Point Pattern Matching in One Dimension: Applications to Music Information Retrieval

Godfried Toussaint

School of Computer Science

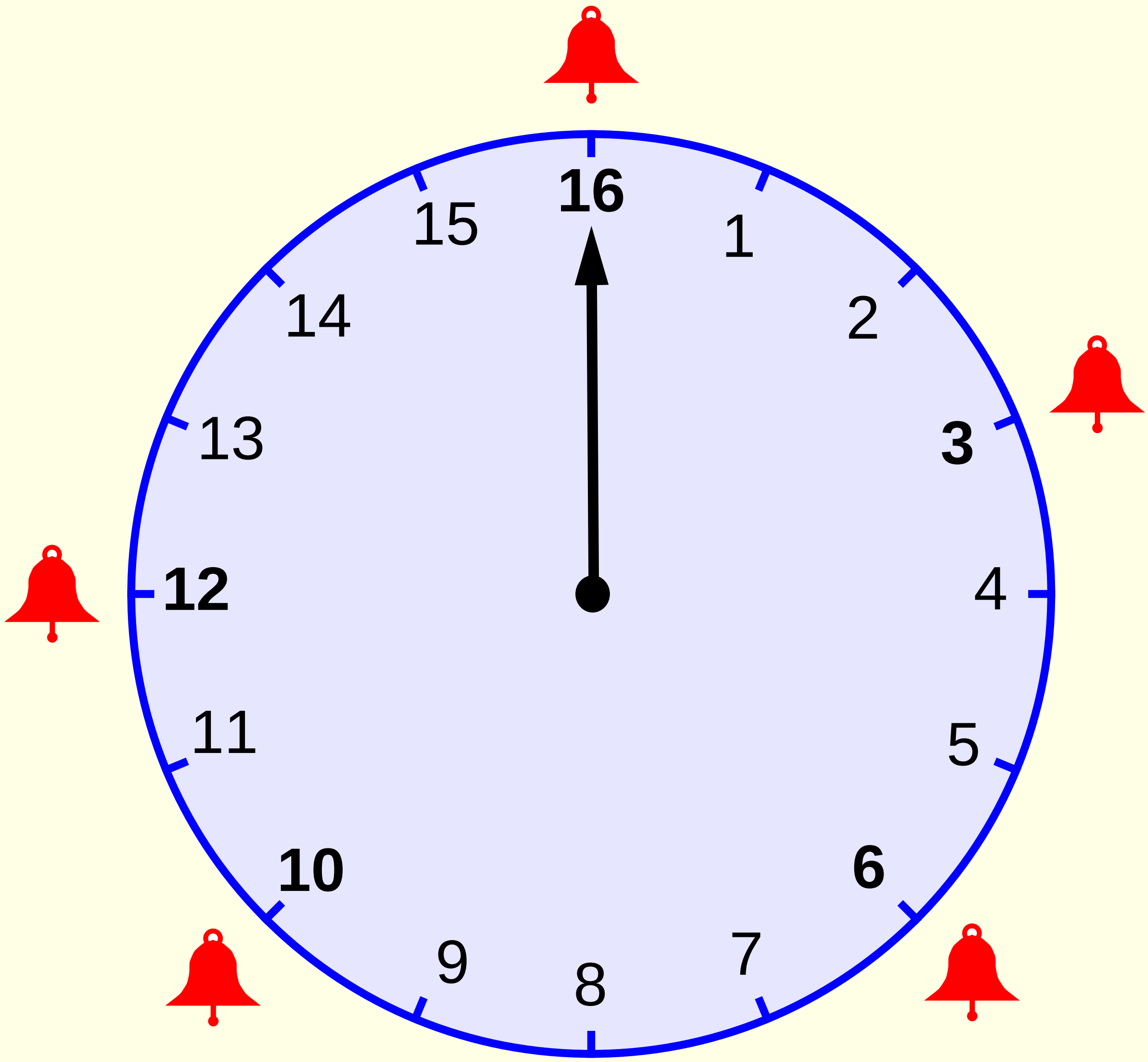
and

Center for Interdisciplinary Research in Music Media and Technology

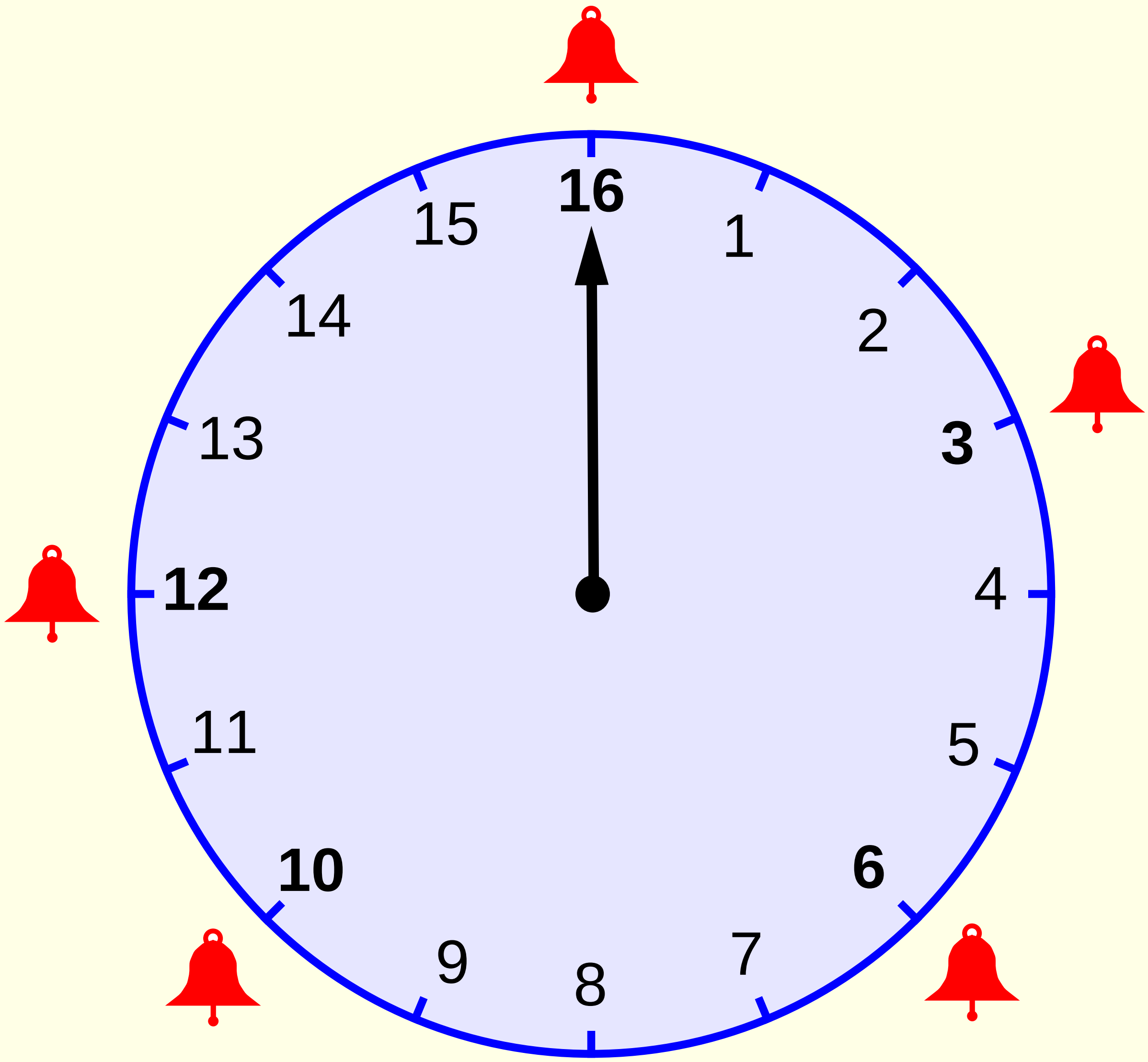
The Schulich School of Music

McGill University

The 16-hour Clock that Strikes a Bell at the Hours of 16, 3, 6, 10, and 12.



