether the problems will be easier for more restrictive classes we investigate maximal and marginal formulas. A MU formula F is *maximal* if adding a literal to any clause of F, which is new for this clause, results in a satisfiable formula. MAX is the set of maximal MU formulas and $MAX(k) = MU(k) \cap MAX$.

A MU formula F is *marginal* if removing an occurrence of a literal from F results a non–minimal unsatisfiable formula. $MARG$ is the set of marginal formulas and $MARG(k) = MU(k) \cap MARG$.

MAX and $MARG$ are known to be D^P-complete, whereas the problems for fixed k are in P because $MU(k)$ is solvable in polynomial time. In this paper, we will show that for $MAX(1)$ and $MARG(1)$ in contrast to $MU(1)$ the renaming problems are solvable in polynomial time. For $MAX(k)$ $(k \geq 2)$ and the $MARG(k)$ $(k \geq 1)$ it is open whether the homomorphism problems are NP–complete, as hard as the graph isomorphism problem, or solvable in polynomial time. For $MAX(k)$ and $MARG(k)$ $(k \geq 2)$ the variable and literal renaming problems are equivalent to a restricted graph isomorphism problem denoted as $GI(n,m)$, where n and m depend on the deficiency k. The exact relationship will be presented in a forthcoming paper.

For $\mathcal{C}= MU$, MAX, or $MARG$ the problems $\mathcal{VR}{-}\mathcal{C}$, $\mathcal{LR}{-}\mathcal{C}$ and $\mathcal{HOM}{-}\mathcal{C}$ are so called promise problems, because MU, MAX, and $MARG$ are D^P–complete. We write instead of GI and NPC the terms GI^* and NPC^* in order to make clear that given MU formulas resp. MAX or $MARG$ formulas as instances the remaining problem is as hard as GI resp. as an NP–complete problem.

An overview of the results can be found in following table. The proofs of the complexities given in boldface will be presented in the paper or in the appendix.

classes	*variable renaming*	*literal renaming*	*homomorphism*
MU	**GI***	**GI***	**NPC***
$HORN-MU$	**GI**	**GI**	**NPC**
$MU(k)$	**GI**	**GI**	**NPC**
$MU(k,t)$	–	–	**NPC**
MAX	*GI**	*GI**	*Open*
$MAX(1)$	**P**	**P**	**P**
$MAX(k)$	$GI(n,m)$	$GI(n,m)$	*Open*
$MARG$	*GI**	*GI**	*Open*
$MARG(1)$	**P**	**P**	*Open*
$MARG(k)$	$GI(n,m)$	$GI(n,m)$	*Open*

A literal is a propositional variable or a negated propositional variable. A clause C is a disjunction of literals, $C = (L_1 \vee \cdots \vee L_m)$, and sometimes written as a set of literals, $C = \{L_1, \cdots, L_m\}$. A formula F in conjunction normal form (CNF) is a conjunction of clauses, $F = (C_1 \wedge \cdots \wedge C_n)$, and sometimes we write F as a set of clauses, $F = \{C_1, \cdots, C_n\}$. $var(F)$ is the set of variables occurring in the formula F and $lit(F)$ is the set of literals over the variables of F. We write $A \leq_p B$, if the class B is polynomially many one reducible to the class A. $A \equiv_p B$ is an abbreviation for $A \leq_p B$ and $B \leq_p A$.

2 Minimal Unsatisfiable Formulas

At first we show that the variable and literal renaming problems $\mathcal{VR}{-}HORN{-}MU$, $\mathcal{LR}{-}HORN{-}MU$, $\mathcal{VR}{-}MU(k)$, and $\mathcal{LR}{-}MU(k)$ are equivalent to the graph isomorphism problem GI. Further, we show that the homomorphism problem $\mathcal{HOM}{-}Horn{-}MU$ and for fixed k and t the problem $\mathcal{HOM}{-}MU(k,t)$ are NP–complete. These results imply that the variable and literal renaming resp. homomorphism problem for MU formulas as instances are as hard as the graph isomorphism problem resp. a NP-complete problem.

Lemma 1 *$GI \leq_p \mathcal{VR}{-}HORN{-}MU$ and $GI \leq_p \mathcal{LR}{-}HORN{-}MU$*

Proof: Let $G = (V, E)$ be an undirected graph without self-loops, where $V = \{x_1, \cdots, x_n\}$ and $E = \{(x_{i_1}, x_{j_1}), \cdots, (x_{i_m}, x_{j_m})\}$ by some order of edges. We view vertex x_i as a variable and transform an edge (x_{i_p}, x_{j_p}) into a clause $\neg x_{i_p} \vee \neg x_{j_p}$. For new variables $z_1, \cdots, z_m$ we associate to G the formula F_G as follows: $F_G = [(\neg z_1 \vee \cdots \vee \neg z_m \vee \neg x_1 \vee \cdots \vee \neg x_n), (z_1 \vee \neg x_{i_1} \vee \neg x_{j_1}), \cdots, (z_m \vee \neg x_{i_m} \vee \neg x_{j_m}), x_1, \cdots, x_n]$. The formula F_G is a minimal unsatisfiable Horn formula over $n + m$ variables and consists of $n + m + 1$ clauses. We immediately obtain for two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ that G_1 is isomorphic to G_2 iff there is a variable renaming φ with $\varphi(F_{G_1}) = F_{G_2}$. The variable renaming is also a literal renaming φ with $\varphi(F_{G_1}) = F_{G_2}$. ■

For the upper bound we can show the following theorems. The proof can be found in the appendix.

Theorem 1 *We can associate in polynomial time to any MU–formula F a directed graph G_F, such that for every pair F_1 and F_2 of MU–formulas we have: There exists a variable renaming ϕ with $\phi(F_1) = F_2$ if and only if G_{F_1} is isomorphic to G_{F_2}.*

Theorem 2 *We can associate in polynomial time to any MU–formula F an undirected graph G_F, such that for every pair F_1 and F_2 of MU–formulas we have: There exists a literal renaming ϕ with $\phi(F_1) = F_2$ if and only if G_{F_1} is isomorphic to G_{F_2}.*

Let F_1 and F_2 be formulas in $MU(1)$. We can easily extend F_1 and F_2 to F_1^* and F_2^* in $MU(k)$, respectively, such that: There exists a variable (literal) renaming ϕ with $\phi(F_1) = F_2$ if and only if there exists a variable (literal) renaming ϕ' with $\phi'(F_1^*) = F_2^*$. Since every minimal unsatisfiable Horn formula is in $MU(1)$, by Lemma 1, Theorem 1, and Theorem 2 we obtain our desired result:
$\mathcal{VR}{-}HORN{-}MU \equiv_p \mathcal{LR}{-}HORN{-}MU \equiv_p \mathcal{VR}{-}MU(k) \equiv_p \mathcal{LR}{-}MU(k) \equiv_p GI.$

