Covering Bitmap with Trapezoids is Hard

Changzhou Wang, *X. Sean Wang*
Department of Information and Software Engineering
George Mason University, Fairfax, Virginia, USA
{cwang, xywang}@gmu.edu

Abstract

Each sub-series of a time series can be mapped to a point on a plane whose two coordinates represent the starting time and the ending time, respectively. In certain applications, points with similar properties are likely to be clustered in a special trapezoid shape. To save these points in a compact way, the following optimization problem needs to be studied: find the minimum number of trapezoids which cover a given set of points. This paper formalizes the optimization problem as a decision problem, namely, Bitmap Cover with special Trapezoids (BCT), and proves that BCT is NP-hard.

1 Introduction

In applications such as data compression, it is helpful to solve the problem of covering (two-dimensional) bitmap with a minimum number of simple shapes. When rectangular shapes are used, this problem is proved to be NP-hard [GJ79]. In this paper, we focus on a kind of trapezoids, and prove that the bitmap cover problem with this kind of trapezoids is NP-hard. This problem arises from the application of time series databases. In the rest of the introduction, we first describe the motivation application, and then discuss the bitmap cover problem in more detail.

In time series databases, it is often desirable to support sub-series queries, namely, to retrieve time series containing sub-series that have certain properties. The time period of the sub-series is often not specified in the query (although the time span may be given). For example, in a stock price database, users may look for stocks whose prices go up for at least five months consecutively. The system needs to search for all sub-series of any five consecutive months or longer that show an up trend.

A direct way to evaluate such queries is to scan each series in the database and verify all of its sub-series against the query condition. This method is obviously costly. To reduce query response time, pre-computation of interesting properties for all sub-series of all series can be helpful. However, when the number of series as well as the length of series is very large, the amount of information obtained from this pre-computation is often too large to be stored directly. Fortunately, this problem can be alleviated via the following observations:

Observation 1 *For a given time series, a sub-series is likely to have similar properties as those sub-series with almost the same time period.* For example, if a stock price is considered (e.g., using the moving average method) going up starting from January 1, 1998 and ending on February 28, 1998, it is likely to be considered going up starting from some day around January 1, 1998 and ending on some day around February 28, 1998. The fundamental reason for this observation is that many properties of sub-series are derived from all the points in the sub-series, and such properties change slowly over time.

Observation 2 *Usually exact values of properties are not required. Instead, all the possible values of a property can be partitioned into groups, and information of groups (instead of exact values) may be used to categorize the property of each sub-series.* For example, an exact slope of upward trend such as $23.4°$ is not of particular interest in many applications. Instead, all possible trends may be partitioned, for instance, into three groups: up-trend, down-trend, and stable-trend.

When the above observations are valid in an application domain, the pre-computation results belonging to the same group can be clustered, and the information of the clusters (instead of each individual sub-series) can be stored and used to approximate the results. Formally, for a time series S, each of its sub-series, denoted $S[i,j]$ where i and j are positive integers with $i<j$, can be mapped to a point (i,j) on a plane whose two axes represent the starting time and the ending time, respectively. This plane is called the **sub-series plane**. For each time series, there is a sub-series plane for each property group G (in a particular partition of values). For any point (i,j) on the sub-series plane, it is called a **solid** point if $i<j$ and $S[i,j]$ has the property specified by G; otherwise, it is called a **hollow** point. Points with non-integer coordinates are not consider here. *The problem is to find a minimum set of clusters of the solid points on each sub-series plane.*

From the above observations and other research results [QWW98], the clusters usually show a special trapezoid shape, as illustrated in Figure 1.[1] Specifically, each trapezoid of this special kind is isosceles, its two legs are parallel to the two axes respectively, its bases are on lines $y = x + c_i$ where c_i's are positive integers, and its shorter base is to the left-top of its longer base. Such a trapezoid can therefore be denoted by the tuple (s,e,a,b) (see Figure 1). The points in the trapezoid (s,e,a,b) correspond to all the sub-series which start no earlier than s, end no later than e, and are of lengths between a and b inclusively. Note that such a trapezoid may degenerate to a line when $a=b$, or a triangle when $e-s=b$.

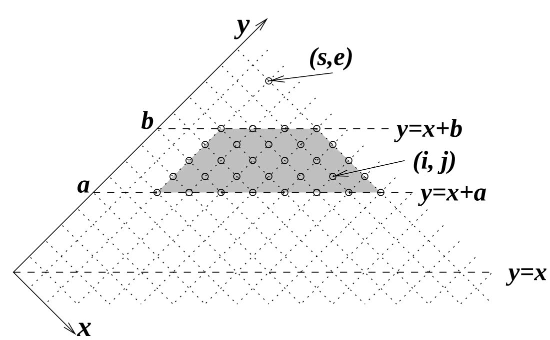


Figure 1: An example of the special trapezoids (with coordinate system rotated 45° clockwise)

The rest of the paper consists of two sections. Section 2 formalizes the optimization problem as a decision problem, and proves that the decision problem is NP-hard; and Section 3 concludes the paper.

[1]To save space, the coordinate system is rotated 45° clockwise in Figure 1 as well as in the subsequent figures. In addition, for simplicity, the word "trapezoid" used in the following discussion always means this special kind of trapezoids, if not mentioned otherwise.

2 Cover with trapezoids

If we use "1" to indicate a solid point and "0" to indicate a hollow point, we can use a (two-dimensional) bitmap to represent a sub-series plane. Hence, the aforementioned optimization problem can be easily transformed into the following decision problem:

Bitmap Cover with special Trapezoids (BCT)
Instance: An $l \times l$ bitmap B (i.e., B is an $l \times l$ matrix of 0's and 1's) where $B_{ij} = 0$ for $1 \leq j \leq i \leq l$, and a positive integer K.
Question: Is there a collection of K or fewer trapezoids that cover precisely those points in B that are 1's, i.e., is there a set of K' quadruples (s_m, e_m, a_m, b_m), where $K' \leq K$, s_m, e_m, a_m and b_m are positive integers and $a_m \leq b_m \leq e_m - s_m$, such that for every pair of integers (i, j), $B_{ij} = 1$ if and only if there exists a k, $1 \leq k \leq K'$, satisfying $s_k \leq i < j \leq e_k$ and $a_k \leq j - i \leq b_k$.

This section proves the following:

Theorem 1 *The BCT decision problem is NP-hard.*

Before going into details of proof, we first introduce some basic definitions. For a point (i, j) where i and j are integers and $1 \leq i, j \leq l$, it is **solid** if $B_{ij} = 1$; otherwise (i.e., $B_{ij} = 0$) it is **hollow**. For a trapezoid (s, e, a, b), a point (i, j) is **covered** by this trapezoid if and only if $s \leq i < j \leq e$ and $a \leq j - i \leq b$. A trapezoid T is said to **cover** another trapezoid T' if for any point (i, j) covered by T', (i, j) is also covered by T. A **candidate** trapezoid is a trapezoid that covers only solid points. A candidate trapezoid is a **maximum** one if no other candidate trapezoid can cover it. Two solid points are said to be **invisible to each other** if no single candidate trapezoid can cover both points. Clearly, given a set of *solid* points which are pairwise invisible, each of them requires a unique candidate trapezoid to cover. Figure 2 shows some patterns of pairwise-invisible solid point sets. In each pattern, solid point A is invisible to each solid point B due to the existence of hollow point C. These patterns will be used in the subsequent component constructions to ensure a minimum number of trapezoids which are needed for a covering instance.

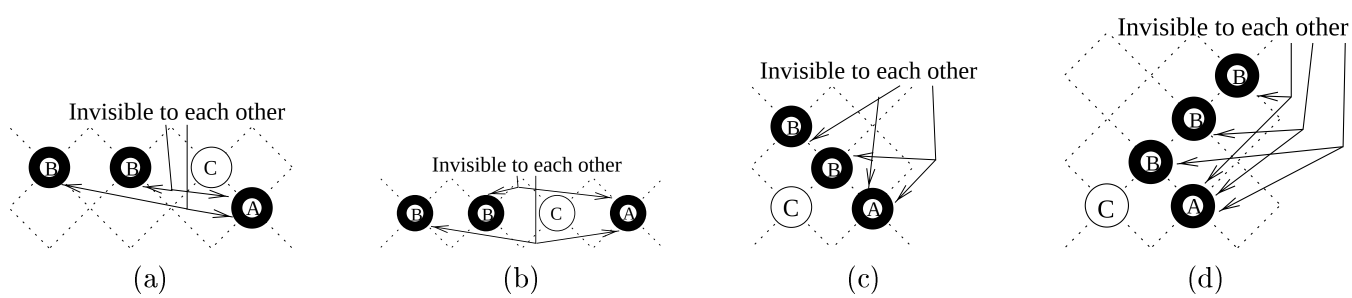


Figure 2: Invisible solid point pairs

The proof of the theorem is based on a reduction from the three-satisfiability problem, 3SAT[GJ79]:
Instance: A set $\mathcal{U}$ of variables and a collection $\mathcal{C}$ of clauses over $\mathcal{U}$ such that each clause c in $\mathcal{C}$ has exactly three (positive or negative) occurrences of variables in $\mathcal{U}$, i.e., $|c| = 3$.
Question: Is C satisfiable, i.e., is there a satisfying truth assignment for $\mathcal{C}$?

