



HAL
open science

On the complexity of Submap Isomorphism

Christine Solnon, Guillaume Damiand, Colin de La Higuera, Jean-Christophe Janodet

► **To cite this version:**

Christine Solnon, Guillaume Damiand, Colin de La Higuera, Jean-Christophe Janodet. On the complexity of Submap Isomorphism. Graph-Based Representation in Pattern Recognition, May 2013, Vienna, Austria. pp.21-30, 10.1007/978-3-642-38221-5_3. hal-00825794

HAL Id: hal-00825794

<https://hal.science/hal-00825794>

Submitted on 24 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the complexity of Submap Isomorphism^{*}

Christine Solnon^{1,2}, Guillaume Damiand^{1,2}, Colin de la Higuera³, and
Jean-Christophe Janodet⁴

¹ INSA de Lyon, LIRIS, UMR 5205 CNRS, 69621 Villeurbanne, France

² Université de Lyon, France

³ LINA, UMR CNRS 6241, Université de Nantes, France

⁴ IBISC, Université d'Evry, France

Abstract. Generalized maps describe the subdivision of objects in cells, and incidence and adjacency relations between cells, and they are widely used to model 2D and 3D images. Recently, we have defined submap isomorphism, which involves deciding if a copy of a pattern map may be found in a target map, and we have described a polynomial time algorithm for solving this problem when the pattern map is connected. In this paper, we show that submap isomorphism becomes NP-complete when the pattern map is not connected, by reducing the NP-complete problem Planar-4 3-SAT to it.

1 Motivations

Combinatorial maps and generalized maps [1] are very nice data structures to model the topology of nD objects subdivided in cells (*e.g.*, 0D vertices, 1D edges, 2D faces, 3D volumes, ...) by means of incidence and adjacency relationships between these cells. In 2D, maps may be used to model the topology of an embedding of a planar graph in a plane. In particular, these models are very well suited for scene modeling [2], and for 2D and 3D image segmentation [3].

In [4], we have defined a basic tool for comparing 2D maps, *i.e.*, submap isomorphism (which involves deciding if a copy of a pattern map may be found in a target map), and we have proposed an efficient polynomial-time algorithm for solving this problem when the pattern map is connected. This work has been generalized to nD maps in [5]. The subisomorphism defined in [5] is based on induced submap relations, such that submaps are obtained by removing some darts and all their seams, just like induced subgraphs are obtained by removing some vertices and all their incident edges. In [6], we have introduced a new kind of submap relation, called partial submap: partial submaps are obtained by removing not only some darts (and all their seams), but also some other seams, just like partial subgraphs are obtained by removing not only some vertices (and their incident edges), but also some other edges. The polynomial time

^{*} Paper published in Proceedings of 9th Workshop on Graph-Based Representation in Pattern Recognition, LNCS 7877, pp. 21-30, May 2013. Thanks to Springer Berlin Heidelberg. The original publication is available at http://dx.doi.org/10.1007/978-3-642-38221-5_3

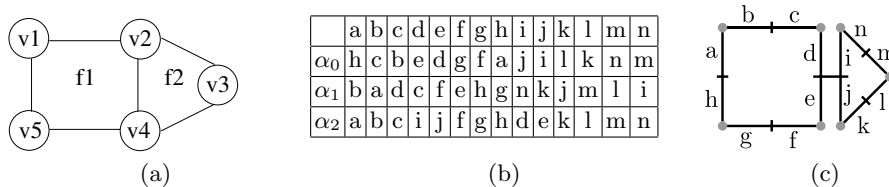


Fig. 1. (a) A plane graph. (b) The corresponding 2G-map. (c) Its graphical representation: darts are represented by segments labeled with letters, consecutive darts separated with a little segment are 0-sewn (e.g., $\alpha_0(b) = c$ and $\alpha_0(c) = b$), consecutive darts separated with a dot are 1-sewn (e.g., $\alpha_1(a) = b$ and $\alpha_1(b) = a$), parallel adjacent darts are 2-sewn (e.g., $\alpha_2(d) = i$ and $\alpha_2(i) = d$).

algorithm described in [5] for solving the induced submap isomorphism problem may be extended to the partial case in a very straightforward way. However, this algorithm still assumes that the pattern map is connected. In this paper, we show that the submap isomorphism problem becomes NP-complete when the pattern map is not connected, both for partial and induced submaps.