A graph $G = (V, E)$ is homomorphic to a graph $G' = (V', E')$ if there is a mapping $\varphi : V \to V'$, such that $V' = \varphi(V)$ and $E' = \{(\varphi(u), \varphi(v)) \mid (u, v) \in E\}$, that means $\varphi(G) = G'$. For graphs G and H, we say that G has a H-coloring if there exists a homomorphism φ with $\varphi(G) = H$. The H-coloring problem has as instances a graph G and the question is whether there exists a homomorphism φ with $\varphi(G) = H$. In [3], Hell and Nešetřil have shown that if H is not bipartite then the H-coloring problem is

NP-complete. Let K_3 be the complete graph with 3 vertices. It can be shown by means of the NP-completeness of the K_3-coloring problem that the homomorphism problem for $MU(1)$ formulas and even for Horn formulas is NP-complete. Based on this result, it can be shown that for fixed $k, t \geq 1$ the problem whether for a $MU(k)$ formula F and a $MU(t)$ formula H there is a homomorphism ϕ with $\phi(F) = H$ is NP-complete.

Theorem 3 *The problem $\mathcal{HOM}-Horn-MU$ is NP–complete.*

Theorem 4 *For fixed $k, t \geq 1$ the problem $\mathcal{HOM}-MU(k,t)$ is NP–complete.*

Detailed proofs can be found in the appendix.

As mentioned above we only present the proofs for maximal and marginal formulas with $n+1$ clauses, where n is number of variables. In contrast to the results for $MU(1)$ we will show that the variable and literal renaming problems for $MAX(1)$ and $MARG(1)$ are solvable in polynomial time. Moreover, we will see that the homomorphism problem for $MAX(1)$ is as complex as the literal renaming problem for this class and therefore solvable in polynomial time, too.

Let F be a formula over n variables in $MAX(1)$. Based on the well-known structure of $MU(1)$ formulas [1] it is quite easy to prove that there is a unique variable $x \in var(F)$ which occurs in every clause. The number of clauses with an occurence of the literal x resp. $\neg x$ is denoted as $pos(x, F)$ resp. $neg(x, F)$. We call the variable x the axial variable of F. For the axial variable x the formula F can be written as $F = [(x \vee f_1), \cdots, (x \vee f_p), (\neg x \vee g_1), \cdots, (\neg x \vee g_q)]$, where $F|_x := \{f_1, \cdots, f_p\}$, $F|_{\neg x} := \{g_1, \cdots, g_q\}$ are maximal and $var(F|_x) \cap var(F|_{\neg x}) = \emptyset$.

Let F_1 and F_2 be formulas over n variables in $MAX(1)$ with axial variables a_1 and a_2. Then we have:
1) There exists a variable renaming φ with $\varphi(F_1) = F_2$ if and only if $\varphi(a_1) = a_2$, $pos(a_1, F_1) = pos(a_2, F_2)$, and there are two variable renamings φ_+ and φ_- with $\varphi_+(F_1|_{a_1}) = F_2|_{a_2}$ and $\varphi_-(F_1|_{\neg a_1}) = F_2|_{\neg a_2}$.
2) There exists a literal renaming φ with $\varphi(F_1) = F_2$ if and only if $pos(a_1, F_1) = pos(a_2, F_2)$ or $pos(a_1, F_1) = neg(a_2, F_2)$ and there are two literal renamings φ_+ and φ_- such that at least one of the following cases holds: (1) $\varphi_+(F_1|_{a_1}) = F_2|_{a_2}$ and $\varphi_-(F_1|_{\neg a_1}) = F_2|_{\neg a_2}$, or (2) $\varphi_+(F_1|_{a_1}) = F_2|_{\neg a_2}$ and $\varphi_-(F_1|_{\neg a_1}) = F_2|_{a_2}$.
From these observations we immediately obtain

Theorem 5 *The problems $\mathcal{VR}-MAX(1)$ and $\mathcal{LR}-MAX(1)$ are in P.*

For $MAX(1)$ formulas F_1 and F_2 any homomorphism ϕ with $\phi(F_1) = F_2$ must be a literal renaming. That can be seen as follows: Suppose there is some homomorphism which is not a literal renaming. Then there exists variables x and y, such that $\phi(x) = \phi(y)$ and at least two clauses $f_1 \neq f_2$ in F_1 will be identified. That means $\phi(f_1) = \phi(f_2)$. But for $MAX(1)$ formulas by means of the axial variable it can easily be seen that two different clauses have complementary literals, e.g. $L \in f_1$ and $\neg L \in f_2$. Because of Theorem 5 we obtain

Corollary 1 *The homomorphism problem $\mathcal{HOM}-MAX(1)$ is in P.*

Now we deal with marginal formulas. Using the well–known structure of $MU(1)$ formulas again, the following proposition can be shown easily for a $MARG(1)$ formula F: F contains at least two unit clauses and every variable in F occurs exactly once positively and once negatively. Further, if $(L_1 \vee L_2)$ is a subclause of a clause of F for distinct literals L_1 and L_2, then $(\neg L_1 \vee \neg L_2)$ is not a subclause of a clause of F.

Based on this observation, we can prove the following lemma. The proof can be found in the appendix.

Lemma 2 *Let F_1 and F_2 be formulas in $MARG(1)$. For the associated trees T_1 and T_2 we have: There exists a literal renaming φ with $\varphi(F_1) = F_2$ if and only if T_1 is isomorphic to T_2.*

Since the tree isomorphism problem is solvable in polynomial time, we obtain

Theorem 6 *$\mathcal{LR}-MARG(1)$ is decidable in polynomial time.*

For the variable renaming problem we have to modify our associated graph, because we have to distinguish between positive and negative occurrences.

Similarly, we can associate trees and show the following lemma. (The construction and the proof are the appendix)

Lemma 3 *Let F_1 and F_2 be formulas in $MARG(1)$. For the associated trees T_1^* and T_2^* we have: There exists a variable renaming φ with $\varphi(F_1) = F_2$ if and only if T_1^* is isomorphic to T_2^*.*

Theorem 7 *The problem $\mathcal{VR}-MARG(1)$ is decidable in polynomial time.*

3 Future Work

For fixed deficiency k the complexity of the homomorphism problem remains open for maximal formulas with deficiency $k \geq 2$ and for marginal formulas with deficiency $k \geq 1$.