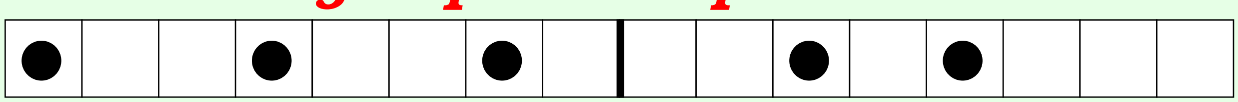
The 16-hour Clock that Strikes a Bell at the Hours of 16, 3, 6, 10, and 12.



The *clave* Son rhythm of Cuba.

Some Ways of Representing the *Clave* Son.

binary sequence representation

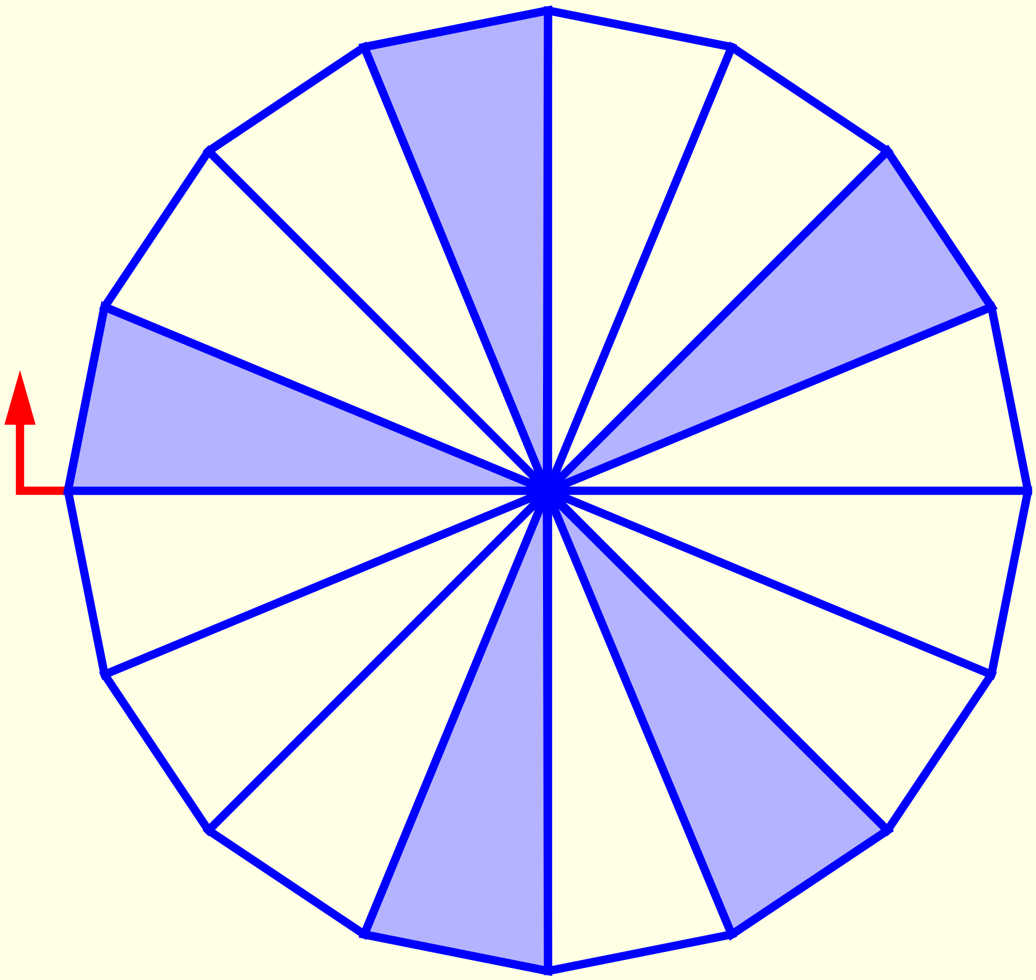


[3 3 4 2 4]

The 5*th* is the *Box Notation Method* developed by Philip Harland at UCLA in 1962 also known as **TUBS** (**T**ime **U**nit **B**ox **S**ystem).

The “*Clave* Son” in Ancient *Persian* Notation.

Safi-al-Din, “Al-sharafiyyeh,” 1252.
The *Al-saghil-al-avval* rhythm.



Measuring the Similarity of Rhythms.

The Hamming distance between two rhythms represented as binary sequences is the sum of the number of places in the sequence where the symbols in both rhythms differ.

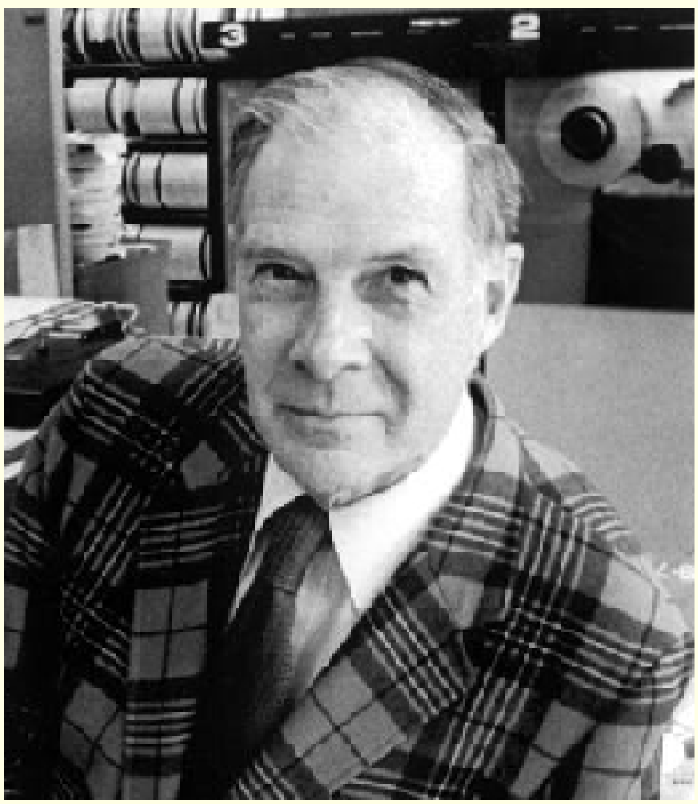
Put another way: it is the minimum number of substitutions required to change one sequence to the

Example:

Rumba	●			●				●			●		●			
Gahu	●			●			●				●				●	
							✓	✓					✓		✓	

distance = 4

Introduced by Richard Hamming in 1950.