Outline of the paper. In Section 2, we recall definitions related to generalized maps. In Section 3, we define the submap isomorphism problem and recall some complexity results about this problem. In Section 4, we describe the planar-4 3-SAT problem, which is NP-complete. In Section 5, we describe a polynomial-time reduction of planar-4 3-SAT to submap isomorphism, thus showing that submap isomorphism is NP-complete.

2 Recalls and basic definitions on generalized maps

In this work we consider generalized maps, which are more general than combinatorial maps, and we refer the reader to [1] for more details.

Definition 1. (*nG-map*) Let $n \geq 0$. An n -dimensional generalized map (or nG -map) is defined by a tuple $G = (D, \alpha_0, \dots, \alpha_n)$ such that (i) D is a finite set of darts; (ii) $\forall i \in \{0, \dots, n\}$, α_i is an involution⁵ on D ; and (iii) $\forall i, j \in \{0, \dots, n\}$ such that $i + 2 \leq j$, $\alpha_i \circ \alpha_j$ is an involution.

2G-maps may be used to model the embedding of a planar graph into a plane. For example, Fig. 1 displays a plane graph and the corresponding 2G-map. We say that a dart d is i -sewn with a dart d' whenever $d = \alpha_i(d')$ and $d \neq d'$, whereas it is i -free whenever $d = \alpha_i(d)$. A seam is a tuple (d, i, d') such that d' is i -sewn to d . For example, $(a, 0, h)$ is a seam of the map displayed in Fig. 1 because $\alpha_0(a) = h$.

Definition 2. (*seams of a set of darts in an nG-map*) Let $G = (D, \alpha_0, \dots, \alpha_n)$ be an nG -map and $E \subseteq D$ be a set of darts. The set of seams associated with E in G is: $\text{seams}_G(E) = \{(d, i, \alpha_i(d)) \mid d \in E, i \in \{0, \dots, n\}, \alpha_i(d) \in E, \alpha_i(d) \neq d\}$.

⁵ An involution f on D is a bijective mapping from D to D such that $f = f^{-1}$.

A map is connected if any pair of darts is connected by a path of sewn darts.

Definition 3 (Connected map). A generalized map $G = (D, \alpha_0, \dots, \alpha_n)$ is connected if $\forall d \in D, \forall d' \in D$, there exists a path between d and d' , i.e., a sequence of darts (d_1, \dots, d_k) such that $d_1 = d$, $d_k = d'$ and $\forall i \in \{1, \dots, k-1\}$, $\exists j_i \in \{0, \dots, n\}, d_{i+1} = \alpha_{j_i}(d_i)$.

Map isomorphism [1] allows us to decide the equivalence of two maps.

Definition 4. (*nG-map isomorphism [1]*) Two *nG*-maps $G = (D, \alpha_0, \dots, \alpha_n)$ and $G' = (D', \alpha'_0, \dots, \alpha'_n)$ are isomorphic if there exists a bijection $f : D \rightarrow D'$, such that $\forall d \in D, \forall i \in [0, n], f(\alpha_i(d)) = \alpha'_i(f(d))$.

In [4], induced submaps have been defined: G is an induced submap of G' if G preserves all seams of G' , i.e., for every couple of darts (d_1, d_2) of G , d_1 is i -sewn to d_2 in G' if and only if d_1 is i -sewn to d_2 in G .