References

[1] G. Davydov, I. Davydova, H. Kleine Büning, An efficient algorithm for the minimal unsatisfiability Problem for a Subclass of CNF, *Annals of Mathematics and Artificial Intelligence*, **23** (1998), pp. 229-245.

[2] H. Fleischner, O. Kullmann, S. Szeider, Polynomial-time recognition of minimal unsatisfiable formulas with fixed clause-variable difference, to appear in *Theoretical Computer Science.*

[3] P. Hell and J. Nešetřil, On the complexity of H-Coloring, *Journal of Combinatorial Theory, Series B*, **48** (1990), pp. 92-110.

[4] H. Kleine Büning, Xishun Zhao, Polynomial time algorithms for computing a representation for minimal unsatisfiable formulas with fixed deficiency, appears in *Information Processing Letters.*

[5] J. Köbler, U. Schöning, J. Toran, The graph isomorphism problem: its structural complexity, *Birkhäuser Verlag*, 1993.

[6] B. Krishnamurthy, Short proofs for tricky formulas, *Acta Informatica*, **22** (1985), pp. 253-275.

[7] C. H. Papadimitriou, D. Wolfe, The complexity of facets resolved, *Journal of Computer and System Sciences*, **37** (1988), pp. 2-13.

[8] S. Szeider, Homomorphisms of conjunctive normal forms, to appear in *Discrete Applied Mathematics.*

[9] A. Urquhart, The symmetry rule in propositional logic, *Discrete Applied Mathematics* **96-97** (1999), pp. 177–193.

4 Appendix

Please note that the isomorphism problem for undirected graphs is as hard as the problem for directed graphs.

Theorem 1 *We can associate in polynomial time to any MU–formula F a directed graph G_F, such that for every pair F_1 and F_2 of MU–formulas we have: There exists a variable renaming ϕ with $\phi(F_1) = F_2$ if and only if G_{F_1} is isomorphic to G_{F_2}.*

Proof: Let $F = \{C_1, \cdots, C_m\}$ be a formula over the variables $x_1, \cdots, x_n$ where the clauses are given in an arbitrary but fixed order. We associate to F the directed graph $G_F = (V_F, E_F)$ as follows:
(1) $V_F = V_{var} \cup V_{cl} \cup V_{2deg} \cup V_{1deg}$, where $V_{var} = \{x_1, \cdots, x_n\}$ (corresponding to variables in F), $V_{cl} = \{c_1, \cdots, c_m\}$ (corresponding to clauses in F), $V_{2deg} = \{t_1^1, t_1^2, \cdots, t_n^1, t_n^2\}$, and $V_{1deg} = \{s_1^1, s_1^2, s_1^3, \cdots, s_m^1, s_m^2, s_m^3\}$.
(2) $E_F = E_0 \cup E_1 \cup E_2$, where
$E_0 = \{(c_i, x_k), (x_k, c_j) |\ x_k \in C_i \text{ and } \neg x_k \in C_j \text{ for } 1 \leq k \leq n \text{ and } 1 \leq i, j \leq m\}$,
$E_1 = \{(c_i, s_i^1), (c_i, s_i^2), (c_i, s_i^3) | 1 \leq i \leq m\}$, and $E_2 = \{(x_k, t_k^1), (t_k^1, t_k^2), (t_k^2, x_k) | 1 \leq k \leq n\}$.
Please note that we construct a $'triangle'$ cycle at each vertex x_k and a $'star'$ at each vertex c_i to distinguish x_k and c_i.
Then we have:

(1) For any $t \in V_{2deg}$, $deg_{in}(t) = deg_{out}(t) = 1$, where $deg_{in}(t)$ denotes the indegree of t and $deg_{out}(t)$ the outdegree of t.
(2) For any $s \in V_{1deg}$, $deg_{in}(s) = 1$ and $deg_{out}(s) = 0$.
(3) For any $c \in V_{cl}$, $deg_{out}(c) \geq 3$.
(4) For any $x \in V_{var}$, $deg_{in}(x) \geq 2$ and $deg_{out}(x) \geq 2$.
(5) For any $c \in V_{cl}$, there are three vertices, $s, s', s'' \in V_{1deg}$, such that $(c, s), (c, s'), (c, s'') \in E_F$.
(6) For any $x \in V_{var}$, there are four vertices, $c, c' \in V_{cl}$ and $t, t' \in V_{2deg}$, such that $(c, x), (x, c'), (x, t), (t, t'), (t', x) \in E_F$.

W.l.o.g., let F_1 and F_2 be MU-formulas over variables $x_1, \cdots, x_n$, $F_1 = \{C_1, \cdots, C_m\}$ and $F_2 = \{C_1', \cdots, C_m'\}$. Let G_{F_1} and G_{F_2} be the associated directed graphs of F_1 and

F_2, respectively, where $G_{F_k} = (V_k, E_k)$ and $V_k = V_{var} \cup V_{cl}^{(k)} \cup V_{2deg} \cup V_{1deg}$ $(k = 1, 2)$. $(V_{cl}^{(1)} = \{c_1, \cdots, c_m\}$ and $V_{cl}^{(2)} = \{c'_1, \cdots, c'_m\}$)

Then for any isomorphism φ from G_{F_1} to G_{F_2}, we have

$$\varphi(V_{1deg}) = V_{1deg},\ \varphi(V_{2deg}) = V_{2deg},\ \varphi(V_{var}) = V_{var} \text{ and } \varphi(V_{cl}^{(1)}) = V_{cl}^{(2)}.$$

We now show: There exists a variable renaming φ from F_1 to F_2 such that $\varphi(F_1) = F_2$ if and only if G_{F_1} is isomorphic to G_{F_2}.

If there exists a variable renaming φ with $\varphi(F_1) = F_2$, we have a permutation π_v over $\{1, \cdots, n\}$ and a permutation π_c over $\{1, \cdots, m\}$ such that $\varphi(x_k) = x_{\pi_v(k)}$ for $1 \le k \le n$ and $\varphi(C_i) = C'_{\pi_c(i)}$ for $1 \le i \le m$.

We now define an isomorphism ψ from G_{F_1} to G_{F_2} as follows:

$\psi(x_k) = x_{\pi_v(k)}$, $\psi(t_k^p) = t_{\pi_v(k)}^p$ for $(1 \le k \le n)$ and $p = 1, 2$

$\psi(c_i) = c'_{\pi_c(i)}$, $\psi(s_i^q) = s_{\pi_c(i)}^q$ for $(1 \le i \le m)$ and $q = 1, 2, 3$.