Another view of the Hamming distance.

Such Binary sequences may also be viewed as vectors in a 16-dimensional space:

$$X = x_1, x_2, \ldots, x_{16}$$

Example:

Rumba	1	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0
Gahu	1	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0

Then the Hamming distance between X and Y is:

$$d_H(X,Y) = \sum_{i=1}^{16} \left| x_i - y_i \right|$$

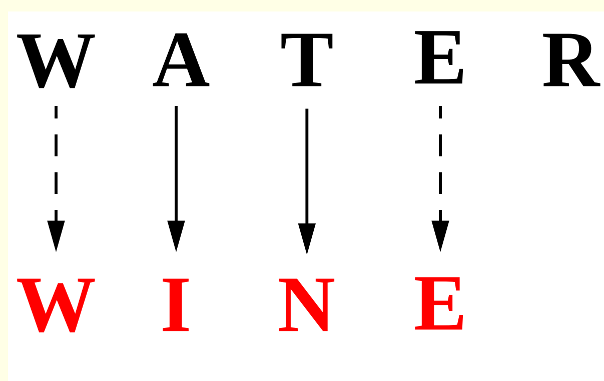
Sequence Comparison: Levenshtein Distance

Vladimir I. Levenshtein, “Binary codes capable of correcting deletions, insertions, and reversals,” *Cybernetics and Control Theory*, 1966.

Popularly known as the **“edit” distance**.

Given two sequences (strings) of symbols:
A= $a_1, a_2, \ldots, a_n$ and ***B***=$b_1, b_2, \ldots, b_m$, the Levenshtein distance between ***A*** and ***B*** is the smallest number of *insertions*, *deletions*, and *substitutions* (replacements) required to change ***A*** into ***B***.

example:



The father of Russian *information theory.*

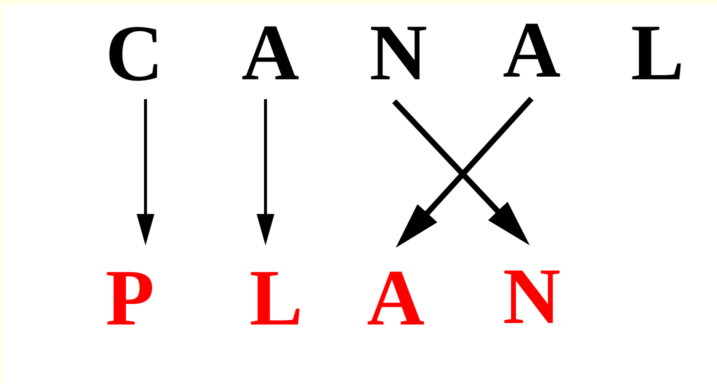
Sequence Comparison: **Edits + Swaps Distance**

R. A. Wagner and M. J. Fisher, "The *string-to-string* correction problem" *J. of the ACM*, 1974.

Give an O(*mn*) dynamic programming algorithm for computing the **edit distance**.

R. Lowrance and R. A. Wagner, "An extension of the string-to-string correction problem" *J. of the ACM*, 1975.

They added the **swap** operation to the **edit distance**.
A **swap** interchanges two adjacent characters.
They also give an O(*mn*) dynamic programming algorithm for computing it.



The Longest Common Subsequence Problem and the Edit (*Levenshtein*) Distance

Given two sequences $\boldsymbol{A} = a_1, a_2, \ldots, a_n$ and $\boldsymbol{S} = s_1, s_2, \ldots, s_m$, with $m \leq n$.
$\boldsymbol{S}$ is a subsequence of $\boldsymbol{A}$ if for some $1 \leq i_1 < i_2 < \ldots < i_m \leq n$ we have $a_{i_h} = s_h$ for all $1 \leq h \leq m$.

Given two sequences $\boldsymbol{A} = a_1, a_2, \ldots, a_n$ and $\boldsymbol{B} = b_1, b_2, \ldots, b_n$, let $\boldsymbol{v}(\boldsymbol{A},\boldsymbol{B})$ denote the maximum length of any subsequence common to $\boldsymbol{A}$ and $\boldsymbol{B}$.

Let $\boldsymbol{d}(\boldsymbol{A},\boldsymbol{B})$ denote the edit (*Levenshtein*) distance (total cost) between $\boldsymbol{A}$ and $\boldsymbol{B}$ with the following operation costs: *insertion* = 1, *deletion* = 1, and *substitution* = *deletion* + *insertion* = 2.

Then:

$$\boldsymbol{d}(\boldsymbol{A},\boldsymbol{B}) = 2\left(n - \boldsymbol{v}(\boldsymbol{A},\boldsymbol{B})\right)$$

How Similar are two Random Sequences?

V. Chvatal and D. Sankoff, "Longest common subsequences of two random sequences," *J. Applied Probability*, 1975.

Suppose $\boldsymbol{A} = a_1, a_2, \ldots, a_n$ and $\boldsymbol{B} = b_1, b_2, \ldots, b_n$, are sequences created by random draws from an alphabet of size k,

(all letters occur with equal probability, and successive draws are independent).

They proved that the expected value of $\boldsymbol{v}(\boldsymbol{A},\boldsymbol{B})$ is is asymptotically proportional to n:

$$\lim_{n \to \infty} \frac{E\{\boldsymbol{v}(\boldsymbol{A},\boldsymbol{B})\}}{n} = c_k$$

Let

$$d_k = \lim_{n \to \infty} \frac{E\{\boldsymbol{d}(\boldsymbol{A},\boldsymbol{B})\}}{n}$$

Since

$$\boldsymbol{d}(\boldsymbol{A},\boldsymbol{B}) = 2\left(n - \boldsymbol{v}(\boldsymbol{A},\boldsymbol{B})\right)$$

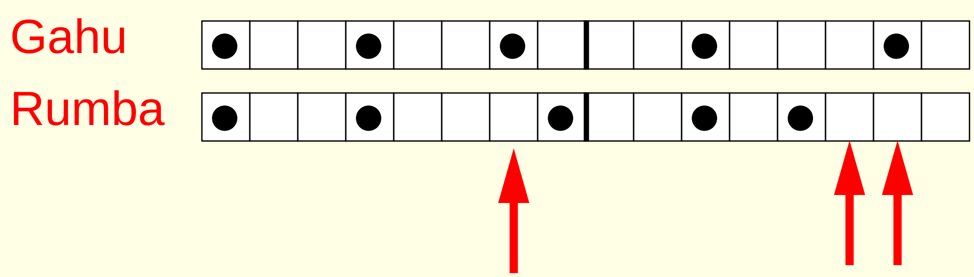
We have

$$\boldsymbol{d}_k = 2\left(n - \boldsymbol{c}_k\right)$$

The **Swap-distance** Between **Rumba** and **Gahu**.

The ***swap-distance*** *between two rhythms represented as binary sequences of symbols is defined as the minimum number of position interchanges between adjacent one's and zero's required to transform one rhythm into the other.*

Example:



$d(R,G) = 3$

Computing the Swap-distance in Linear Time.