Definition 5. (*induced submap*) A map $G' = (D', \alpha'_0, \dots, \alpha'_n)$ is an induced submap of $G = (D, \alpha_0, \dots, \alpha_n)$ if $D' \subseteq D$ and $\text{seams}_{G'}(D') = \text{seams}_G(D')$.

In [6], we have introduced another submap relation, called partial submap by analogy with existing work on graphs. Indeed, induced subgraphs are obtained by removing some nodes (and all their incident edges) whereas partial subgraphs are obtained by removing not only some nodes (and all their incident edges) but also some edges. In our map context, partial submaps are obtained by removing not only some darts (and all their seams) but also some other seams.

Definition 6. (*partial submap*) A map $G' = (D', \alpha'_0, \dots, \alpha'_n)$ is a partial submap of $G = (D, \alpha_0, \dots, \alpha_n)$ if $D' \subseteq D$ and $\text{seams}_{G'}(D') \subseteq \text{seams}_G(D')$.

3 The submap isomorphism problem

The submap isomorphism problem involves deciding if a pattern map is isomorphic to a submap of a target map, and it is formally defined as follows:

Problem: Partial (resp. induced) submap isomorphism
Instance: A triple (n, G, G') such that $n > 0$, and G and G' are *nG*-maps.
Question: Does there exist a partial (resp. induced) submap of G' which is isomorphic to G ?

We note $G \sqsubseteq^p G'$ (resp. $G \sqsubseteq^i G'$) when the answer is yes. Note that $G \sqsubseteq^i G' \Rightarrow G \sqsubseteq^p G'$. Fig. 2 displays examples of submap isomorphisms.

The complexity of the submap isomorphism problem depends on the connectedness of the pattern map. For example, the map G_1 of Fig. 2 is not connected, and is composed of two connected components, whereas the maps G_2, G_3 and G_4 are connected. In [5], we have described a polynomial-time algorithm which solves the submap isomorphism problem when the pattern map G is connected. When the pattern map G is not connected, we may use this algorithm to search for all occurrences of each connected component of G in the target map G' . Let us consider, for example, the map G_1 of Fig. 2. Its left hand side component

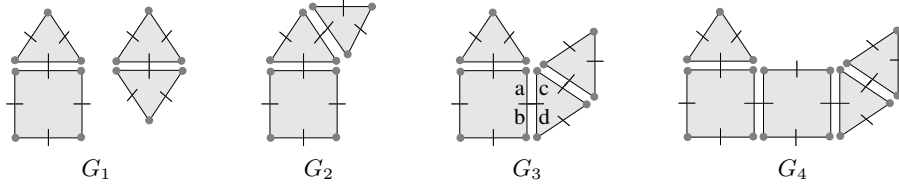


Fig. 2. Submap isomorphism examples. G_1 is not isomorphic to a submap of G_2 (i.e., $G_1 \not\sqsubseteq^p G_2$ and $G_1 \not\sqsubseteq^i G_2$), though each connected component of G_1 is a submap of G_2 . G_1 is isomorphic to a partial submap of G_3 , but not to an induced one (i.e., $G_1 \sqsubseteq^p G_3$ and $G_1 \not\sqsubseteq^i G_3$), because the seams $(a, 2, c)$ and $(b, 2, d)$ of G_3 are not preserved in G_1 . G_1 is isomorphic to an induced submap of G_4 and, therefore, it is also isomorphic to a partial submap of G_4 (i.e., $G_1 \sqsubseteq^p G_4$ and $G_1 \sqsubseteq^i G_4$).

occurs once in G_2 and twice in G_3 and G_4 , and its right hand side component occurs twice in G_2 , G_3 and G_4 (as it is automorphic). To solve the submap isomorphism problem from these occurrence lists, we have to solve the following combinatorial problem: Can we select one occurrence in G' of each connected component of G so that the selected occurrences do not overlap in G' ?

Theorem 1 claims that this combinatorial problem is NP-complete.