Let $I = (\mathcal{U}, \mathcal{C})$ be an arbitrary instance of 3SAT. We will show how to polynomially transform I

to an instance $J = (B, K)$ of BCT, such that B can be covered by K or fewer trapezoids if and only if I is satisfiable. We use the basic component design approach [GJ79, CR94]. Figure 3 illustrates the final structure of J. Intuitively, for each variable v_i in $\mathcal{U}$ we construct a "variable structure" V_i. For each clause c_i in $\mathcal{C}$ we construct a "clause checker" C_i. For each variable occurrence f_i of variable v_{j_1} in clause c_{j_2}, we construct a "variable occurrence shifter" F_i to associate the variable structure V_{j_1} with the clause checker C_{j_2}. For example, in Figure 3, F_1 associates V_n with C_1, while F_r associates V_1 with C_m. The association is achieved via "connectors", which are shown as shadow regions in Figure 3. Each connector "connects" two components (and possibly goes through several other components). For example, the connector t_1 connects V_1 and F_r while going through F_1, and the connector t_2 connects F_r and C_m without going through other components.

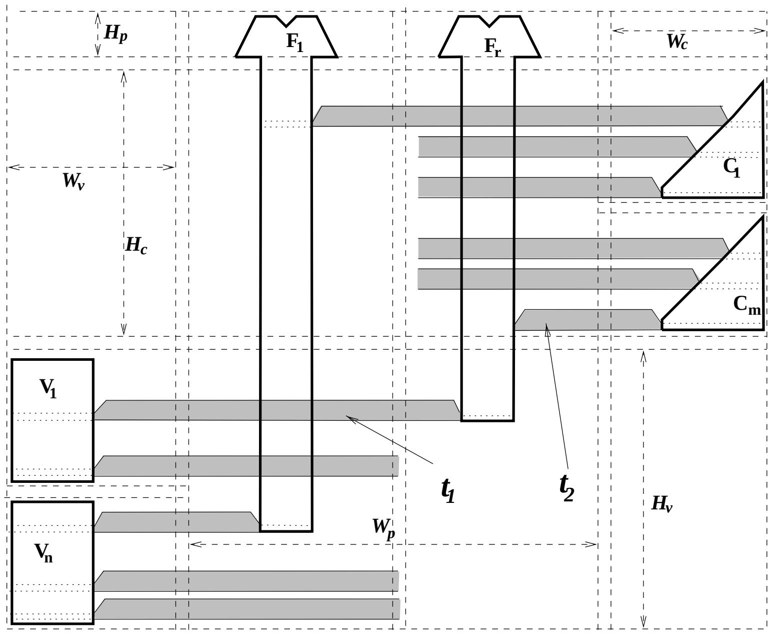


Figure 3: Layout of devices on the sub-series plan (with coordinate system rotated 45° clockwise)

Each of these components (variable structures, clause checkers, shifters and connectors) consists of a set of solid points which are placed in a specific way. Points outside these components are hollow points. Furthermore, for any two components, there exists at least one layer of hollow points to separate them, except that a non-connector component may have several connectors "touching" its boundary without a separating layer. Therefore, a candidate trapezoid cannot cover solid points from two non-connector components at the same time if it does not cover any solid points from a connector, if exists, between them. Hence we divide all candidate trapezoids into two categories, namely **background trapezoids**, each covering solid points only from a single component, and **connecting trapezoids**, each covering solid points from more than one components[2]. The construction process, however, will ensure that a connecting trapezoid must cover some solid points from exactly one connector and from one or two non-connector components connected via that connector. *Note that a connecting trapezoid is a candidate trapezoid used to cover solid points, while a connector is a component consisting of some specially arranged solid points.*

[2]However, there is one exception. A special set of trapezoids, each covering solid points from only one "splitter" component, are called inner connecting trapezoids and are counted as connecting trapezoids instead of background trapezoids (see Section 2.2.2 for details).

The constructed structure intuitively works as follows: If a variable v_i has r_i^+ positive occurrences and r_i^- negative occurrences in the collection $\mathcal{C}$ of clauses, its corresponding variable structure V_i has r_i^+ "positive" connectors and r_i^- "negative" connectors and hence connects with $r_i^+ + r_i^-$ distinctive variable occurrence shifters, which in turn connect with $r_i^+ + r_i^-$ (not necessarily distinct) clause checkers. When constrained appropriately (as will be discussed later), in order to obtain a covering instance of J, either none of the connecting trapezoids which covers solid points from positive connectors can be selected, or none of the connecting trapezoids which covers solid points from negative connectors can be selected. This mimics the truth assignment: It is impossible to have both the positive and the negative occurrences of the same variable assigned true. Furthermore, the selection of connecting trapezoids "propagates" from variable structures to clause checkers by variable occurrence shifters. Finally, a checker can be covered by an appropriate number of trapezoids if at least one of the connecting trapezoids is selected. Under such a setting, a satisfying truth assignment for I implies a covering instance for J and vice versa.

In the remainder of this section, we will first study each type of component in detail, then construct the overall structure from a 3SAT instance, and finally prove that the constructed structure has a covering instance if and only if the 3SAT instance is satisfiable.

2.1 Connectors

A **connector** consists of two rows[3] of solid points, where the bottom row has one more solid point than the top row. Figure 4 illustrates a connector which connects components X and Y (The boundaries of the connector and both components are shown as dashed lines, and components X and Y are not shown in detail except for one row). The connector consists of all the solid points within the shadow area, and can be covered by a single candidate trapezoid (the shadow one). In addition, some neighboring points of the connector are required to be hollow, see Figure 4. Specifically, for a connector with l solid points on the top row and $l+1$ solid points on the bottom row, a specific set of $2l+3$ neighboring points are required to be hollow, namely, the $l+1$ points from the row right above the connector, the l points from the row right below the connector, and the two points at both ends of the top row of the connector. Intuitively, this means that there is one layer of hollow points around the connector, except at both ends of the bottom row. Hence, the left-most solid point of the top row (point B in Figure 4) is not visible to any solid point outside the connector due to the presence of five hollow points (points C's in Figure 4).

Similarly, a layer of surrounding hollow points is also required for both component X and component Y (except where a connector exists). In addition, different connectors will generally be placed at different rows. For any two connectors which *are* placed at the same rows, there is at least one hollow point that separates their bottom rows (their top rows will always be separated due to the required presence of the surrounding hollow points of each connector). Indeed, due to the required hollow points surrounding components X and Y, there are two hollow points (from those that surround components X and Y, i.e., points A and D in Figure 4) at both ends of the bottom row of the connector when extended to the boundaries of X and Y. It is easily seen that, for each connector, there exists one and only one maximum connecting trapezoid (the solid rectangle in Figure 4) that covers solid points from it. This unique maximum connecting trapezoid does not cover solid points from any other connectors due to the hollow points mentioned above. Actually, the maximum connecting trapezoid of a connector is a

[3]We talk about rows or columns in the new coordinate system after rotation, hence *points in two neighboring rows or columns are not aligned*, see figure 1.

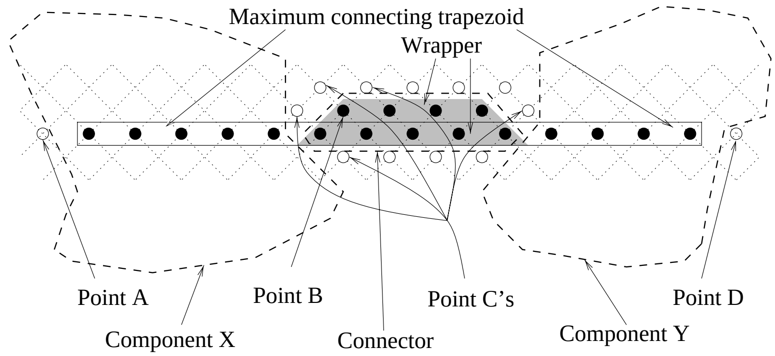


Figure 4: Connector

degenerated trapezoid (a line) which covers the bottom row of the connector, as well as other solid points from the two components which are connected via the connector.