*Executing and counting the **swaps** is bad.*

***A* = x x x x x x x x _ _ _ _ _ _ _ _**
***B* = _ _ _ _ _ _ _ _ x x x x x x x x**

$T(n) = O(n^2)$

O(n) Algorithm

Convert the sequence of k onsets to a k-dimensional vector of x-coordinates.

X = 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
***A* = x _ _ x _ _ x _ _ _ x _ x _ _ _**
***B* = _ _ _ x _ _ _ x _ _ x x _ _ x _**

$X_A = (0, 3, 6, 10, 12)$
$X_B = (3, 7, 10, 11, 14)$

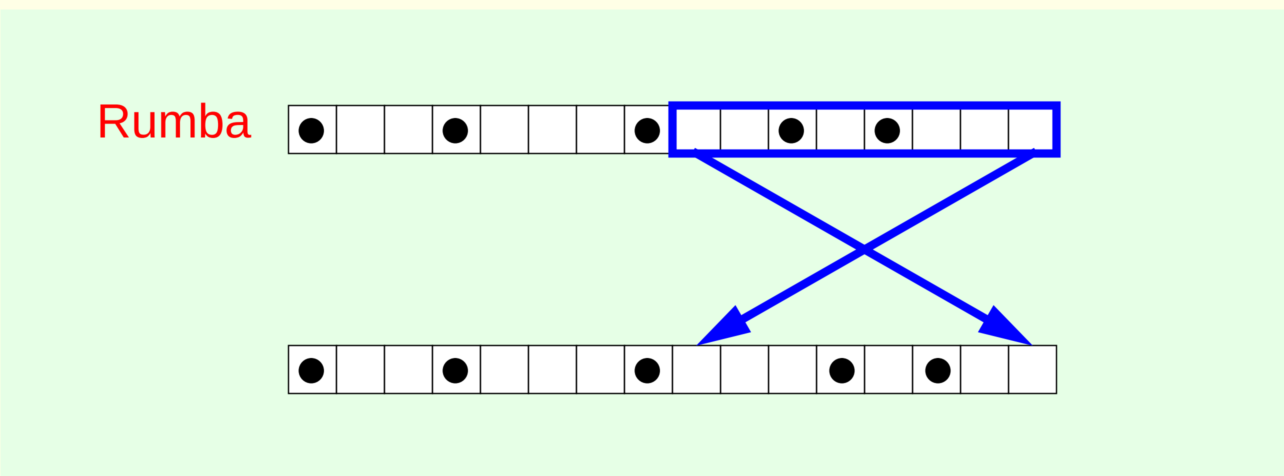
$d_{swap}(A, B) = 3 + 4 + 4 + 1 + 2 = 14$

Sequence Comparison with Block Reversals

Let A and B be two sequences of n elements each, over some alphabet set.

Define d(A,B) as the minimum number of element substitutions and block (subsequence) reversals needed to transform A to B, such that no element is involved in more than one operation.

(a swap is the smallest possible block reversal)



Theorem: *S. Muthukrishnan and S. Cenk Sahinalp* (2004)

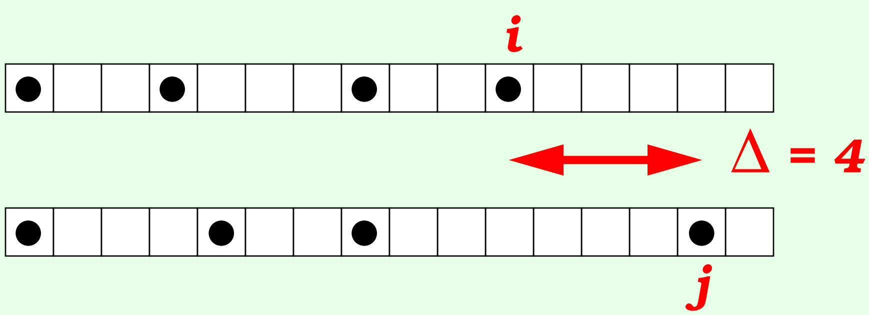
d(A,B) may be computed in time $O(n \log^2 n)$

Sequence Comparison with *Fuzzy* Hamming Distance

A. Bookstein, S. T. Klein and T. Raita, (2001)

The fuzzy (extended, generalized) Hamming distance is an edit distance with three operations:

(1) *insertion*
(2) *deletion*
(3) *shift* (cost as a function of Δ)



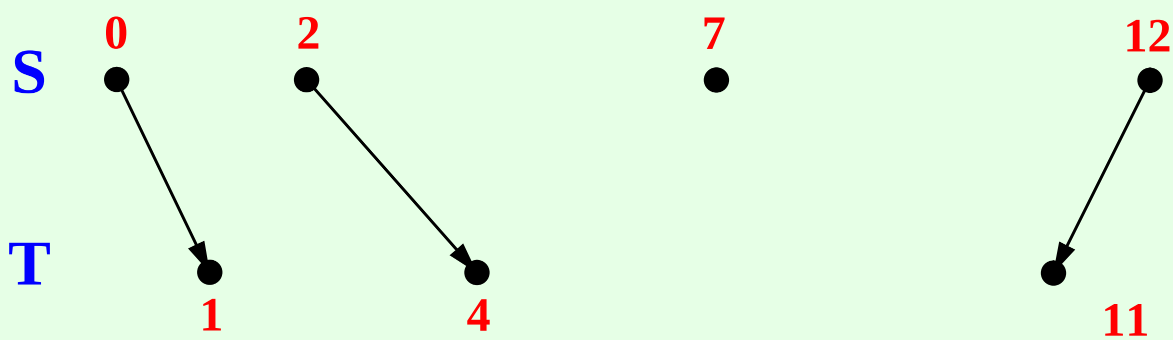
For two binary sequences A and B of n elements each, d(A,B) is computed in time $O(n^2)$ time with dynamic programming.

If only the shift is used with cost of shift = Δ, *then this distance becomes the swap distance.*

The One-to-One Assignment Problem

Richard Karp and Shuo-Yen Li (1975) Discrete Mathematics.