Theorem 1. *The partial (resp. induced) submap isomorphism problem is NP-complete.*

The problem trivially belongs to NP since one can check that a given partial (resp. induced) submap of the target map G' is isomorphic to the pattern map G in polynomial time. We may use for example the polynomial algorithm of [5], which has been defined for non connected maps.

To prove that it is NP-complete, we show in Section 5 that Planar-4 3-SAT, which is known to be NP-complete, may be reduced to it in polynomial time.

4 Planar-4 3-SAT

Planar-4 3-SAT is a special case of the SAT problem, which involves deciding if there exists a truth assignment for a set X of variables such that a boolean formula F over X is satisfied [?]. We assume that F is in Conjunctive Normal Form (CNF), i.e., it is a conjunction of clauses such that each clause is a disjunction of literals which are either variables of X or negations of variables of X .

The formula-graph associated with a CNF formula F over a set of variables X is the bipartite graph $G_{X,F} = (V, E)$ such that V associates a vertex with every variable $x_i \in X$ and every clause c_j of F , and E associates an edge (x_i, c_j) with every variable/clause couple such that variable x_i occurs in clause c_j .

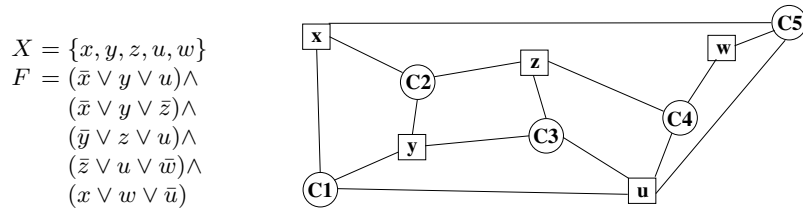


Fig. 3. An instance of Planar-4 3-SAT and its associated formula graph (clauses correspond to circles, and variables to squares).

The planar-4 3-SAT problem is formally defined as follows.

Problem: Planar-4 3-SAT

Instance: A couple (X, F) such that X is a set of boolean variables and F is a CNF formula over X such that (i) every clause of F is a disjunction of 3 literals, (ii) the formula-graph $G_{X,F}$ is planar, and (iii) the degree of every vertex of $G_{X,F}$ is bounded by 4 (*i.e.*, each variable occurs in at most 4 different clauses)

Question: Does there exist a truth assignment for X which satisfies F ?

Planar-4 3-SAT has been shown to be NP-complete in [7]. Fig. 3 displays an instance of Planar-4 3-SAT and its associated formula-graph.

To reduce an instance (X, F) of Planar-4 3-SAT to an instance (n, G, G') of submap isomorphism, we first perform a preprocessing: We iteratively eliminate from (X, F) every variable $x_i \in X$ which occurs in only one clause c_j of F (those whose degree is equal to 1 in the formula-graph), set x_i to the truth value which satisfies c_j , and eliminate c_j from F , until either X and F become empty (thus showing that the answer is trivially yes), or all variables in X occur in 2, 3, or 4 clauses of F .

5 Reduction of Planar-4 3-SAT to submap isomorphism

Let us first show that planar-4 3-SAT can be reduced to induced submap isomorphism in polynomial time: The partial case will be studied at the end of this section. We consider an instance (X, F) of planar-4 3-SAT and we show how to build an instance (n, G, G') such that $G \sqsubseteq^i G'$ iff the answer to (X, F) is yes.

We consider 2G-maps, so that $n = 2$, and the 2G-maps G and G' are constructed by assembling building blocks which are 2G-maps. Fig. 4 displays building blocks associated with variables: For each variable $x_i \in X$ such that the degree of x_i in the formula-graph $G_{X,F}$ is equal to k with $2 \leq k \leq 4$ (as the preprocessing step has removed any variable whose degree is equal to 1), we build two variable patterns V'_k and V_k which will respectively occur in G' and G . Each variable pattern V'_k (resp. V_k) looks like a flower whose core is a $2k$ -edge face and which have $2k$ petals (resp. k petals), where each petal is a 6-edge face. For each petal in each variable pattern V'_k , the edge opposite to the core of the flower is a connecting edge which may be 2-sewn with clause patterns to define G' .