The construction process in the rest of this paper will ensure that any two solid points from two non-connector components respectively are invisible to each other if the two points cannot be covered by a maximum connecting trapezoid, if exists, between the two components. Hence non-connector components "interfere" with each other only via connectors. Given a connector T connecting a component X with some other component, for any connecting trapezoid t that covers at least a solid point of T, we will abuse the language and call t a **connecting trapezoid of X (from T)**, even if it does not cover any solid point of X at all (then it must cover some solid point(s) of the other component, since t is not an inner connecting trapezoids of a splitter component, and hence has to cover points from at least two components).

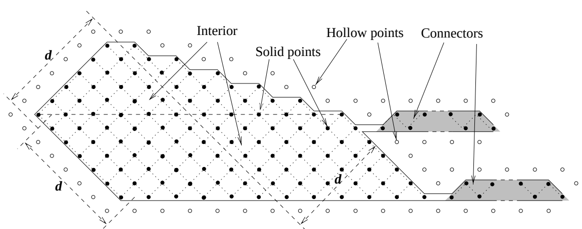
Every connector itself can be covered by a single background trapezoid, and this trapezoid is called the **wrapper** of the connector. Clearly, when the wrapper of a connector is selected, no connecting trapezoid of the connector is needed to cover the connector itself. On the other hand, the top row of each connector has at least one special solid point (e.g., B in Figure 4, actually any point in the top row is eligible) which is invisible to any solid point outside the connector. Hence, for each connector, *the wrapper is both necessary and sufficient in order to obtain a covering instance* for the solid points in the connector itself. A connector can be of arbitrary length, yet a longer connector does not require more candidate trapezoids than a shorter one does in order to obtain a covering instance.

2.2 Variable Structures

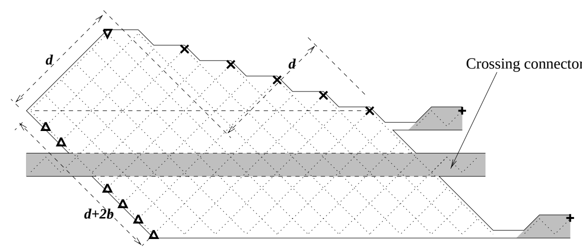
Variable structures (e.g., V_1, ..., V_n in Figure 3) are complex structures, consisting of more basic components, namely inverters and splitters. A **variable structure** for a variable which has $r^+ = 3$ positive occurrences and $r^- = 2$ negative occurrences (in the clauses) is illustrated in Figure 7(a). It has $r^+ + r^- + 1$ inverters $I_1^+, \ldots, I_{r^+}^+$, $I_1^-, \ldots, I_{r^-}^-$, I_0, and two splitters S^+ and S^-. Each inverter connects to two other components via two connectors (there may be some other connectors going through an inverter). Roughly speaking, if constrained appropriately, one and only one connecting trapezoid of an inverter can be selected in order to obtain a covering instance. Each splitter connects to other components via $f + 1$ connectors, where $f = r^+$ for $S+$ and r^- for S^-. The bottom-right connector of the splitter is the "input" while the f top-left connectors are the "outputs". Roughly speaking again, the connecting trapezoids of the input and the output connectors of a splitter shall be selected or dropped together. Hence, a variable structure intuitively works as follows: I_0 first generates two opposite "signals", the two splitters then split the signals into multiple copies for all variable occurrences. The other inverters "transmit" the appropriate signal to the "outside".

2.2.1 Inverters

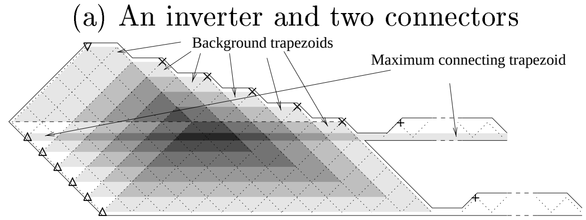
We now show the inverters. Figure 5(a) illustrates a basic **inverter** of **degree** $d = 6$. The (diagonally) dotted line segments are aligned to $x = i$ or $y = j$ where i's and j's are integers. The dashed lines are used to help clarifying the construction of the inverter. As shown in Figure 5(a), all points on the intersections of dotted line segments are solid points. There is also at least one layer of surrounding points which are hollow points. We always *require at least a layer of surrounding hollow points for any constructed components* and will not mention this later. Solid points in the shadow regions in Figure 5(a) are two connectors. The **interior** of the inverter is the set of all other solid points shown in Figure 5(a). The interior can be covered with $d + 1$ candidate trapezoids in two ways, as illustrated in Figure 5(c) and 5(d). In each case, exactly one maximum connecting trapezoid (the line extending into a connector and the component on the other end of the connector) is included.



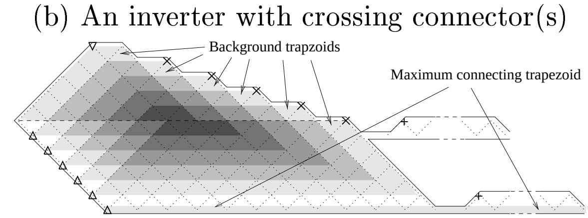
(a) An inverter and two connectors



(b) An inverter with crossing connector(s)



(c) Covering an inverter with d background trapezoids and the top connecting trapezoid



(d) Covering an inverter with d background trapezoids and the bottom connecting trapezoid

Figure 5: Inverter

Intuitively, if we are required to cover the interior of an inverter of degree d with at most $d + 1$ trapezoids, the inverter works as a NOT operator: if one of the two maximum connecting trapezoids is selected, the other must be dropped. Thus an inverter mimics a variable such that either its positive occurrences or its negative occurrences can be true, but not both.

The basic inverter of degree d can be extended to allow $b \geq 0$ connectors to *cross* (i.e. go through) its body. An inverter of degree $d = 6$ with $b = 1$ **crossing connector** is illustrated in Figure 5(b). Crossing connectors are placed between the two normal connectors. Intuitively, for the $d - 3$ rows between the two normal connectors, we can choose any two consecutive rows, split the inverter between these two rows, and insert a crossing connector (i.e., two layers of solid points as far as the inverter is concerned). Hence, $b \leq d - 4$ is required, i.e., there can be at most $d - 4$ connectors going through an inverter of degree d. Similarly, the **interior** is defined as the solid points not belong to any connector. Note that the interior of an inverter may not be continuous and the solid points of crossing connectors are *not* regarded as the interior. The property mentioned above still holds for this extended form of inverters. Formerly:

Lemma 1 *Given an inverter I of degree d with b crossing connectors,*

(1) the interior of I has the width of $O(d+b)$ and the height of $O(d+b)$;

(2) at least $d+1$ (background or connecting) trapezoids are necessary to cover the interior of I;

(3) at least d background trapezoids are necessary to cover the interior of I regardless the selection of its connecting trapezoids;

(4) d background trapezoids are sufficient to cover the interior of I, if at least one of the two (non-crossing) maximum connecting trapezoids of I is selected.

Proof. First we mark some special (solid) points in the interior of an inverter of degree d with b crossing connectors, as illustrated in Figure 5(b). The left-most point of the top-most row is marked with "$\triangledown$". For each of the $d-1$ consecutive rows under the top-most row, the right-most point is marked with "$\times$". Below these d rows, there are $d+2b$ rows including $2b$ rows for b crossing connectors. For each of the d non-crossing rows, the left-most points is marked with "$\triangle$".

(1) From Figure 5(b), it is clear that the interior of I has $2d+2b$ rows and $5d+2b-1$ columns of points. Since the coordinate system is rotated 45°, the interior has the height of $\frac{\sqrt{2}}{2}(2d+2b-1)$ and the width of $\frac{\sqrt{2}}{2}(5d+2b-2)$.

(2) To see that at least $d+1$ trapezoids are necessary, we observe that the $d+1$ points marked with "$\triangledown$" or "$\triangle$" in Figure 5 are pairwise invisible, and each of these points demands a unique trapezoid to cover.