Conversely, if there exists an isomorphism ψ from G_{F_1} and G_{F_2} then we have $\psi(V_{var}) = V_{var}$. Thus, the restriction $\psi|_{V_{var}}$ is the desired variable renaming, since $x_k \in C_i$ and $\neg x_k \in C_j$ if and only if G_{F_1} contains the edges (c_i, x_k) and (x_k, c_j). ∎

Theorem 2 *We can associate in polynomial time to any MU–formula F an undirected graph G_F, such that for every pair F_1 and F_2 of MU–formulas we have: There exists a literal renaming ϕ with $\phi(F_1) = F_2$ if and only if G_{F_1} is isomorphic to G_{F_2}.*

Proof: Let $F = \{C_1, \cdots, C_m\}$ be a formula over the variables $x_1, \cdots, x_n$ where the clauses are given in a fixed order. We associate to F the directed graph $G_F = (V_F, E_F)$ as follows:

(1) $V_F = V_{var} \cup V_{var}^+ \cup V_{var}^- \cup V_{cl} \cup V_{2deg} \cup V_{1deg}$, where
   $V_{var} = \{x_1, \cdots, x_n\}$ (corresponding to variables in F),
   $V_{var}^+ = \{x_1^+, \cdots, x_n^+\}$ (corresponding to positive literals in F),
   $V_{var}^- = \{x_1^-, \cdots, x_n^-\}$ (corresponding to negative literals in F),
   $V_{cl} = \{c_1, \cdots, c_m\}$ (corresponding to clauses in F),
   $V_{2deg} = \{t_1^1, t_1^2, \cdots, t_n^1, t_n^2\}$ and
   $V_{1deg} = \{s_1^1, s_1^2, s_1^3, \cdots, s_m^1, s_m^2, s_m^3\}$.

(2) $E_F = E_0 \cup E_1 \cup E_2 \cup E_3$, where
   $E_0 = \{(c_i, x_k^+), (x_k^-, c_j) |\ x_k \in C_i \text{ and } \neg x_k \in C_j \text{ for } 1 \le k \le n \text{ and } 1 \le i, j \le m\}$,
   $E_1 = \{(c_i, s_i^1), (c_i, s_i^2), (c_i, s_i^3) | 1 \le i \le m\}$,

$E_2 = \{(x_k, t_k^1), (t_k^1, t_k^2), (t_k^2, x_k) | 1 \le k \le n\}$and
$E_3 = \{(x_k^+, x_k), (x_k, x_k^-) | 1 \le k \le n\}$.

Then we have:

(1) For any $t \in V_{2deg}$, $deg(t) = 2$.

(2) For any $s \in V_{1deg}$, $deg(s) = 1$.

(3) For any $c \in V_{cl}$, $deg(c) \ge 4$.

(4) For any $x \in V_{var}$, $deg(x) = 4$.

(5) For any $x^+ \in V_{var}^+$, $deg(x^+) = pos(x, F) + 1$.

(6) For any $x^- \in V_{var}^-$, $deg(x^-) = neg(x, F) + 1$.

(7) For any $c \in V_{cl}$, there are three vertices, $s, s', s'' \in V_{1deg}$, such that $(c, s), (c, s')$, $(c, s'') \in E_F$.

(8) For any $x \in V_{var}$, there are four vertices, $x^+ \in V_{var}^+, x^- \in V_{var}^-$ and $t, t' \in V_{2deg}$, such that $(x^+, x), (x, x^-), (x, t), (t, t'), (t', x) \in E_F$.

W.l.o.g., let F_1 and F_2 be MU-formulas over variables $x_1, \cdots, x_n$, $F_1 = \{C_1, \cdots, C_m\}$, and $F_2 = \{C_1', \cdots, C_m'\}$. Let G_{F_1} and G_{F_2} be the associated directed graphs of F_1 and F_2, respectively, where $G_{F_k} = (V_k, E_k)$ and $V_k = V_{var} \cup V_{var}^+ \cup V_{var}^- \cup V_{cl}^{(k)} \cup V_{2deg} \cup V_{1deg}$ $(k = 1, 2)$. ($V_{cl}^{(1)} = \{c_1, \cdots, c_m\}$ and $V_{cl}^{(2)} = \{c_1', \cdots, c_m'\}$)

Then for any isomorphism φ from G_{F_1} to G_{F_2}, we have

$\varphi(V_{1deg}) = V_{1deg}$, $\varphi(V_{2deg}) = V_{2deg}$, $\varphi(V_{var}) = V_{var}$,
$\varphi(V_{cl}^{(1)}) = V_{cl}^{(2)}$, and $\varphi(V_{var}^+ \cup V_{var}^-) = V_{var}^+ \cup V_{var}^-$.

We now show: There exists a literal renaming φ from F_1 to F_2 such that $\varphi(F_1) = F_2$ if and only if G_{F_1} is isomorphic to G_{F_2}.

If there exists a literal renaming φ with $\varphi(F_1) = F_2$, we have a permutation π_v over $\{1, \cdots, n\}$, a permutation π_c over $\{1, \cdots, m\}$ and a renaming f over $\{x_1, \cdots, x_n\}$ such that $\varphi(x_k) = f(x_{\pi_v(k)})$ for $1 \le k \le n$ and $\varphi(C_i) = C'_{\pi_c(i)}$ for $1 \le i \le m$.

We now define an isomorphism ψ from G_{F_1} to G_{F_2} as follows:

$\psi(x_k) = x_{\pi_v(k)}$, $\psi(t_k^p) = t_{\pi_v(k)}^p$ for $(1 \le k \le n)$ and $p = 1, 2$

$\psi(c_i) = c'_{\pi_c(i)}$, $\psi(s_i^q) = s_{\pi_c(i)}^q$ for $(1 \le i \le m)$ and $q = 1, 2, 3$

$\psi(x_k^+) = x_{\pi_v(k)}^+$ and $\psi(x_k^-) = x_{\pi_v(k)}^-$ if $f(x_{\pi_v(k)}) = x_{\pi_v(k)}$ for $1 \le k \le n$

$\psi(x_k^+) = x_{\pi_v(k)}^-$ and $\psi(x_k^-) = x_{\pi_v(k)}^+$ if $f(x_{\pi_v(k)}) = \neg x_{\pi_v(k)}$ for $1 \le k \le n$.