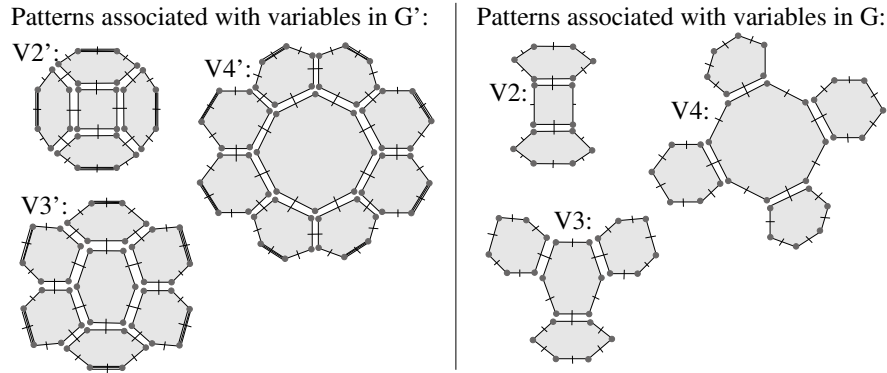


Fig. 4. Variable patterns used as building blocks to define G' and G . Connecting edges in G' are displayed in bold.

For each clause, we build two clause patterns C' and C which will respectively occur in G' and G . The clause pattern C' is composed of a 3-edge central face which has 3 adjacent 4-edge faces, whereas the clause pattern C is composed of a 3-edge face which has 1 adjacent 4-edge face, as displayed below:



Edges of C' displayed in bold are connecting edges which are 2-sewn with variable patterns to define G' .

Definition of the 2G-map G' . For each variable $x_i \in X$ such that the degree of x_i in the formula-graph $G_{X,F}$ is equal to k , G' contains an occurrence of the variable pattern V'_k . Each petal of this occurrence of V'_k is alternatively labeled with x_i and \bar{x}_i . For each clause c_j of F , G' contains an occurrence of the clause pattern C' . Each 4-edge face of this occurrence of C' is labeled with a different literal of c_j . Variable and clause patterns are 2-sewn to define a connected 2G-map: every connecting edge of each clause pattern is 2-sewn with a different connecting edge of a variable pattern such that the two faces which become adjacent by this seam are labeled with the same literal. We can easily check that this 2G-map can always be built in polynomial time as the formula-graph $G_{X,F}$ is planar, and there exist polynomial-time algorithms for embedding a planar graph in a plane [8]: We can use the same embedding for constructing G' . Fig. 5 displays the 2G-map associated with the formula displayed in Fig. 3.

Definition of the 2G-map G . If the SAT instance has n variables and c clauses, then G is composed of $n + c$ different components: a component V_k is associated with every variable $x_i \in X$, where k is the degree of x_i in $G_{X,F}$; a component C is associated with every clause. For example, the 2G-map G associated with the formula displayed in Fig. 3 contains 10 components: 3 occurrences of V_3 , 1 occurrence of V_4 , 1 occurrence of V_2 , and 5 occurrences of C .