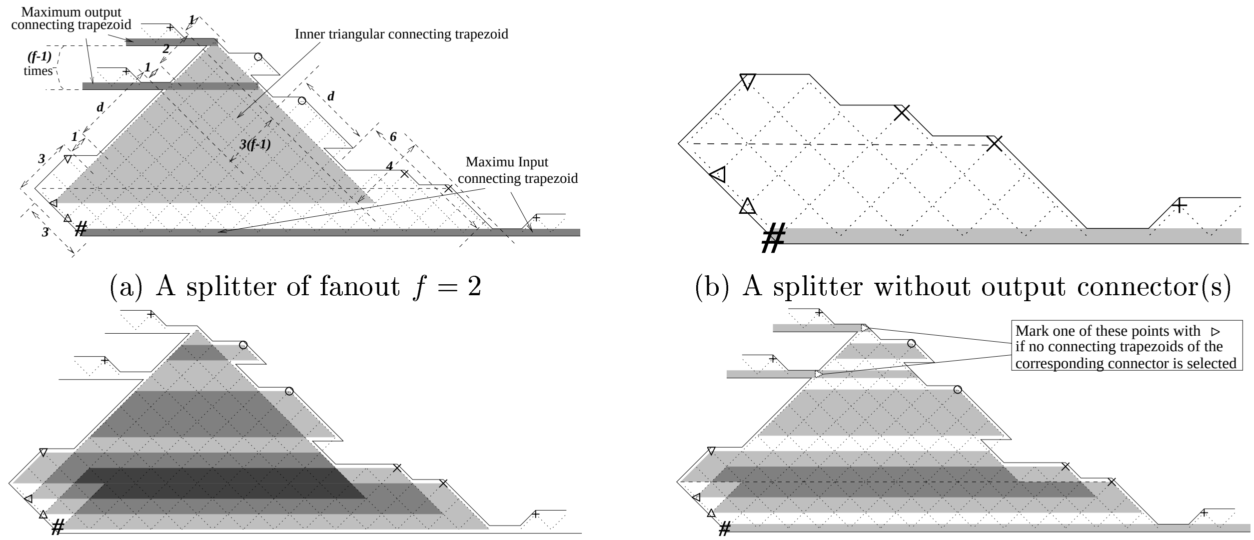
(3) To see that at least d background trapezoids are necessary, we observer that the d points marked with "$\triangledown$" or "$\times$" in Figure 5 are pairwise invisible, and none of them can be covered by any connecting trapezoid.

(4) This is illustrated in Figure 5(c) and (d). $\square$

In the following, we will use $\mathtt{Bg}\#(X)$ to denote the minimum number of background trapezoids needed regardless the selection of connecting trapezoids to cover the interior of a non-connector component X. From Lemma 1(3), for an inverter I of degree d, we have $\mathtt{Bg}\#(I) = d$.

2.2.2 Splitters

The **(variable occurrence) splitter**, as illustrated in Figure 6, is designed as a fanout device to mimic the possible multiple positive or negative occurrences of the same variable. We prefer to think of the bottom-right connector in Figure 6(a) as the input signal, and the two (in this case) top-left connectors as the output signals. The number of output connectors is the **fanout** of the splitter. Note that the bottom six rows of every splitter are similar to an inverter, but with only one connector. In general, a splitter of fanout f and degree d is constructed as follows: First construct a special triangle of $3f{+}d{+}2$ rows (shown as the shadow triangle in Figure 6(a)) and the bottom six rows of the splitter in the way illustrated in Figure 6(a). (For subsequent reference, we mark six special points from the bottom six rows with "$\triangledown$", "$\triangle$","$\triangleleft$", "#" or "$\times$", as shown in Figure 6.) Then we place f output connectors to the top-left of the triangle such that the bottom row of the i-th connector (counted from top down) is aligned to the $(3i{-}2)$-th row (again, counted from top down) of the triangle. For every two consecutive output connectors, there are two rows between their bottom rows (exclusively). For each of these two rows, we add an extra solid point at the right of the triangle, and mark the extra point for the upper row of these two rows with "$\circ$" for subsequent reference. Altogether there are $f{-}1$ points marked with "$\circ$" in this way (since there are f output connectors and $f-1$ pairs of consecutive output connectors). Finally, there are d rows between the bottom-most output connector and the bottom six rows of the splitter (exclusively). For these d rows, we add an extra solid point for each row to the right of the triangle, and similarly we mark the extra solid point of the top-most one of these d rows with "$\circ$".



(a) A splitter of fanout $f = 2$

(b) A splitter without output connector(s)

(c) Covering a splitter with the inner (triangular) connecting trapezoid

(d) Covering a splitter with the maximum input connecting trapezoids and all the maximum output connecting trapezoids

Figure 6: Splitter

Again, the **interior** of a splitter consists of the solid points outside of both the input and the output connectors. The shadow triangle shown in Figure 6(a) is called the **inner (triangular) connecting trapezoid** since it "connects" the input and output connecting trapezoids[4], as discussed shortly. In the following, *we will count an inner connecting trapezoid as a connecting trapezoid instead of a background trapezoid* (this is the only exception for the definition of background trapezoids and connecting trapezoids, and it does not change the proof of Lemma 1 since inner connecting trapezoids only exist in splitters).

We also require that $d > 0$ whenever $f > 0$. In the special case where $f = d = 0$, as shown in Figure 6(b), the interior of the splitter can be covered by 3 candidate trapezoids plus the only connecting trapezoid. Thus we have:

[4]By "input (output) connecting trapezoid", we actually mean the connecting trapezoid from the input (output) connector. In the following discussion, we will follow this convention without formally introducing new definitions.

Lemma 2 *Given a splitter S of fanout f and degree d,*

(1) the interior of S has width of $O(d+f)$ and height of $O(d+f)$.

(2) at least $f+4$ (background, input or inner connecting) trapezoids are necessary to cover the interior of S regardless the selection of its output connecting trapezoids;

(3) if there exists an output connector none of whose connecting trapezoids is selected, no input connecting trapezoid can be selected in order to cover the interior of S with at most $f+4$ trapezoids;

(4) at least $f+3$ background trapezoids are necessary to cover the interior of S regardless the selection of its connecting trapezoids;

(5) $f+3$ background trapezoids are sufficient to cover the interior of S, if either (a) the maximum input connecting trapezoids and all the maximum output connecting trapezoids are selected, or (b) the inner connecting trapezoid is selected.

Proof. (1) From Figure 6(a), it is clear that the width is $\frac{\sqrt{2}}{2}(2d+6f+11)$ and the height is $\frac{\sqrt{2}}{2}(3f+d+3)$ when $f>0, d>0$. If $f=d=0$, the width and height are both constant from Figure 6(b).

(2) Consider the $f+4$ points marked with "$\circ$", "$\triangledown$", "$\triangle$", "$\triangleleft$" or "$\#$". They are pairwise invisible and none of them can be covered by any output connecting trapezoid.

(3) Consider one output connector such that none of its connecting trapezoids is selected. For the bottom row of this output connector, mark the left-most point of the inner connecting trapezoid at this row with "$\triangleright$" (see Figure 6(d)). Consider the $f+4$ points marked with "$\circ$", "$\triangledown$", "$\triangleright$" or "$\times$": they are pairwise invisible and none of them can be covered by any input connecting trapezoid. Hence, if an input connecting trapezoid were selected, at least $f+4+1=f+5$ trapezoids would have to be used to cover the interior of S. Therefore, no input connecting trapezoid can be used.

(4) Consider the $f+3$ points marked with "$\circ$", "$\triangledown$" or "$\times$": they are pairwise invisible and cannot be covered by any connecting trapezoid.