Let S and T with with $T \leq S$ *be two sets of points on the line. They give an algorithm that computes a minimum-weight one-to-one assignment from S to T that runs in $O(n \log n)$ time* (and $O(n)$ *after sorting*) *where n is the sum of the cardinalities of the sets.*



A minimum one-to-one assignment from S to T

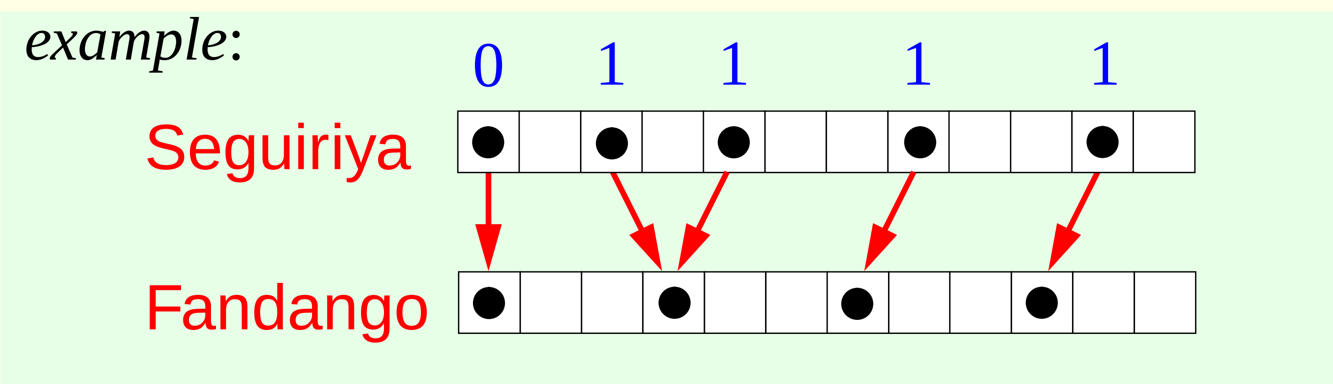
The Directed Swap-Distance Between Two Rhythms of Different Density

Miguel Diaz-Bañez, Giovanna Farigu, Francisco Gomez, David Rappaport, and Godfried T. Toussaint, "*El compas flamenco: A phylogenetic analysis*," 2004.

The directed swap-distance between two rhythms of different density is the minimum number of position interchanges between adjacent elements, required to transform the "larger" rhythm into the "smaller" rhythm under two constraints:

(1) *every onset of the "larger" rhythm must travel to some onset of the "smaller" rhythm,*

(2) *every onset of the "smaller" rhythm must receive at least one onset from the "larger" rhythm.*

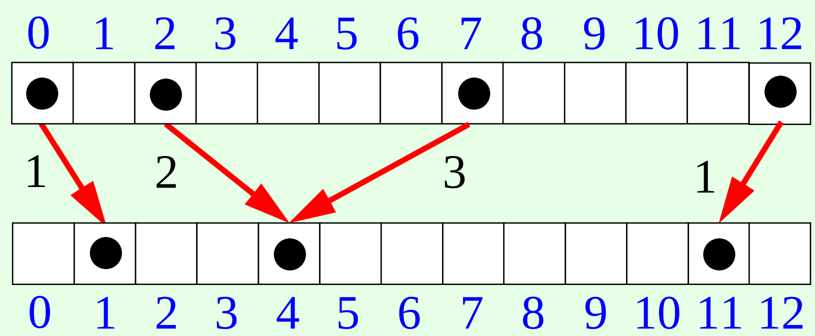


swap-distance (*S*,*F*) = 4

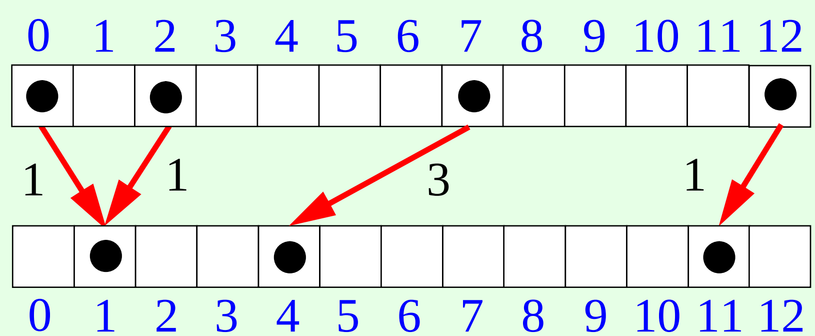
The Directed Swap-distance = the Restriction Scaffold Assignment Problem in Computational Biology

A. Ben-Dor, R. Karp, B. Schwikowski, and R. Shamir
(2003) Journal of Computational Biology.
Algorithm that runs in O(n) time for sorted points.

incorrect
7



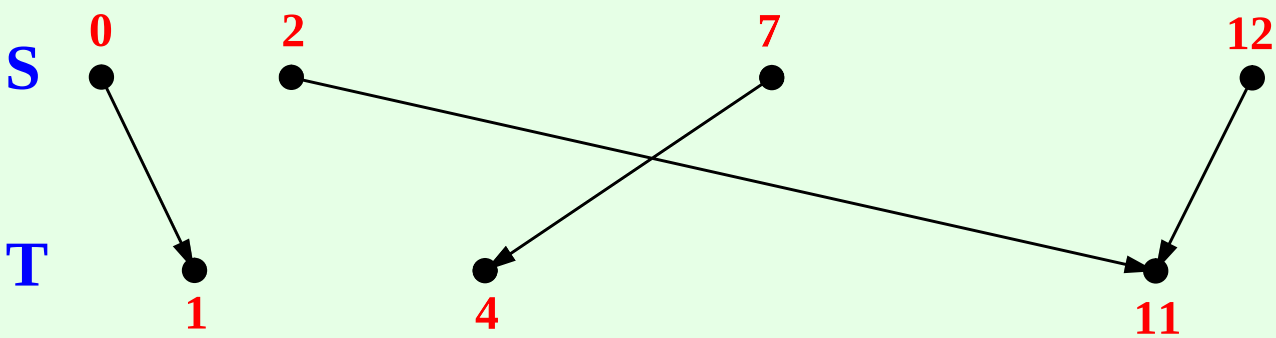
correct
6



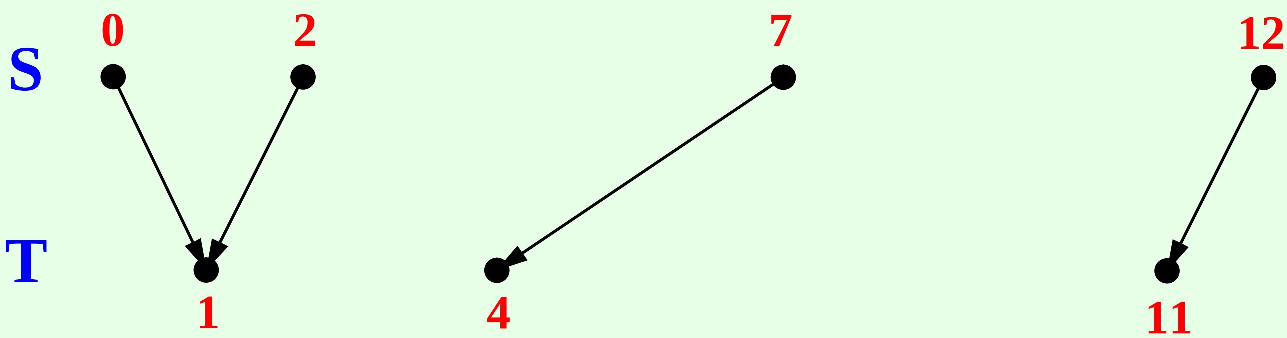
The Surjection Distance Between two Point Sets

Thomas Eiter and Heikki Mannila (1997) Acta Informatica.

Give an algorithm that computes a minimal surjection from S to T that runs in $O(n^3)$ time, where n is the sum of the cardinalities of the sets.