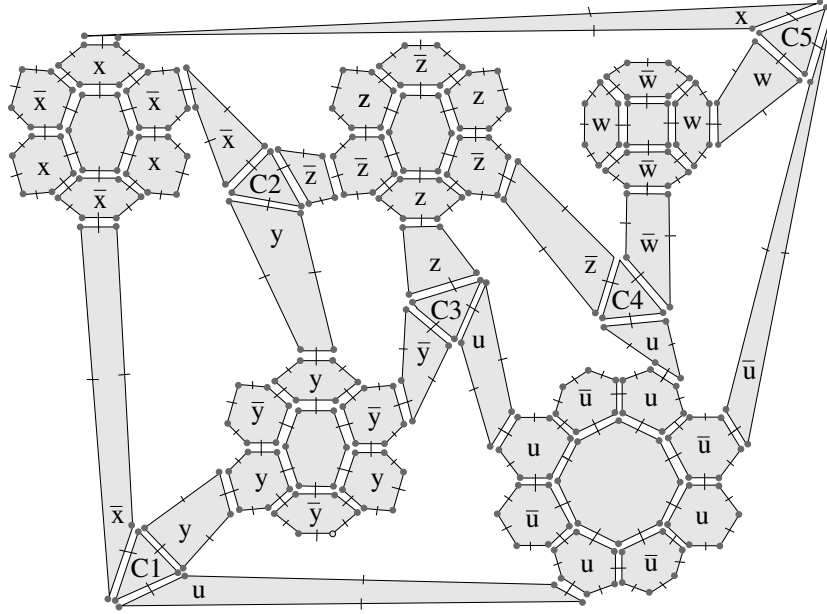


Fig. 5. 2G-map G' associated with the SAT instance displayed in Fig. 3. Note that this map contains holes (corresponding to white parts in the figure): each dart d adjacent to these holes is 2-free so that $\alpha_2(d) = d$.

Proof of $(G \sqsubseteq^i G') \Rightarrow (\exists \text{ truth assignment of } X \text{ which satisfies } F)$. Let us first assume that there exists an induced submap G'' of G' which is isomorphic to G , and let us show that there exists a truth assignment of X which satisfies F .

If G'' is isomorphic to G then, according to Def. 4, there exists a bijection f which matches darts of G'' with darts of G and which preserves all seams. By extension, we say that f matches faces of G'' with faces of G . As we consider induced submap isomorphism, two faces of G which belong to two different connected components cannot be matched by f with faces which are 2-sewn in G'' (according to Def. 5). Fig. 6 displays an example of such a solution for the instance $(2, G, G')$ of the induced submap isomorphism problem associated with the instance (X, F) of Planar-4 3-SAT displayed in Fig. 3.

G contains c occurrences of C , where c is the number of clauses of F . Each occurrence of C has a 3-edge face adjacent to a 4-edge face. These faces can only be matched with faces which belong to occurrences of C' in G' as 3-edge faces in G' only come from C' patterns. As there are c occurrences of C in G , each occurrence of C' in G' is matched with a different occurrence of C in G . For the same reasons, each occurrence of a variable pattern V_k in G is matched with a different occurrence of a variable pattern V'_k in G' : Petal and core faces in G can only be matched with petal and core faces in G' , and an occurrence of V_i cannot be matched with faces of an occurrence of V'_j if $i \neq j$. For each variable pattern V'_k , the label of the petals of V'_k which are not matched with petals of variable

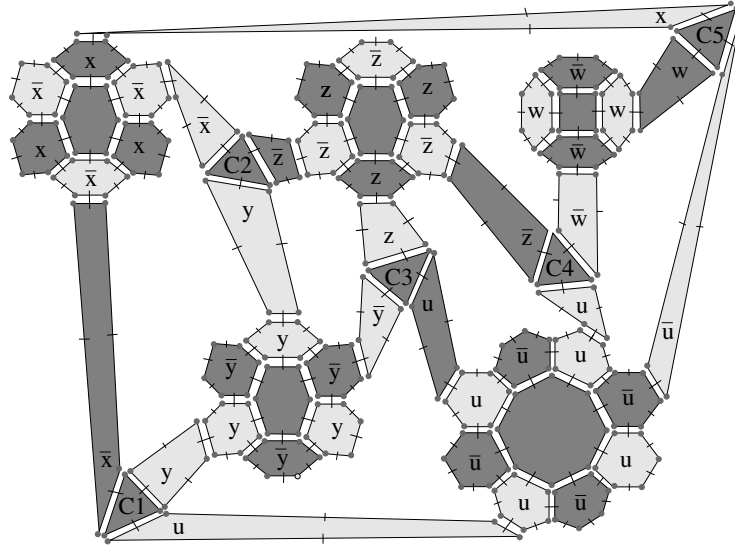


Fig. 6. Solution of the induced submap isomorphism instance $(2, G, G')$ associated with the Planar-4 3-SAT instance displayed in Fig. 3. The induced submap of G' which is isomorphic to G is displayed in dark grey. Note that two different components of this submap cannot be 2-sewn in G' as we consider induced submap isomorphism.