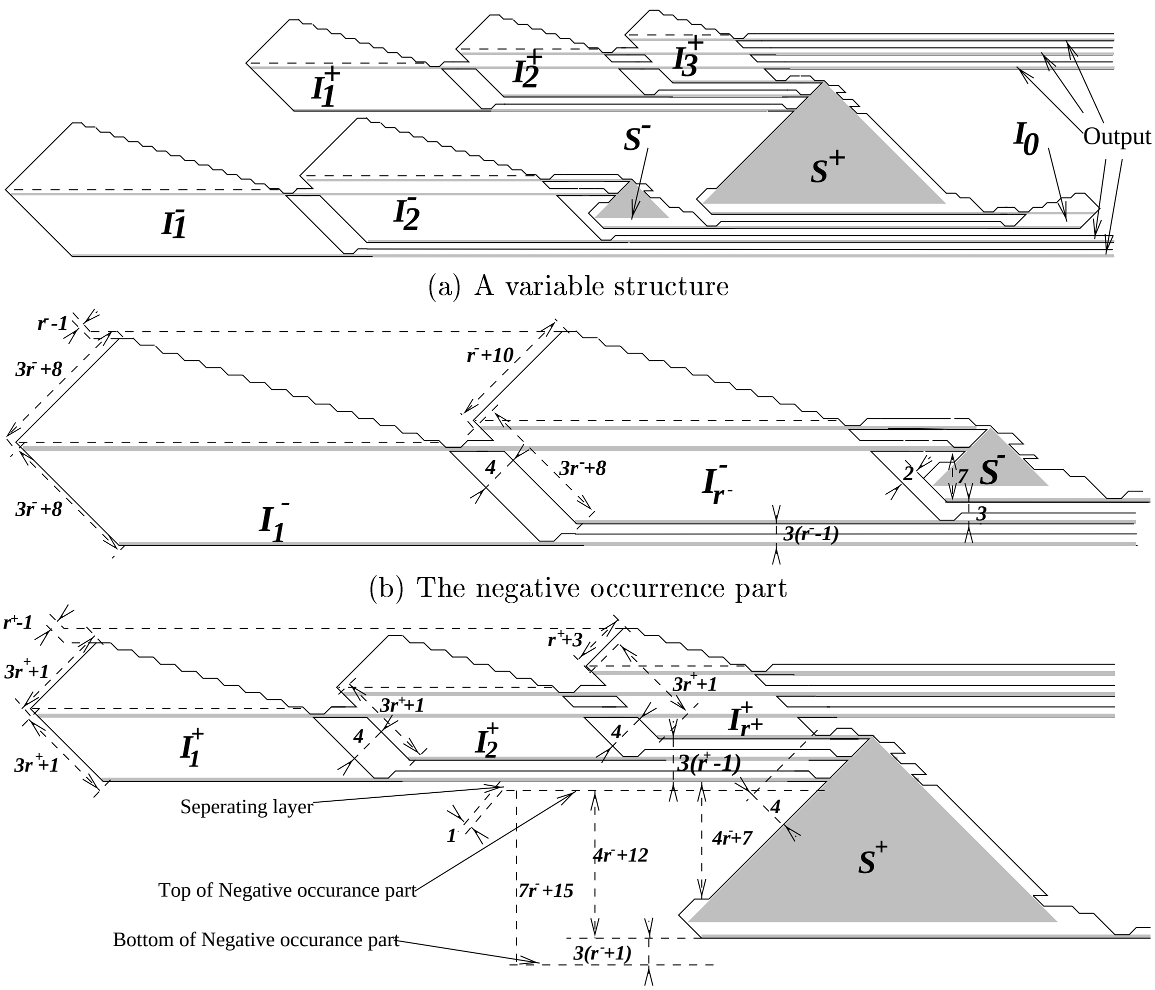
(5) The sufficiency is illustrated in Figure 6(c) and (d).

$\square$

From Lemma 2(4), $\mathtt{Bg}\#(S)=f+3$ where S is a splitter of fanout f. Furthermore, if we are allowed to use only $\mathtt{Bg}\#(S)$ background trapezoids plus *only one* of the inner or input connecting trapezoids to cover S, then the output "depends on" the input. Indeed, by Lemma 2(3), if an input connecting trapezoid is selected, i.e., the inner connecting trapezoid is not selected, then at least one connecting trapezoid for each output connector must be selected.

2.2.3 Assemble variable structures with inverters and splitters

We are now ready to construct a variable structure, as illustrated in Figure 7. Suppose the variable we wish to represent has r^+ positive occurrences and r^- negative occurrences in the clauses of $\mathcal{C}$. The variable structure consists of three parts: one for the positive occurrences, one for the negative occurrences, and a basic inverter I_0 of degree 4 connecting the above two parts. The positive occurrence part has a splitter S^+ of fanout r^+, and r^+ inverters $I_1^+, \ldots, I_{r^+}^+$. Similarly the negative occurrence part has a splitter S^- of fanout r^-, and r^- inverters $I_1^-, \ldots, I_{r^-}^-$. These components are placed, as illustrated in Figure 7, such that (1) the bottom connector of I_i^+ is the $(r^+{-}i{+}1)$-th output connector of S^+ (counted from top down); (2) the top connector of I_i^- is the $(r^-{-}i{+}1)$-th output connector of S^-; (3) the input connector of S^+ is the top connector of I_0; (4) the input connector of S^- is the bottom connector of I_0; (5) the top connector of I_i^+ goes through $I_{i+1}^+, \ldots, I_{r^+}^+$; and (6) the top connector of I_i^- goes through $I_{i+1}^-, \ldots, I_{r^-}^-$. Hence, other parameters of these inverters and splitters are specified as the following: I_i^+ has $i-1$ crossing connectors and is of degree $3r^+{-}2i{+}3$ for $1{\le}i{\le}r^+$. I_i^- has $i{-}1$ crossing connectors and is of degree $3r^-{-}2i{+}10$ for $1{\le}i{\le}r^-$. The extra seven rows in I^-'s are due to the splitter S^-, whose degree is one when $r^- > 0$, zero otherwise. The degree of S^+ is set to $4r^-{+}7$ to make sure the input connectors of S^+ and S^- can be seamlessly connected to the inverter I_0. The detail is omitted here for simplicity.



(a) A variable structure

(b) The negative occurrence part

(c) The positive occurrence part

Figure 7: Variable Structure

The upper r^+ connectors from the positive occurrence part, i.e., the top connectors of $I_1^+, \ldots, I_{r^+}^+$, are called **positive output connectors**, while the bottom r^- connectors from the negative occurrence part, i.e., the bottom connectors of $I_1^-, \ldots, I_{r^-}^-$, are called **negative output connectors**. All of them may extend arbitrarily long, and are called **output connectors** for the variable structure. Other $r^+ + r^- + 2$ connectors are fixed at both sides within the variable structure, and are called **inner connectors** for the variable structure. The interior of the variable structure consists of the solid points that are not in *any* connector (including inner connectors).

Lemma 3 *Given a variable structure V with r^+ positive output connectors and r^- negative output connectors,*

(1) the interior of V has the height of $O(r^+ + r^-)$ and the width of $O((r^+)^2 + (r^-)^2)$;

(2) at least[5] $\mathtt{Bg\#}(V) = 2(r^+)^2 + 2(r^-)^2 + 3r^+ + 10r^- + 10$ *background trapezoids are necessary to cover the interior of V regardless of the connecting trapezoids of V.*

(3) at least $\mathtt{Bg\#}(V) + (r^+ + r^- + 2)$ (background and connecting) trapezoids are necessary to cover the interior of V;

(4) to cover the interior of V with $\mathtt{Bg\#}(V) + (r^+ + r^- + 2)$ trapezoids, either (a) none of the positive output connecting trapezoids of V is selected, or (b) none of the negative output connecting trapezoids of V is selected.

(5) it is sufficient to cover the interior of V with $\mathtt{Bg\#}(V)$ background trapezoids if either (a) the maximum input connecting trapezoid and all the r^+ maximum output connecting trapezoids of S^+, the inner connecting trapezoid of S^-, and the r^- maximum negative output connecting trapezoids of V are selected, or (b) the maximum input connecting trapezoid and all the r^- maximum output connecting trapezoids of S^-, the inner connecting trapezoid of S^+, and the r^+ maximum positive output connecting trapezoids of V are selected;

Proof. (1) See Figure 7(b) and (c). Details are omitted here.

(2) First, if we mark special points in each inverter and each splitter in the same way as we do previously, then any two points marked with "$\triangledown$", "$\times$" or "$\circ$" in different components are invisible to each other. Hence the background trapezoids required for one component cannot cover points marked with "$\triangledown$", "$\times$" or "$\circ$" in any other component, and the total number of minimum required background trapezoids is

$$\begin{aligned}
\mathtt{Bg\#}(V) &= \sum_{i=1}^{r^+} \mathtt{Bg\#}(I_i^+) + \sum_{i=1}^{r^-} \mathtt{Bg\#}(I_i^-) + \mathtt{Bg\#}(S^+) + \mathtt{Bg\#}(S^-) + \mathtt{Bg\#}(I_0) \\
&= \sum_{i=1}^{r^+} (3r^+ + 3 - 2i) + \sum_{i=1}^{r^-} (3r^- + 10 - 2i) + (r^+ + 3) + (r^- + 3) + 4 \\
&= r^+(3r^+ + 3) - 2\frac{r^+(1+r^+)}{2} + r^-(3r^- + 10) - 2\frac{r^-(1+r^-)}{2} + r^+ + r^- + 10 \\
&= 2(r^+)^2 + 2(r^-)^2 + 3r^+ + 10r^- + 10.
\end{aligned}$$

(3) We partition all the background and connecting trapezoids of V into $r^+ + r^- + 3$ groups as follows: for inverter I_0, put all its background trapezoids into a unique group $\mathcal{G}_0^I$; for inverter I_i^+ $(1 \le i \le r^+)$,

[5]The use of the symbol $\mathtt{Bg\#}(V)$ is justified by (5) below.

put all its background and connecting trapezoids into a unique group $\mathcal{G}_i^I$; for inverter I_i^- $(1 \le i \le r^-)$, put all its background and connecting trapezoids into a unique group $\mathcal{G}_{-i}^I$; for splitter S^+, put all its background trapezoids, as well as input and inner connecting trapezoids (the output connecting trapezoids, which are also connecting trapezoids of I_i^+'s and I_i^-'s, are already in the previous groups), into a unique group $\mathcal{G}_+^S$; and for splitter S^-, put all its background trapezoids, as well as input and inner connecting trapezoids (excluding output connecting trapezoids), into a unique group $\mathcal{G}_-^S$.