A surjection between *S* and *T*

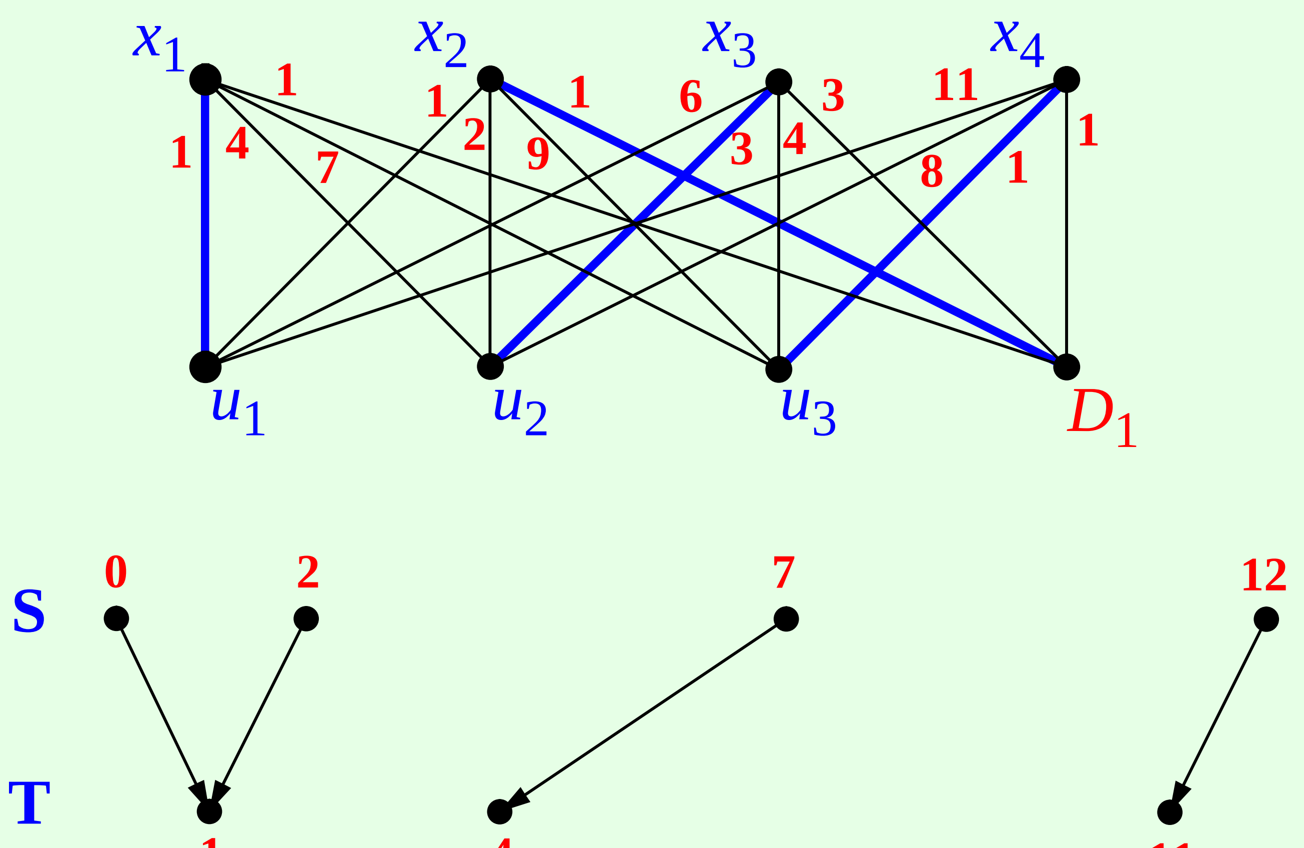


A minimal surjection between *S* and *T*

The Algorithm of Eiter & Mannila

Lemma:
Let F be a minimal surjection from S to T. Then for any s_i not equal to s_j if $F(s_i) = F(s_j)$, then the distance from s_i to $F(s_i)$ is not more than the distance from s_i to any oyther element of T.

Their algorithm uses a reduction to a minimum-weight perfect matching in a suitable bipartite graph.



A minimal surjection between *S* and *T*

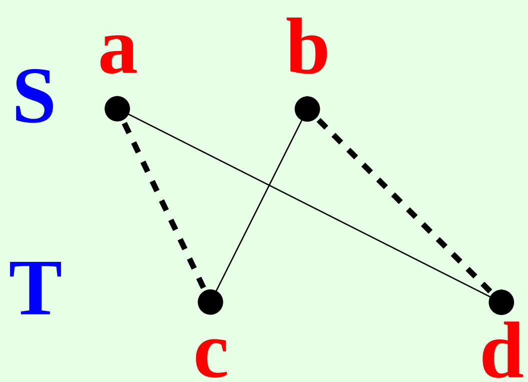
The O(n^2) Algorithm

J. Colannino and G. T. Toussaint (2005)

Quadrangle inequality lemma:

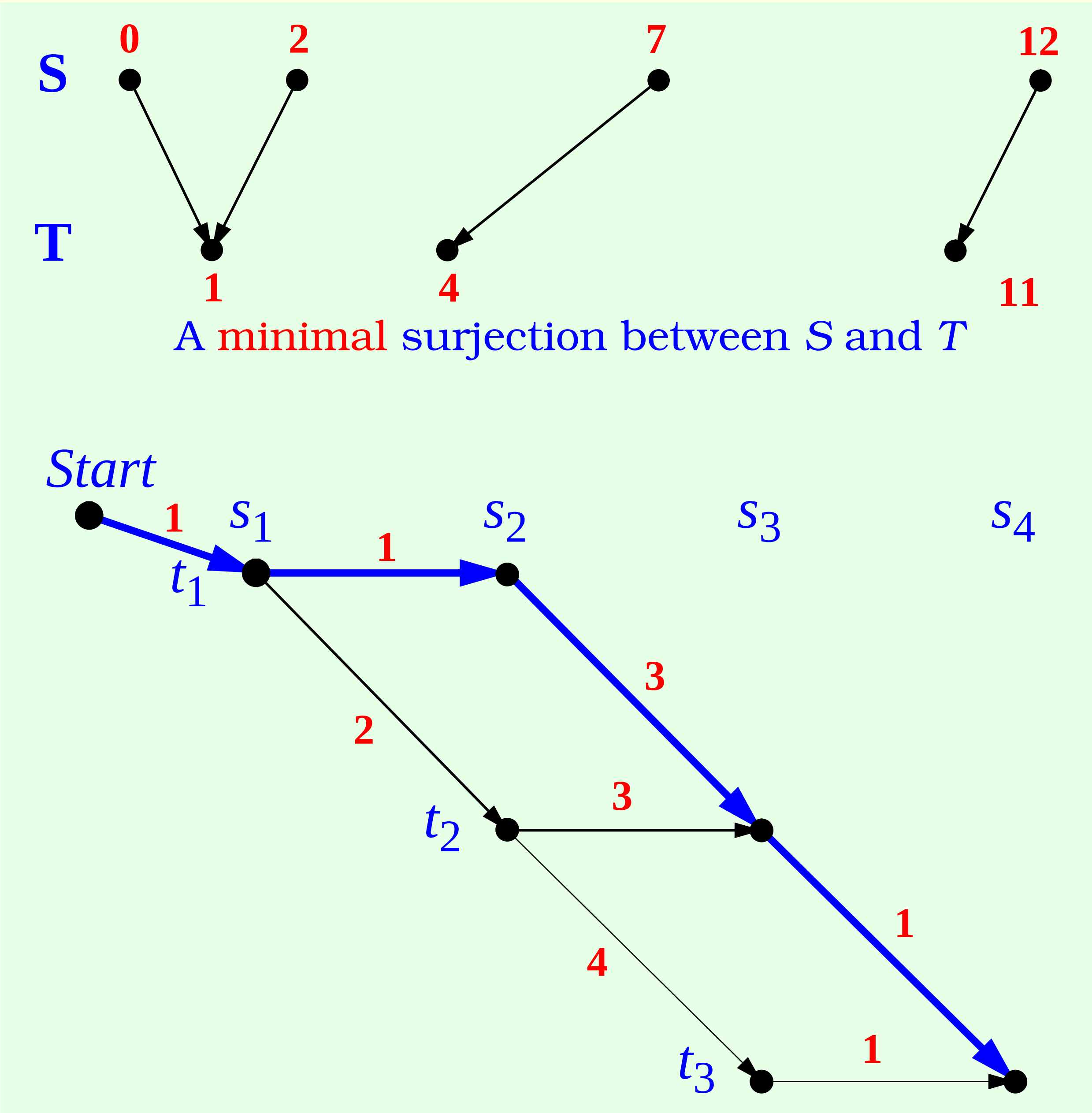
Let ***S*** *and* ***T*** *be two sets of points on the line.*
Let $\delta(s,t) = |s-t|$.
Then for $a < b$ *in* ***S*** *and* $c < d$ *in* ***T***