Conversely, if there exists an isomorphism ψ from G_{F_1} and G_{F_2} then we have two permutations from ψ: π_v over $\{1, \cdots, n\}$ and π_c over $\{1, \cdots, m\}$, satisfying $\psi(x_k) = x_{\pi_v(k)}$ and $c'_{\pi_c(i)} = \psi(c_i)$. Further, we can require $\psi(x_k^+) = x_{\pi_v(k)}^+$ or $x_{\pi_v(k)}^-$ (up to exchanging new auxiliary variables).

Let $\pi = \psi|_{V_{var}}$, which is the restriction of ψ over V_{var}. Clearly, π is a variable renaming over F_1.

Define a formula $F_2^* = \{\pi(C_1), \cdots, \pi(C_m)\}$ from F_1 by π. By the construction of G_{F_2} and the above observations, we have a renaming φ^* over F_2^* such that for every $(1 \leq i \leq m)$ $\varphi^*(\pi(C_i)) = C'_{\pi_c(i)}$. Finally, we define a literal renaming $\varphi = \pi \circ \varphi^*$. It is the desired literal renaming. ■

Theorem 3 *$\mathcal{HOM}$−Horn−MU is NP-complete.*

Proof: Obviously, the problem is in NP. The NP–hardness will be shown by means of a reduction from the NP–complete K_3–coloring problem for graphs.

Let $G = (V, E)$ be an undirected connected graph without self-loops, where $V = \{x_1, \cdots, x_n\}$ and $E = \{(x_{i_1}, x_{j_1}), \cdots, (x_{i_m}, x_{j_m})\}$. Let $K_3 = (V_0, E_0)$ be a complete graph with three vertices, where $V_0 = \{a, b, c\}$, $E_0 = \{(a, b), (a, c), (b, c)\}$, and $\{a, b, c\} \cap \{x_1, \cdots, x_n\} = \emptyset$. Further, let $G^+ = G + K_3$. Clearly, G is K_3-colorable if and only if the vertices of G^+ can be colored by exactly three different colors. In the next step we associate to the graph G^+ two minimal unsatisfiable Horn formulas F' and F'_0.

We transform an edge (x_i, x_j) into a clause $(\neg x_i \vee \neg x_j)$ and introduce $m + 3$ new variables $y_1, y_2, y_3, z_1, \cdots, z_m$:

$F'_0 = [(\neg y_1 \vee \neg y_2 \vee \neg y_3 \vee \neg a \vee \neg b \vee \neg c), (y_1 \vee \neg a \vee \neg b), (y_2 \vee \neg a \vee \neg c), (y_3 \vee \neg b \vee \neg c), a, b, c]$

$F' = [(\neg y_1 \vee \neg y_2 \vee \neg y_3 \vee \neg z_1 \vee \cdots \vee \neg z_m \vee \neg a \vee \neg b \vee \neg c \vee \neg x_1 \vee \cdots \vee \neg x_n),$
$\quad (y_1 \vee \neg a \vee \neg b), (y_2 \vee \neg a \vee \neg c), (y_3 \vee \neg b \vee \neg c),$
$\quad (z_1 \vee \neg x_{i_1} \vee \neg x_{j_1}), \cdots, (z_m \vee \neg x_{i_m} \vee \neg x_{j_m}),$
$\quad a, b, c, x_1, \cdots, x_n].$

For sake of readability, we replace every variable in F'_0 and F' by its complement. Then, the resulting formulas are

$F_0 = [(y_1 \vee y_2 \vee y_3 \vee a \vee b \vee c), (\neg y_1 \vee a \vee b), (\neg y_2 \vee a \vee c), (\neg y_3 \vee b \vee c), \neg a, \neg b, \neg c]$

$F = [(y_1 \vee y_2 \vee y_3 \vee z_1 \vee \cdots \vee z_m \vee a \vee b \vee c \vee x_1 \vee \cdots \vee x_n),$
$\quad (\neg y_1 \vee a \vee b), (\neg y_2 \vee a \vee c), (\neg y_3 \vee b \vee c),$
$\quad (\neg z_1 \vee x_{i_1} \vee x_{j_1}), \cdots, (\neg z_m \vee x_{i_m} \vee x_{j_m}),$
$\quad \neg a, \neg b, \neg c, \neg x_1, \cdots, \neg x_n].$

Clearly, F' is homomorphic to F'_0 if and only if F is homomorphic to F_0.

Now we show that G^+ can be colored by exactly three different colors if and only if there exists a homomorphism ϕ with $\phi(F) = F_0$.
($\Rightarrow$) Assume that G^+ can be colored by exactly three different colors. Then the vertices a, b, and c must be colored by different colors. Suppose that the vertex a (reps. b and c) is colored by the color *a_color* (resp. *b_color* and *c_color*). The set $V = \{x_1, \cdots, x_n\}$ can be divided into three disjoint subsets as follows:

$$V_a = \{x \in V | \text{ the vertex } x \text{ is colored by the color } a\_color\},$$
$$V_b = \{x \in V | \text{ the vertex } x \text{ is colored by the color } b\_color\},$$
$$V_c = \{x \in V | \text{ the vertex } x \text{ is colored by the color } c\_color\}.$$

For every $1 \leq k \leq m$, the vertices x_{i_k} and x_{j_k} must be colored by the different colors since $(x_{i_k}, x_{j_k}) \in E$. In other words, the vertices x_{i_k} and x_{j_k} are colored by one of the groups $\{a\_color, b\_color\}, \{a\_color, c\_color\}, \{b\_color, c\_color\}$.
We define a mapping for the x-variables $\phi_X : \{x_1, \cdots, x_n\} \to \{a, b, c\}$ as follows:

$$\phi_X(x) = \begin{cases} a & \text{if } x \in V_a \\ b & \text{if } x \in V_b \\ c & \text{if } x \in V_c \end{cases}.$$

Please note the edge (x_{i_k}, x_{j_k}) corresponds to a 2-CNF clause $(x_{i_k} \vee x_{j_k})$ $(1 \leq k \leq m)$. By the mapping ϕ_X, the index set $\{1, \cdots, m\}$ of clauses set $\{(x_{i_1} \vee x_{j_1}), \cdots, (x_{i_m} \vee x_{j_m})\}$ will be divided into the three disjoint subsets as follows:

$$I_{ab} = \{k | \ \phi_X(\{x_{i_k}, x_{j_k}\}) = \{a, b\}\},$$
$$I_{ac} = \{k | \ \phi_X(\{x_{i_k}, x_{j_k}\}) = \{a, c\}\},$$
$$I_{bc} = \{k | \ \phi_X(\{x_{i_k}, x_{j_k}\}) = \{b, c\}\}.$$