Clearly, from Lemma 1(3), at least $\mathtt{Bg\#}(I_0)$ trapezoids from $\mathcal{G}_0^I$ are required to cover I_0; from Lemma 1(2), at least $\mathtt{Bg\#}(I_i^+)+1$ trapezoids from $\mathcal{G}_i^I$ are required to cover I_i^+, and at least $\mathtt{Bg\#}(I_i^-)+1$ trapezoids from $\mathcal{G}_i^I$ are required to cover I_i^-; from Lemma 2(2), at least $\mathtt{Bg\#}(S^+)+1$ trapezoids from $\mathcal{G}_+^S$ to cover S^+, and at least $\mathtt{Bg\#}(S^-)+1$ trapezoids from $\mathcal{G}_-^S$ to cover S^-. In summary, at least $\mathtt{Bg\#}(V)+r^++r^-+2$ trapezoids are necessary, since $\mathtt{Bg\#}(V) = \sum_{i=1}^{r^+}\mathtt{Bg\#}(I_i^+) + \sum_{i=1}^{r^-}\mathtt{Bg\#}(I_i^-) + \mathtt{Bg\#}(S^+)+\mathtt{Bg\#}(S^-)+\mathtt{Bg\#}(I_0)$ (from the proof of (2)).

(4) Partition all background and connecting trapezoids of V as above. Clearly, from the proof of (3) above, the number $\mathtt{Bg\#}(V)+r^++r^-+2$ is tight. That is, in order to cover V with $\mathtt{Bg\#}(V)+r^++r^-+2$ trapezoids, exactly $\mathtt{Bg\#}(I_0)$ trapezoids has to be selected from $\mathcal{G}_0^I$; exactly $\mathtt{Bg\#}(I_i^+)+1$ trapezoids has to be selected from $\mathcal{G}_i^I$ for $1 \le i \le r^+$; exactly $\mathtt{Bg\#}(I_i^-)+1$ trapezoids has to be selected from $\mathcal{G}_{-i}^I$ for $1 \le i \le r^-$; exactly $\mathtt{Bg\#}(S^+)+1$ trapezoids has to be selected from $\mathcal{G}_+^S$; and exactly $\mathtt{Bg\#}(S^-)+1$ trapezoids has to be selected from $\mathcal{G}_-^S$.

If both conditions (a) and (b) are violated, there is a selected positive output connecting trapezoid, denoted t^+, and a selected negative negative output connecting trapezoid, denoted t^-. Consider t^+ first. Without loss of generality, let t^+ be a connecting trapezoid of I_{i0}^+ from connector T_1, hence $t^+ \in \mathcal{G}_{i_0}^I$. Denote the other connector of I_{i0}^+ as T_2. Since exactly $\mathtt{Bg\#}(I_{i0}^+)+1$ trapezoids can be selected from $\mathcal{G}_{i_0}^I$, and at least $\mathtt{Bg\#}(I_{i0}^+)$ background trapezoids (also from $\mathcal{G}_{i0}^I$) are necessary to cover I_i^+(from Lemma 1(3)), then at most *one* connecting trapezoid from $\mathcal{G}_{i_0}^I$ can be selected. Now t^+ is selected, no connecting trapezoid of I_{i0}^+ from T_2 can be selected. Note that T_2 is also an output connector of S^+, thus S^+ has an output connector none of whose connecting trapezoids is selected. Again, exactly $\mathtt{Bg\#}(S^+)+1$ trapezoids can be selected from $\mathcal{G}_+^S$, therefore from Lemma 2(3), no input connecting trapezoid of S^+ can be selected. As the input connector of S^+ is also the top connector of I_0, hence, none of the connecting trapezoids of I from its top connector can be selected due to the selection of r^+. Similarly, consider t^-, none of the connecting trapezoids of I from its bottom connector can be selected due to the selection of r^-. Since only $\mathtt{Bg\#}(I_0)$ trapezoids can be selected from $\mathcal{G}_0^I$, from Lemma 1(2), I_0 cannot be covered.

(5) From Lemma 1(4) and Lemma 2(5), if either (a) or (b) is satisfied, the interior of V, i.e., the interiors of all inverters and splitters, can be covered.

□

By the above lemma, when we are allowed to cover a variable structure with only the minimum number of background and connecting trapezoids given in the lemma, the variable structure mimics the truth assignment of the multiple occurrences of the same variable: If one of the positive output connecting trapezoids is selected, none of the negative output connecting trapezoids can be selected, and vice versa.

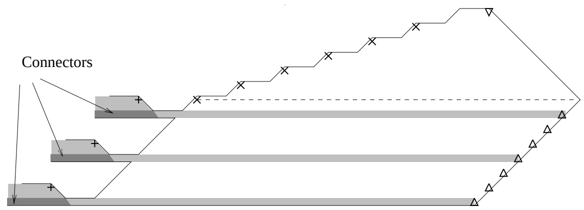


Figure 8: Clause checker

2.3 Clause checkers and variable occurrence shifters

The next device we need is a **(clause) checker**, as illustrated in Figure 8. A checker is similar to a basic inverter of degree seven. The checker has three connectors. There are no parameters for checkers. Similar to the inverter, we have:

Lemma 4 *Given a clause checker C,*

(1) the interior of C has a constant width and height;

(2) at least seven background trapezoids are necessary to cover the interior of C regardless the selection of its connecting trapezoids;

(3) a connecting trapezoid of C from one of the three connectors must be selected in order to cover the interior of C with at most seven background trapezoids;

(4) it is sufficient to cover the interior of C with seven background trapezoids, if at least one of the three maximum connecting trapezoids is selected.

Proof. The proof is similar to that of Lemma 1 and is omitted here. □

The last device, **variable occurrence shifter**, illustrated in Figure 9, is used to shift the output connectors of variable structures to the positions which are aligned to the input connectors of appropriate clause checkers. A shifter of degree d with b crossing connectors consists of two almost reflexive parts. Each part is actually an inverter of degree d with $b+1$ crossing connectors. The interior of these two inverters are $6 \times \frac{\sqrt{2}}{2}$ apart (horizontally). The left bottom connector is called the **input connector**. The input connector is a crossing connector for the left part inverter and a regular connector for the right part inverter. Another connector, which is a crossing connector for the right inverter and a regular connector for the left part, is called the **output connector**. The top connector, which is a regular connector for both inverters, is totally within the shifter and is called the **inner connector**. Other connectors are called **crossing connector** for the shifter. The interior of the shifter consists of the points not in any (even the inner!) connector. The interior of a shifter is divided into two parts: the top part is called **head** and the bottom part called **body**. They are separated by the inner connector, which belongs to the head.

Again, some special points are marked: on the top-most row, the left-most point from the left inverter and the right-most point from the right inverter are marked with "∇"; on each of the following $d-1$ rows, the right-most point from the left inverter and the left-most point from the right inverter

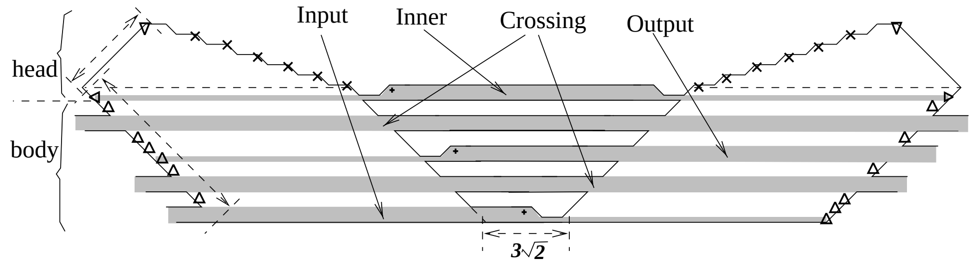


Figure 9: A variable occurrence shifter of degree 6 with 2 crossing connectors

are marked with "×"; on the next row (the bottom row of the inner connector), the left-most point is marked with "◁" and the right-most point is marked with "▷"; on each of the following $d+2b+1$ rows, the left-most point from the left inverter is marked with "△" if it is in a connector of the left inverter, also the right-most point from the right inverter is marked with "△" if it is not in a connector of the right inverter.