$$\delta(a,c) + \delta(b,d) \leq \delta(a,d) + \delta(b,c)$$



The O(n^2) Algorithm - *continued*

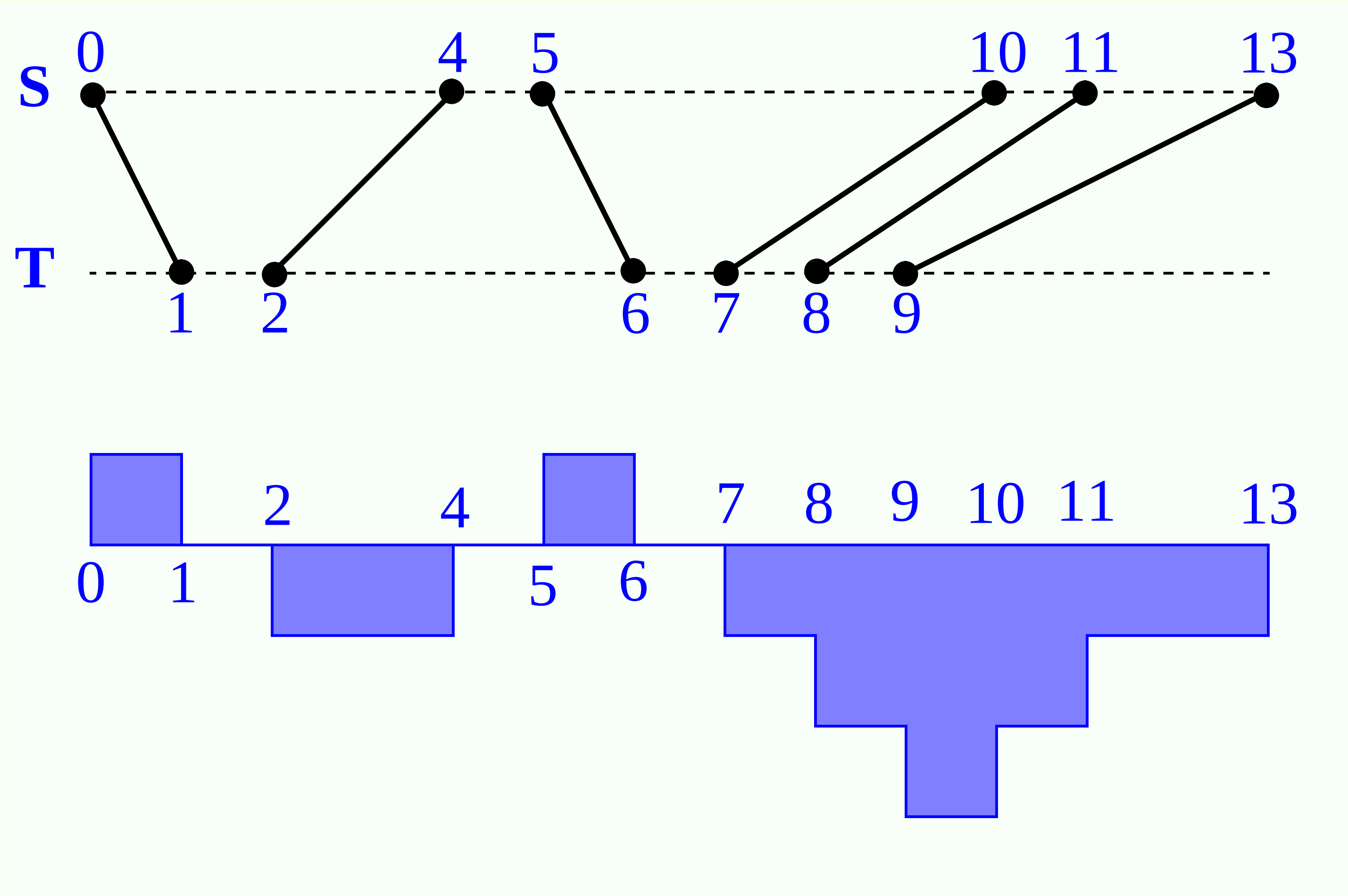
The algorithm uses a reduction of the directed swap-distance problem to a shortest path problem in a suitable directed acyclic graph.



An O(n) Algorithm for the Directed-Swap Distance for Sorted Points

J. Colannino, M. Damian, F. Hurtado, J. Iacono, H. Meijer, S. Ramaswami, and G. T. Toussaint (2006)

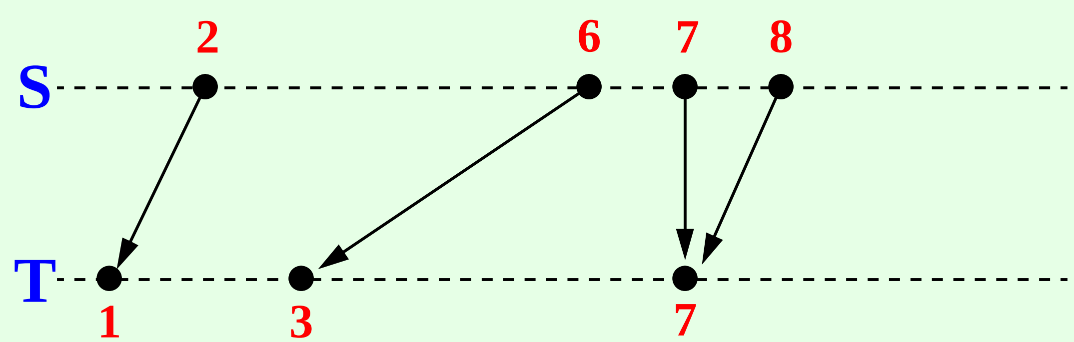
Via an extension of the one-to-one assignment algorithm of Karp and Li (1975).
Shaded area is the cost of the assignment.