We define a mapping for the z-variables $\phi_Z : \{z_1, \cdots, z_m\} \to \{y_1, y_2, y_3\}$ as follows:

$$\phi_Z(z_k) = \begin{cases} y_1 & \text{if } k \in I_{ab} \\ y_2 & \text{if } k \in I_{ac} \\ y_3 & \text{if } k \in I_{bc} \end{cases}.$$

We now define a mapping ϕ from $var(F)$ to $var(F_0)$ by ϕ_X and ϕ_Z as follows:

$$\phi(x) = \begin{cases} x & \text{if } x \in \{y_1, y_2, y_3, a, b, c\} \\ \phi_X(x) & \text{if } x \in \{x_1, \cdots, x_n\} \\ \phi_Z(x) & \text{if } x \in \{z_1, \cdots, z_m\} \end{cases}.$$

Finally, we extend ϕ to a homomorphism from $lit(F)$ to $lit(F_0)$ defining $\phi(\neg x) = \neg\phi(x)$ for every variable $x \in var(F)$. Then we see that ϕ is a homomorphism with $\phi(F) = F_0$.

($\Leftarrow$) Suppose, there exists a homomorphism ϕ with $\phi(F) = F_0$. By the structure of F_0 and F we see: $\phi(x) \in \{a, b, c\}$ for every $x \in V = \{x_1, \cdots x_n\}$ and $\phi(z) \in \{y_1, y_2, y_3\}$ for every $z \in \{z_1, \cdots, z_m\}$. Then for every clause $(\neg z_k \vee x_{i_k} \vee x_{j_k})$ in F we have $\phi(\neg z_k \vee x_{i_k} \vee x_{j_k})$ is in $\{(\neg y_1 \vee a \vee b), (\neg y_2 \vee a \vee c), (\neg y_3 \vee b \vee c)\}$.
In particular we see that $\phi(x_{i_k}) \neq \phi(x_{j_k})$ for every clause $(x_{i_k} \vee x_{j_k})$ and therefore for every edge (x_{i_k}, x_{j_k}) . We divide V into three subsets as follows:

$V_a = \{x \in V | \ \phi(x) = a\}$, $V_b = \{x \in V | \ \phi(x) = b\}$, $V_c = \{x \in V | \ \phi(x) = c\}$.

The partition induces a coloring of G^+ with three different colors. ■

Theorem 4 *For fixed $k, t \geq 1$ the problem $\mathcal{HOM}{-}MU(k,t)$ is NP–complete.*

Proof: Obviously, the problem is in NP. Since any minimal unsatisfiable Horn formula is in $MU(1)$ we immediately obtain the NP–completeness for $k = t = 1$. For other values of k and t we establish a reduction to the problem $\mathcal{HOM}{-}MU(1,1)$.
Let k and t be fixed. For $MU(1)$ formulas F_1 and F_2 we construct formulas $R_1 \in M(k)$ and $R_2 \in MU(t)$, such that $(F_1, F_2) \in \mathcal{HOM}{-}MU(1,1) \Leftrightarrow (R_1, R_2) \in \mathcal{HOM}{-}MU(k,t)$.
At first for any formula F we construct some formulas with new variables
$x, x_i, x'_i, z_j, z'_j, a, b$, and z:
1. $F^x := x \wedge \{(f \vee \neg x) \mid f \in F\}$. We have $F \in MU(k) \Leftrightarrow F^x \in MU(k)$. Please note that F^x is not homomorphic to $y \wedge \neg y$ for $F \in MU$.

2. $F^{\{a,b\}} = \{(a \vee b), (\neg a \vee b)\} + \{(f \vee \neg b) \mid f \in F\}$. We have $F \in MU(k) \Leftrightarrow F^{\{a,b\}} \in MU(k)$.

3. $S(F,k) := \{(h \vee z) \mid h \in Q_k\} + \{(\neg z \vee f) \mid f \in F\}$, where
$Q_k := [(x_k \vee x_1),$
$(\neg x_1 \vee \neg z_1), (z_1 \vee x_2), \cdots, (\neg x_{k-1} \vee \neg z_{k-1}), (z_{k-1} \vee x_k), (\neg x_k \vee \neg z'_{k-1}), (z'_{k-1} \vee x'_{k-1})$
$(\neg x'_{k-1} \vee \neg z'_{k-2}), (z'_{k-2} \vee x'_{k-2}), \cdots, (\neg x'_2 \vee \neg z'_1), (z'_1 \vee x'_1), (\neg x'_1 \vee \neg x_k)]$
The idea of the construction of Q_k is to start with the $MU(1)$ formula
$(\neg x_k \rightarrow x_1 \rightarrow x_2 \rightarrow \cdots \rightarrow x_k \rightarrow x'_{k-1} \rightarrow \cdots \rightarrow x'_1 \rightarrow \neg x_k)$. Then we introduce auxiliary variables z_i and z'_i, such that any clause consists only of negative or positive literals. The formula Q_k is in $MU(1)$.
We have $F \in MU(k) \Leftrightarrow S(F,k) \in MU(k)$.

4. $G(F,k) = \{h \vee z) \mid h \in P_k\} + \{(\neg z \vee f) \mid f \in F\}$, where

$$\begin{aligned} P_k := [&(x_1 \vee x_k), (\neg x_1 \vee \neg x_k) \\ &(x_1 \vee z_1'), (\neg z_1' \vee \neg x_2), (\neg x_1 \vee \neg z_1), (z_1 \vee x_2), \\ &(x_2 \vee z_2'), (\neg z_2' \vee \neg x_3), (\neg x_2 \vee \neg z_2), (z_2 \vee x_3), \\ &\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ &(x_{k-1} \vee z_{k-1}'), (\neg z_{k-1}' \vee \neg x_k), (\neg x_{k-1} \vee \neg z_{k-1}), (z_{k-1} \vee x_k)], \end{aligned}$$

The idea of P_k is to start with the formula $(\neg x_k \leftrightarrow x_1 \leftrightarrow x_2 \leftrightarrow x_3 \leftrightarrow \cdots \leftrightarrow x_{k-1} \leftrightarrow x_k)$. Then we introduce auxiliary variables such that any clause has only positive or only negative literals. The formula P_k is in $MU(k)$. We obtain $F \in MU(r) \Leftrightarrow G(F,k) \in MU(r-1+k)$.