Lemma 5 *Given a shifter F of degree d with b crossing connectors,*

(1) both the height and the width of its interior are bounded by $O(H_b)$, where H_b is the height of its body, and $H_b = O(d+b)$;

(2) at least $2d+1$ (background, inner or output connecting) trapezoids are necessary to cover the interior of F regardless the selection of its input connecting trapezoids;

(3) if no input connecting trapezoid of F is selected, none of its output connecting trapezoids can be selected in order to cover its interior by at most $2d+1$ trapezoids;

(4) $2d$ background trapezoids are sufficient to cover the interior of F if either (a) both the maximum input connecting trapezoid and the maximum output connecting trapezoid are selected, or (b) the maximum inner connecting trapezoid is selected.

(5) at least $2d$ background trapezoids are necessary to cover the interior of F regardless of its connecting trapezoids.

Proof. (1) From Lemma 1, the height is $\frac{\sqrt{2}}{2}(2d+2b+1)$ and the width is $\frac{\sqrt{2}}{2}(10d+4b+2)$. The calculation is straightforward given that the inverters in the shifter has $b+1$ crossing connectors and are $6 \times \frac{\sqrt{2}}{2}$ apart. On the other hand, $H_b = \frac{\sqrt{2}}{2}(d+2b+1)$. Hence, the height and the width are both bounded by $O(H_b)$.

(2) Consider the $2d+1$ points, namely (i) the $d+1$ points marked with "▽", "△" or "◁" from the left inverter, and (ii) the d points marked with "▽" or "×" from the right inverter. These $2d+1$ points are pairwise invisible, and none of them can be covered any input connecting trapezoids.

(3) If no input connecting trapezoid is selected, consider the $2d+1$ points, namely (i) the $d+1$ points marked with "▽", "△" or "▷" from the right inverter, and (ii) the d points marked with "▽" or "×" from the left inverter. These $2d+1$ points are pairwise invisible, and none of them can be covered by any output connecting trapezoids.

(4) If either condition is satisfied, from Lemma 1(4), each inverter can be covered by d background trapezoids (in addition to the selected connectors), hence the shifter can be covered.

(5) Consider the $2d$ points marked with "$\triangledown$" or "$\times$", they are pairwise invisible, and none of them can be covered by any connecting trapezoid. □

From Lemma 5, $\texttt{Bg\#}(F) = 2d$ where F is a shifter of degree d. Furthermore, if we are allowed to cover the interior of F with only $\texttt{Bg\#}(F)$ background trapezoids and either an inner or an output connecting trapezoid, an input connecting trapezoid must be selected whenever an output connector is selected.

2.4 Proof of Theorem 1

We are now ready for the main theorem. For any instance I of 3SAT, we first construct an instance J of BCT, and then prove that any satisfying truth assignment of I implies a covering of J, and vice versa.

2.4.1 Construction

For any instance $I = (\mathcal{U}, \mathcal{C})$ of 3SAT, we can construct an instance $J = (B, K)$ of BCT using devices described above. Let $n = |\mathcal{U}|$, $m = |\mathcal{C}|$ and r be the total number of variable occurrences ($r = 3m$ and $r \geq n$). Let $v_1, \ldots, v_n$ be a permutation of $\mathcal{U}$ and $c_1, \ldots, c_m$ be a permutation of $\mathcal{C}$. Let $f_1, \ldots, f_r$ be a permutation of all occurrences of all variables such that $f_{3i-2}, f_{3i-1}, f_{3i}$ are the variable occurrences in the clause c_i. For each variable v_i ($1 \leq i \leq n$), we construct a unique variable structures V_i; for each clause c_i ($1 \leq i \leq m$), we construct a unique clause checker C_i; and for each variable occurrence f_i ($1 \leq i \leq r$) of variable v_{j_1} in clause c_{j_2}, we construct a unique variable occurrence shifters F_i to associate V_{j_1} with C_{j_2} as illustrated in Figure 3. This is done in the following way:

Firstly, we construct all variable structures. Specifically, the variable structure V_n for v_n is first constructed according to the number of positive and negative occurrence of v_n in all clauses, and is put on the left bottom. The variable structure V_i is then constructed in the same way according to v_i, and is put on the top of V_{i+1} allowing four separating rows between them. From Lemma 3(1), all variable structures $V_1, \ldots, V_n$ together have the height $H_v = O(r)$ and width $W_v = O(r^2)$. Also, there are totally r output connectors, $r + 2n$ inner connectors (since each variable structure V_i has $r_i^+ + r_i^-$ output connectors and $r_i^+ + r_i^- + 2$ inner connectors).

Secondly, we construct all clause checkers. Specifically, the clause checker C_m for c_m is first constructed and put on the top of the variable structure V_1 allowing four separating rows, but on the right. The clause checker C_i for c_i is then constructed and put on the top of C_{i+1} allowing four separating rows. The horizontal distance between all the variable structures and the checkers will be the width of all the shifters (which are put between them as described below) plus eight columns for separating shifters from both sides (four columns for each side). From Lemma 4(1), the height for all the checkers is $H_c = O(m) = O(r)$ and the width is $W_c = O(1)$. Also, there are $3m (= r)$ connectors. The construction is valid, since the width of all the shifters can be calculated precisely regardless the horizontal positions of clause checkers by studying the following step, and the vertical positions are fixed in this step.

Thirdly, we connect the variable structures with the clause checkers by constructing shifters. Specifically, the shifter F_1 for f_1 is first constructed and placed on the right of all variable structures allowing four separating columns. The shifter is constructed so that the inner connector is aligned at the top of the checker for c_1, the input connector (the left bottom connector) is the same as the output connector for f_1 (an variable occurrence) from the corresponding variable structure, and the output connector is the same as the connector for f_1 from the corresponding clause checker. For example, if c_1 is $v_n\overline{v_1}v_2$, then f_1 is the first occurrence of v_n, the top-most positive output connector of V_n becomes the input connector of F_1 and thus connect V_n and F_1, while the top-most input connector of C_1 becomes the output connector of F_1 and thus connect C_1 and F_1. All other output connectors from variable structures become crossing connectors of F_1 if they are on top of the input connector of F_1, otherwise, they simply go below F_1. In the same way, the shifter for f_i is constructed and put on the right of f_{i-1} allowing four separating columns. Since different output connectors of variable structures and connectors for clause checkers are on different rows and are well separated, connectors will not cross each other. The height of the body of any shifter is bounded by $(H_v + H_c) = O(r)$. From Lemma 5(1), the height and width of each shifter is bounded by $O(H_v + H_c) = O(r)$, and the width of all shifters as a whole will be $O(r \times r) = O(r^2)$. In addition, these shifters have r inner connectors. Note that in the structure constructed above, different connectors are placed in different rows, except that all inner connectors of shifters are placed in the same rows. Since shifters are separated by four columns, these inner connectors are well separated by hollow points.

Finally, we put the whole structure into the coordinate system, as illustrated in Figure 10. The overall structure will have $O(r^3)$ points. Then we find the minimum l so that all the points are within the square from the origin to the point (l, l). Clearly, l is bounded by $O(r^3)$ and can be found in $O(r^3)$ time. Now let B be the $l \times l$ matrix such that $B_{i,j} = 1$ if and only if point (i, j) is a solid point, let $K = (\sum_{i=1}^{n} \mathtt{Bg}\#(V_i) + r + 2n) + \sum_{i=1}^{m} \mathtt{Bg}\#(C_i) + (\sum_{i=1}^{r} \mathtt{Bg}\#(F_i) + r) + (4r + 2n)$. Thus the overall transform takes $O((r^3)^2) = O(r^6)$ time.