patterns of G gives the truth assignment for the corresponding variable. For each clause pattern C' , the label of the 4-edge face of C' which is matched with a 4-edge face of C corresponds to a literal which satisfies the clause associated with C' . As two faces of G which belong to two different connected components cannot be matched by f with faces which are 2-sewn in G'' , we ensure that when a 4-edge face of a clause pattern is matched, then the adjacent petal is not matched, *i.e.*, when a clause is satisfied by a literal l , then no other clause can be satisfied by the negation of this literal so that the truth assignment deduced from the flower matching actually satisfies all clauses of F . For example, the truth assignment corresponding to the solution displayed in Fig. 6 is $\{\bar{x}, y, \bar{z}, u, w\}$.

Proof of $(\exists \text{ truth assignment of } X \text{ which satisfies } F) \Rightarrow (G \sqsubseteq^i G')$. Let us assume that there exists a truth assignment of X which satisfies F and let us show that there exists an induced submap G'' of G' which is isomorphic to G , *i.e.*, that there exists a dart matching which preserves all seams of G . For each variable pattern V_k in G associated with a variable x_i , we match the darts of the core face with the darts of the core face of the variable pattern associated with x_i in G' and we match the darts of the k 6-edge petals of V_k with the darts of the k 6-edge petals which are labeled with the negation of the truth value of x_i . For each clause pattern C in G associated with a clause c_j , we match the darts of the 3-edge face of C with the darts of the 3-edge face of the clause pattern associated with c_j in G' and we match the darts of the 4-edge face of C with the darts of one of the three 4-edge faces: We choose a 4-edge face which is labeled

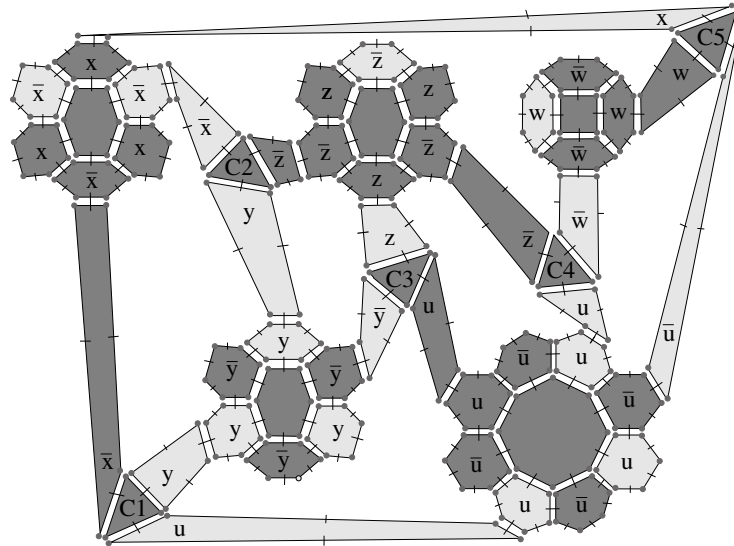
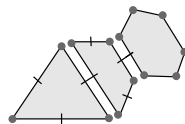


Fig. 7. Solution of the partial submap isomorphism instance $(2, G, G')$ associated with the Planar-4 3-SAT instance displayed in Fig. 3. The partial submap of G' which is isomorphic to G is displayed in dark grey.

with a literal which is satisfied by the truth assignment (this 4-edge face cannot be 2-sewn with a matched 6-edge petal).