We will show for $k, t \geq 1$ and formulas $F_1, F_2 \in MU(1)$: F_1 is homomorphic to $F_2 \Leftrightarrow F_1^x$ is homomorphic to $F_2^x \Leftrightarrow S(F_1^x, t)$ is homomorphic to $G(F_2^x, t) \Leftrightarrow G((S(F_1^x, t))^y, k)$ is homomorphic to $((G(F_2^x, t))^y)^{\{a,b\}}$.
Please note, that $G((S(F_1^x, t))^y, k) \in MU(k)$ and $((G(F_2^x, t))^y)^{\{a,b\}} \in MU(t)$.

It suffices to show for formulas $F_1, F_2 \in MU(1)$ and $H_1, H_2 \in MU$:

I) F_1 is homomorphic to $F_2 \Leftrightarrow F_1^x$ is homomorphic to F_2^x
II) F_1^x is homomorphic to $F_2^x \Leftrightarrow S(F_1^x, t)$ is homomorphic to $G(F_2^x, t)$
III) H_1 is homomorphic to $H_2 \Leftrightarrow G(H_1^y, k)$ is homomorphic to $(H_2^y)^{\{a,b\}}$

Ad I: Simple observation.
Ad II: ($\Rightarrow$) Suppose we have $\phi(F_1^x) = F_2^x$. For $\varphi_t(x_i') = x_i$, $\varphi_t(x_j) = x_j$, $\varphi_t(z_j) = z_j$, and $\varphi_t(z_j') = z_j'$ we obtain $\varphi_t(Q_t) = P_t$. For the homomorphism $\varphi(a) := \phi(a)$ for $a \in var(F_1^x)$, $\varphi(z) := z$, and $\varphi(y) := \varphi_t(y)$ for $y \in var(Q_t)$ we obtain $\varphi(S(F_1^x, t)) = G(F_2^x, t)$.
($\Leftarrow$) Suppose we have $\phi(S(F_1^x, t)) = G(F_2^x, t)$. Since the variable z is the only variable occurring in any clause of $S(F_1^x, t)$ and $G(F_2^x, t)$, we have $\phi(z) \in \{z, \neg z\}$. Therefore two cases are possible: either $\phi(Q_t) = F_2^x$ and $\phi(F_1^x) = P_t$, or $\phi(Q_t) = P_t$ and $\phi(F_1^x) = F_2^x$. However, the case $\phi(F_1^x) = P_t$ can not occur, because P_t is homomorphic to $y \wedge \neg y$ and F_1^x is not homomorphic to $y \wedge \neg y$. Therefore, we have $\phi(F_1^x) = F_2^x$.

Ad III: $(\Rightarrow)$ Suppose we have $\phi'(H_1) = H_2$. Then there is a homomorphism ϕ such that $\phi(H_1^y) = H_2^y$. For the P_k we define $\varphi_k(x_i) := \varphi_k(z_i') := \varphi_k(z_i) := a$. Then we obtain $\varphi_k(P_k) = a \wedge \neg a$. Further, we define $\varphi(z) := b$, and for the remaining variables $\varphi(v) := \phi(v)$. Then we obtain $\varphi(G(H_1^y, k)) = (H_2^y)^{\{a,b\}})$
$(\Leftarrow)$ Suppose we have $\phi(G(H_1^y, k) = (H_2^y)^{\{a,b\}}$.
Since b is the only variable occurring in every clause of $(H_2^y)^{\{a,b\}}$, we obtain $\phi(z) \in \{b, \neg b\}$. Therefore two cases can occur: $\phi(P_k) = H_2^y$ and $\phi(H_1^y) = a \wedge \neg a$, or $\phi(P_k) = a \wedge \neg a$ and $\phi(H_1^y) = H_2^y$. However, the case $\phi(H_1^y) = a \wedge \neg a$ can not occur. Therefore, we have $\phi(H_1^y) = H_2^y$. ■

Lemma 2 *Let F_1 and F_2 be formulas in $MARG(1)$. For the associated trees T_1 and T_2 we have: There exists a literal renaming φ with $\varphi(F_1) = F_2$ if and only if T_1 is isomorphic to T_2.*

Proof: Let F_1 and F_2 be $MARG(1)$ formulas over variables $x_1, \cdots, x_n$, $F_1 = \{C_1, \cdots, C_{n+1}\}$, and $F_2 = \{C_1', \cdots, C_{n+1}'\}$. We associate to a $MARG(1)$ formula F the undirected graph T_F as follows:
T_F is the undirected graph (V_F, E_F), where $V_F = \{x_1, \cdots, x_n, c_1, \cdots, c_{n+1}\}$ and $E_F = \{(c_i, x_k) \mid x_k \in C_i \text{ or } \neg x_k \in C_j\}$.
T_F can be considered as a bipartite graph. The graph contains no cycles, because there are no two clauses in F containing $(L_1 \vee L_2)$ and $(\neg L_1 \vee \neg L_2)$ for literals L_1 and L_2 as subclauses. $MARG(1)$ are acyclic. Hence, T_F is a tree, where no leaf is a node x_i. Further, we have
(a) $deg(x_k) = 2$ for every $1 \leq k \leq n$. (x_k occurs exactly once positively and once negatively in F)
(b) $1 \leq deg(c_k) \leq n$ for every $1 \leq k \leq n+1$. (size of the clause c_k)
$(\Rightarrow)$ Suppose that there exists a literal renaming φ with $\varphi(F_1) = F_2$. It is known that for any variable x_k, that $x_k \in C_i$ and $\neg x_k \in C_j$ if and only if (x_k, c_i) and (x_k, c_j) are two edges in T_1. Please note $deg(x_k) = 2$ and there exists a permutation π over $\{1, \cdots, n{+}1\}$ such that $\varphi(C_i) = C_{\pi(i)}'$ for $(1 \leq i \leq n+1)$. We now define an isomorphism ϕ with $\phi(T_1) = T_2$ as follows:

$$\phi(x_k) = var(\varphi(x_k)) \text{ for } 1 \leq k \leq n \text{ and } \phi(c_i) = c_{\pi(i)}' \text{ for } 1 \leq i \leq n+1.$$

$(\Leftarrow)$ Let ϕ be an isomorphism with $\phi(T_1) = T_2$. We have two permutations from ϕ: π_v

over $\{x_1, \cdots, x_n\}$ and π_c over $\{1, \cdots, n+1\}$, satisfying $\pi_v(x_k) = \phi(x_k)$ and $c'_{\pi_c(i)} = \phi(c_i)$. Define a formula $F_2^* = \{\pi_v(C_1), \cdots, \pi_v(C_{n+1})\}$ from F_1 by π_v. By the definition of T_2, we have a renaming φ^* over F_2^* such that for every $(1 \leq i \leq n+1)$ $\varphi^*(\pi_v(C_i)) = C'_{\pi_c(i)}$. Finally, we define a literal renaming $\varphi = \pi_v \circ \varphi^*$. It is the desired literal renaming. ■

Let $F = \{C_1, \cdots, C_{n+1}\}$ be a formula over $x_1, \cdots, x_n$ in $MARG(1)$. We construct a directed graph $G_F = (V_F, E_F)$, where $V_F = \{x_1, \cdots, x_n, c_1, \cdots, c_{n+1}\}$ and $E_F = \{(c_i, x_k), (x_k, c_j) \mid x_k \in C_i, \neg x_k \in C_j, 1 \leq k \leq n, 1 \leq i, j \leq n+1\}$. The resulting undirected graph G'_F by ignoring directions of edges is a tree. That follows from the structure of $MARG(1)$ formulas.