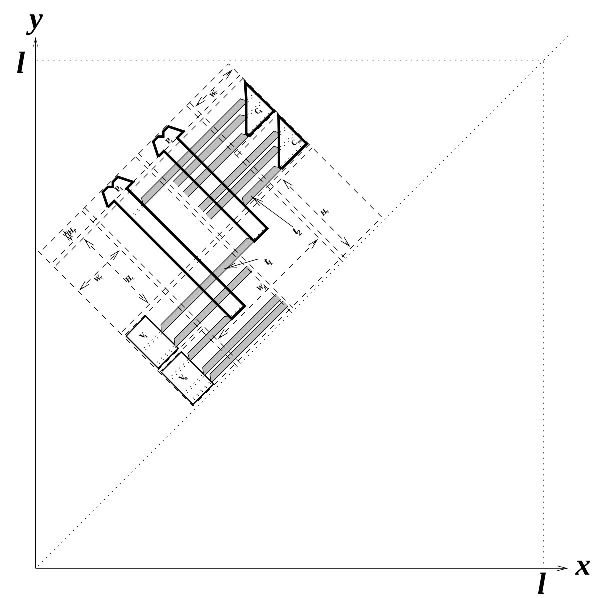


Figure 10: The bitmap

2.4.2 Equivalence

To see that a satisfying truth assignment of I implies a covering instance of J, we first select $2r+2n$ connecting trapezoids and $\sum_{i=1}^{n} \mathtt{Bg\#}(V_i) + \sum_{i=1}^{m} \mathtt{Bg\#}(C_i) + \sum_{i=1}^{r} \mathtt{Bg\#}(F_i)$ corresponding background trapezoids in the following way:

Step 1. For each variable structure V_i, if variable v_i is assigned true (false), we select all the r_i^+ maximum positive (negative) output connecting trapezoids of V, the inner connecting trapezoid from the positive (negative) splitter of V, and the maximum input connecting trapezoid and all the r_i^- maximum output connecting trapezoids from the negative (positive) splitter of V. Totally, there are $r+2n$ such connecting trapezoids for all variable structures. From Lemma 3(5), under such a selection, the interior of all variable structures can be covered by these $r+2n$ connecting trapezoids plus $\sum_{i=1}^{n} \mathtt{Bg\#}(V_i)$ background trapezoids. We hence select these $\sum_{i=1}^{n} \mathtt{Bg\#}(V_i)$ background trapezoids.

Step 2. For each shifter, we select its maximum output connecting trapezoid if its maximum input connecting trapezoid (i.e., the maximum output connecting trapezoid of some variable structure) is selected, otherwise we select its maximum inner connecting trapezoid. There are totally r such connecting trapezoids. From Lemma 5(4), under such a selection, the interior of all shifters can be covered by these r connecting trapezoids plus $\sum_{i=1}^{r} \mathtt{Bg\#}(F_i)$ background trapezoids. We also select these $\sum_{i=1}^{r} \mathtt{Bg\#}(F_i)$ background trapezoids.

Step 3. For each clause checker C_i, the corresponding clause c_i is satisfied (since the truth assignment is satisfying), i.e., either (1) a positive occurrence of a true variable appears in c_i, or (2) a negative occurrence of a false variable appears in c_i. Let the variable be v_j. In case (1), all the maximum positive connecting trapezoids of V_j are selected in Step 1. Hence the maximum output connecting trapezoids for all shifters that shift positive output connectors of V_j are selected in Step 2. According to the construction, one of these output connectors of shifters is one of the connector of C_i, hence one of the maximum connecting trapezoids of C_i is selected. From Lemma 4(4), C_i can be covered by this maximum connecting trapezoid (which is counted in Step 2, not here) plus $\mathtt{Bg\#}(C_i)$ background trapezoids. We thus select these $\mathtt{Bg\#}(C_i)$ background trapezoids. Similarly, in case (2), C_i can be covered by one maximum connecting trapezoid (which is also counted in Step 2) plus $\mathtt{Bg\#}(C_i)$ background trapezoids, and we select these $\mathtt{Bg\#}(C_i)$ background trapezoids.

Clearly all solid points from non-connector components can be covered by the above $(r+2n+r)$ connecting trapezoids and $\sum_{i=1}^{n} \mathtt{Bg\#}(V_i) + \sum_{i=1}^{r} \mathtt{Bg\#}(F_i) + \sum_{i=1}^{m} \mathtt{Bg\#}(C_i)$ background trapezoids. We then select all the $4r+2n$ wrappers (note that a wrapper is background trapezoid covering a connector, and there are $4r+2n$ connectors altogether, including $r+2n$ inner connectors of variable structures, r inner connectors of shifters, r connectors connecting variable structures and shifters, and r connectors connecting shifters and checkers), and all solid points from connectors can be covered by these wrappers. Hence, B can be covered by $K = (\sum_{i=1}^{n} \mathtt{Bg\#}(V_i) + r + 2n) + \sum_{i=1}^{m} \mathtt{Bg\#}(C_i) + (\sum_{i=1}^{r} \mathtt{Bg\#}(F_i) + r) + (4r+2n)$ trapezoids in the way discussed above.

To see that a covering instance of J implies a satisfying truth assignment of I, we obtain the truth assignment from the covering instance in the following way: For each variable v_i, it is assigned a true value if none of the negative output connecting trapezoids of the variable structure V_i is selected, otherwise it is assigned false. We now show that this assignment is satisfying.

Consider the covering instance. First of all, at least $4r+2n$ background trapezoids are necessary to cover all solid points of all connectors (note that these *background* trapezoids will not cover any solid point of non-connector components). We then partition all the background and connecting trapezoids that cover solid point(s) in non-connector components as follows. For each variable structure V_i, all its background and (inner or output) connecting trapezoids are put into one unique group $\mathcal{G}_i^V$. From Lemma 3(3), at least $\mathtt{Bg\#}(V_i) + r_i^+ + r_i^- + 2$ trapezoids need to be selected from $\mathcal{G}_i^V$. For each shifter F_i, all its background trapezoids, as well as its inner and its output connecting trapezoids (excluding its input connecting trapezoids, which are also output connecting trapezoids of variable structures) are put into a unique group $\mathcal{G}_i^F$. From Lemma 5(2), at least $\mathtt{Bg\#}(F_i) + 1$ trapezoids need to be selected from $\mathcal{G}_i^F$. For each clause checker C_i, all its background trapezoids are put into a unique group $\mathcal{G}_i^C$ (its connecting trapezoids are also output connecting trapezoids of shifters and not counted here). From Lemma 4(2), $\mathtt{Bg\#}(C_i)$ trapezoids need to be selected from $\mathcal{G}_i^C$. Since $(4r + 2n) + \sum_{i=1}^{n} \left(\mathtt{Bg\#}(V_i) + r_i^+ + r_i^- + 2\right) + \sum_{i=1}^{r} (\mathtt{Bg\#}(F_i) + 1) + \sum_{i=1}^{m} (\mathtt{Bg\#}(C_i)) = K$, the above numbers are tight for all groups: exactly $\mathtt{Bg\#}(V_i) + r_i^+ + r_i^- + 2$ trapezoids are selected from $\mathcal{G}_i^V$, $\mathtt{Bg\#}(F_i) + 1$ trapezoids are selected from $\mathcal{G}_i^F$, and $\mathtt{Bg\#}(C_i)$ trapezoids are selected from $\mathcal{G}_i^C$.

For each clause c_i, since the corresponding C_i is covered, from Lemma 4(3), at least one input connecting trapezoid of C_i, say t_1, is selected. Since t_1 is the output connecting trapezoid of some shifter, from Lemma 5(3), an input connecting trapezoid, denoted t_2, of the same shifter must be selected. Again, t_2 is the output connecting trapezoid of some variable structure V_j.

If t_2 is a positive output connecting trapezoid, from Lemma 3(4), no negative output connecting trapezoids of V_j can be selected. Thus, variable v_j is assigned true. On the other hand, according to the construction of shifters, there is a positive occurrence of v_j in clause c_i. Hence, c_i is satisfied.

If t_2 is a negative output connecting trapezoid, from Lemma 3(4), no positive output connecting trapezoid of V_j can be selected. Thus, variable v_j is assigned false. On the other hand, according to the construction of shifters, there is a negative occurrence of v_j in clause c_i. Hence, c_i is satisfied.

3 Conclusion

To help queries on all sub-series in time series databases, the following optimization problem was studied: find the minimum number of trapezoids to cover a given set of points on the sub-series plan. We formalized this optimization problem into the decision problem, namely BCT (Bitmap Cover with special Trapezoids), and proved that this decision problem is NP-hard.

We used the basic component design approach in the proof, and a critical step was to identify a set of pairwise invisible points to ensure that a minimum number of trapezoids are required. This technique can be used to prove similar results. For example, we conjecture that it is also NP-hard to cover a bitmap with a set of arbitrary trapezoids (i.e., without requiring that each trapezoid have the special shape as in this paper). To prove this proposition, we may also construct components with special pairwise invisible points as in this paper.

References

[CR94] Joseph C. Culberson and Robert A. Reckhow. Covering polygons is hard. *Journal of Algorithms*, 17, 1994.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and intractability: a guide to the theory of NP-completeness.* W. H. Freeman and Company, 1979.

[QWW98] Yunyao Qu, Changzhou Wang, and Xiaoyang Sean Wang. Supporting fast search in time series for movement patterns in multiple scales. In *International Conference on Information and Knowledge Management*, 1998.