Proof for the partial case. Let us now consider the partial case: We consider an instance (X, F) of planar-4 3-SAT and we show how to build an instance (n, G, G') such that $G \sqsubseteq^p G'$ iff the answer to (X, F) is yes. The proof is similar to the induced case. The difference between the induced and the partial cases is that, when considering induced submap isomorphism, two faces which belong to two different components in G cannot be matched with faces of G' which are 2-sewn whereas, when considering partial submap isomorphism, two faces which belong to two different components in G may be matched with faces of G' which are 2-sewn. Therefore, we modify the clause pattern C so that the 4-edge face is adjacent to a 3-edge face, on one side, and to a 6-edge face on the opposite side, as displayed below:



These 6-edge faces can only be matched with petals. The label of the petal which is matched with the 6-edge face of a clause pattern corresponds to the literal which satisfies the clause. Fig. 7 displays an example of solution for partial submap isomorphism.

6 Conclusion

We have shown that submap isomorphism is NP-complete when the pattern map G is not connected. This implies that there does not exist a polynomial-time algorithm for this problem, unless $P=NP$. The practical tractability of this problem actually depends on the number of different connected components of G . Indeed, if G contains k different connected components, we can use the polynomial-time algorithm of [5] to search for all occurrences of each component of G in the target map G' . Let m be the maximum number of occurrences of a connected component of G in G' (m is bounded by the number of darts of G'). The number of candidate solutions to explore is bounded by m^k so that the problem remains tractable if k is small enough.

A consequence of our NP-completeness proof is that the maximum common submap problem introduced in [9] is NP-hard in the general case, i.e., if the common submap is not necessarily connected (as searching for a common submap is more general than deciding of submap isomorphism). However, the complexity of the maximum common submap problem in the particular case where the common submap must be connected is still an open question: We haven't found a polynomial-time algorithm for solving this problem, neither have we found a polynomial-time reduction from a known NP-complete problem to this problem. Hence, further work will mainly concern the answer to this question.

Acknowledgments. The authors would like to thank Daniel Goncalves (University of Montpellier) for his pointer to problem Planar-3SAT, and fruitful remarks.

References

1. Lienhardt, P.: N-dimensional generalized combinatorial maps and cellular quasi-manifolds. *Computational Geometry and Applications* **4**(3) (1994) 275–324
2. Fradin, D., Meneveaux, D., Lienhardt, P.: A hierarchical topology-based model for handling complex indoor scenes. *Computer Graphics Forum* **25**(2) (June 2006) 149–162
3. Braquelaire, J.P., Brun, L.: Image segmentation with topological maps and inter-pixel representation. *Visual Communication and Image representation* **9**(1) (1998) 62–79
4. Damiand, G., De La Higuera, C., Janodet, J.C., Samuel, E., Solnon, C.: Polynomial algorithm for submap isomorphism: Application to searching patterns in images. In: *GbR*. Volume 5534 of LNCS., Springer (2009) 102–112
5. Damiand, G., Solnon, C., de la Higuera, C., Janodet, J.C., Samuel, E.: Polynomial algorithms for subisomorphism of nd open combinatorial maps. *Computer Vision and Image Understanding (CVIU)* **115**(7) (July 2011) 996–1010
6. Combier, C., Damiand, G., Solnon, C.: From maximum common submaps to edit distances of generalized maps. *Pattern Recognition Letters* **33**(15) (2012) 2020–2028
7. Jansen, K., Müller, H.: The minimum broadcast time problem for several processor networks. *Theoretical Computer Science* **147**(1-2) (1995) 69–85
8. Mohar, B.: A linear time algorithm for embedding graphs in an arbitrary surface. *SIAM Journal on Discrete Mathematics* **12**(1) (1999) 6–26
9. Combier, C., Damiand, G., Solnon, C.: Measuring the distance of generalized maps. In: *GbR*. LNCS, Springer (2011) 82–91