Based on the tree G'_F, we now associate F to a tree $T^* = (V^*, E^*)$ by introducing some auxiliary variables to distinguish positive and negative literals as follows:

(1) Let $V^* = V_{var} \cup V_{var}^+ \cup V_{var}^- \cup V_{cl} \cup V_{var+}^* \cup V_{var-}^* \cup V_{cl}^*$, where

$V_{var} = \{x_1, \cdots, x_n\}$, $V_{cl} = \{c_1, \cdots, c_{n+1}\}$,
$V_{var}^+ = \{x_1^+, \cdots, x_n^+\}$, $V_{var}^- = \{x_1^-, \cdots, x_n^-\}$,
$V_{var+}^* = \{y_1^1, \cdots, y_1^{n+2}, \cdots, y_n^1, \cdots, y_n^{n+2}\}$,
$V_{var-}^* = \{z_1^1, \cdots, z_1^{n+1}, \cdots, z_n^1, \cdots, z_n^{n+1}\}$ and
$V_{cl}^* = \{c_1^1, c_1^2, \cdots, c_{n+1}^1, c_{n+1}^2\}$.

(2) Define $E^* = E_0 \cup E_1^+ \cup E_1^- \cup E_2 \cup E_3$, where

$E_0 = \{(c_i, x_k^+), (x_k^-, c_j) \mid x_k \in C_i, \neg x_k \in C_j, 1 \leq k \leq n, 1 \leq i, j \leq n+1\}$,
$E_1^+ = \{(x_k^+, y_k^i) \mid 1 \leq k \leq n, 1 \leq i \leq n+2\}$,
$E_1^- = \{(x_k^-, z_k^i) \mid 1 \leq k \leq n, 1 \leq i \leq n+1\}$,
$E_2 = \{(c_k, c_k^1), (c_k, c_k^2) \mid 1 \leq k \leq n+1\}$ and
$E_3 = \{(x_k^+, x_k), (x_k, x_k^-) \mid 1 \leq k \leq n\}$.

Please note that $|V^*| = 2n^2 + 11n + 5$ and T^* is a tree. And we have that

(a) $deg(x_k^+) = n + 4$ for every $1 \leq k \leq n$.
(b) $deg(x_k^-) = n + 3$ for every $1 \leq k \leq n$.
(c) $deg(x_k) = 2$ for every $1 \leq k \leq n$.
(d) $3 \leq deg(c_k) \leq n + 2$ for every $1 \leq k \leq n + 1$.
(e) $deg(v) = 1$ for every $v \in V_{var+}^* \cup V_{var-}^* \cup V_{cl}^*$.

(Identify the direction of an edge by different degrees of vertices)

Further, $x_k \in C_i$ and $\neg x_k \in C_j$ if and only if there is a unique path, $c_i x_k^+ x_k x_k^- c_j$, from c_i to c_j through x_k. Please note $deg(x_k^+) = n + 4$ and $deg(x_k^-) = n + 3$. In the other words, we use a path to replace the directed edges (c_i, x_k) and (x_k, c_j) in G_F.

Lemma 3 *Let F_1 and F_2 be formulas in $MARG(1)$. For the associated trees T_1^* and T_2^* we have: There exists a variable renaming φ with $\varphi(F_1) = F_2$ if and only if T_1^* is isomorphic to T_2^*.*

Proof: Let F_1 and F_2 be formulas over variables $x_1, \cdots, x_n$, $F_1 = \{C_1, \cdots, C_{n+1}\}$, and $F_2 = \{C'_1, \cdots, C'_{n+1}\}$.

($\Rightarrow$) Suppose that there exists a variable renaming φ with $\varphi(F_1) = F_2$. We have a permutation π_v over $\{1, \cdots, n\}$ and a permutation π_c over $\{1, \cdots, n+1\}$ such that $\varphi(x_k) = x_{\pi_v(k)}$ for $1 \le k \le n$ and $\varphi(C_i) = C'_{\pi_c(i)}$ for $1 \le i \le n+1$.

It is known that for any variable x_k, $x_k \in C_i$ and $\neg x_k \in C_j$ if and only if (c_i, x_k^+), (x_k^-, c_j), (x_k^+, x_k), (x_k, x_k^-) are edges in T_1^*.

We now define an isomorphism ϕ with $\phi(T_1^*) = T_2^*$ as follows:

(1) $\phi(x_k) = x_{\pi_v(k)}$, $\phi(x_k^+) = x^+_{\pi_v(k)})$, $\phi(x_k^-) = x^-_{\pi_v(k)}$ $(1 \le k \le n)$

(2) $\phi(c_i) = c'_{\pi_c(i)}$ $(1 \le i \le n+1)$.

(3) $\phi(y_k^p) = y^p_{\pi_v(k)}$ for $(1 \le k \le n)$ and $(1 \le p \le n+2)$.

(4) $\phi(z_k^q) = z^q_{\pi_v(k)}$ for $(1 \le k \le n)$ and $(1 \le q \le n+1)$.

(5) $\phi(c_i^p) = c^p_{\pi_c(i)}$ $(1 \le i \le n+1)$ and $p = 1, 2$.

($\Leftarrow$) Let ϕ be an isomorphism with $\phi(T_1^*) = T_2^*$. By the difference of degrees of vertices, we have that $\phi(V_{var}) = V_{var}$, $\phi(V_{var}^+) = V_{var}^+$, $\phi(V_{var}^-) = V_{var}^-$ and $\phi(V_{cl}) = V_{cl}$.

The restriction $\phi|_{V_{var}}$ is the desired variable renaming, since $x_k \in C_i$ and $\neg x_k \in C_j$ if and only if there is a unique path, $c_i x_k^+ x_k x_k^- c_j$, from c_i to c_j through x_k. ■