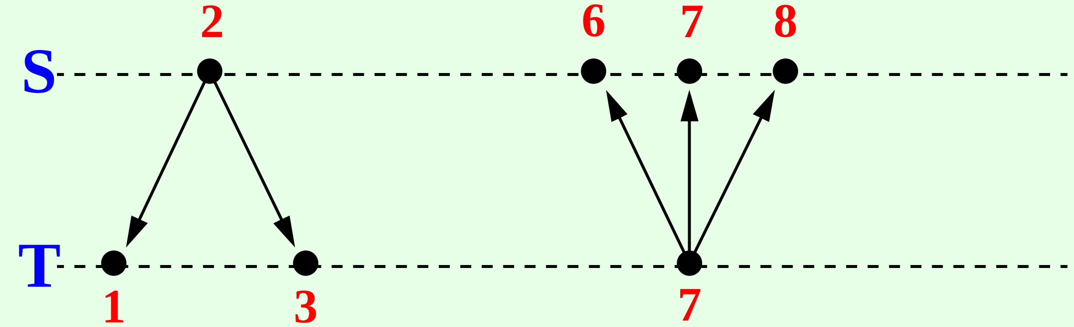
The Link Distance Between two Point Sets (many-to-many matching)

T. Eiter and H. Mannila (1997), Acta Informatica.

Present an algorithm that computes the link distance between S and T that runs in $O(n^3)$ time, where n is the sum of the cardinalities of the sets.



A minimal surjection between S and T

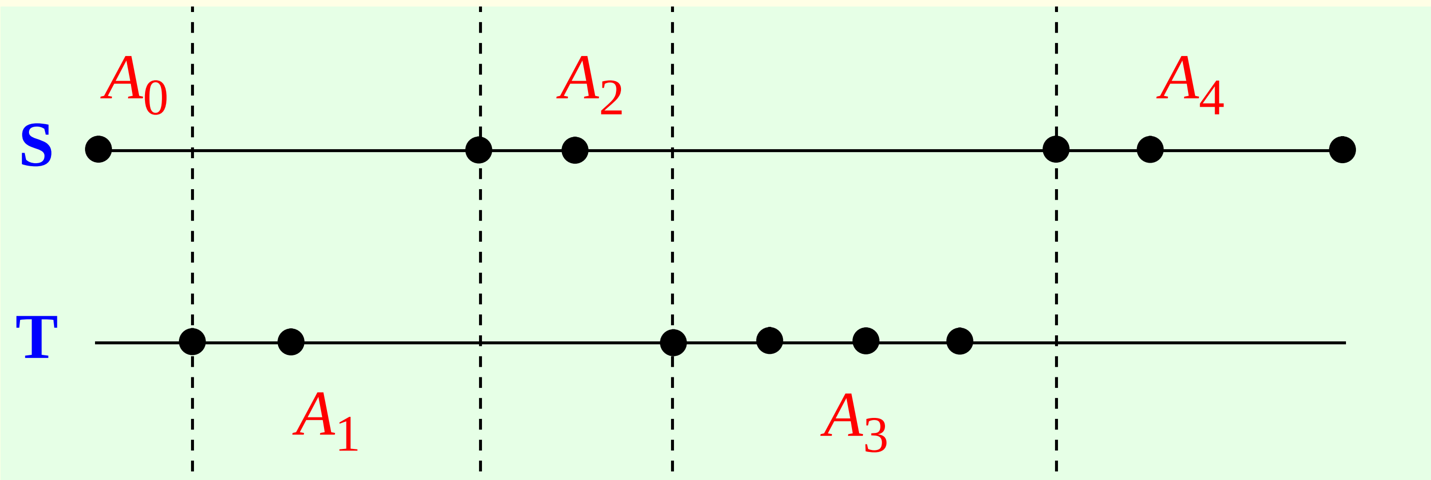


A minimal linking between S and T

An O(*n*) Algorithm for the Link Distance between two Sorted Point Sets

J. Colannino, M. Damian, F. Hurtado, S. Langerman, H. Meijer, S. Ramaswami, D. Souvaine, and G. T. Toussaint (2007)

*Present a dynamic programming algorithm that computes the link distance between **S** and **T** in O(n) time, where n is the sum of the cardinalities of two sorted sets.*

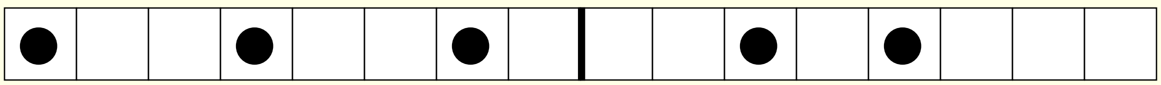


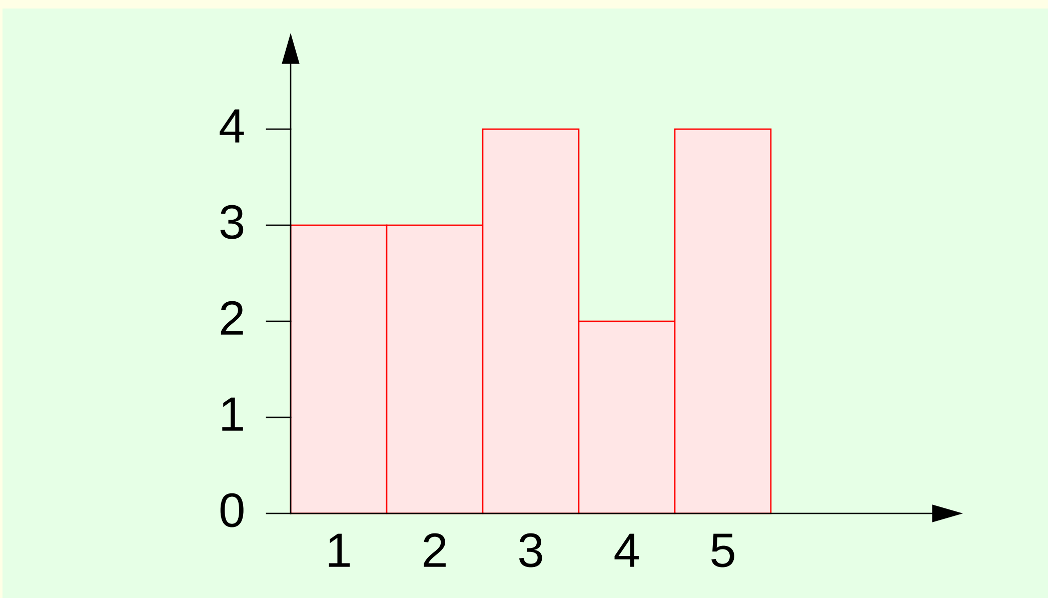
Lemma: *For each i > 0, A_i contains a point q_i such that, in a minimum-cost many-to-many matching, all points in A_i less than q_i are matched to points in A_{i-1}, and all points in A_i greater than q_i are matched to points in A_{i+1}.*

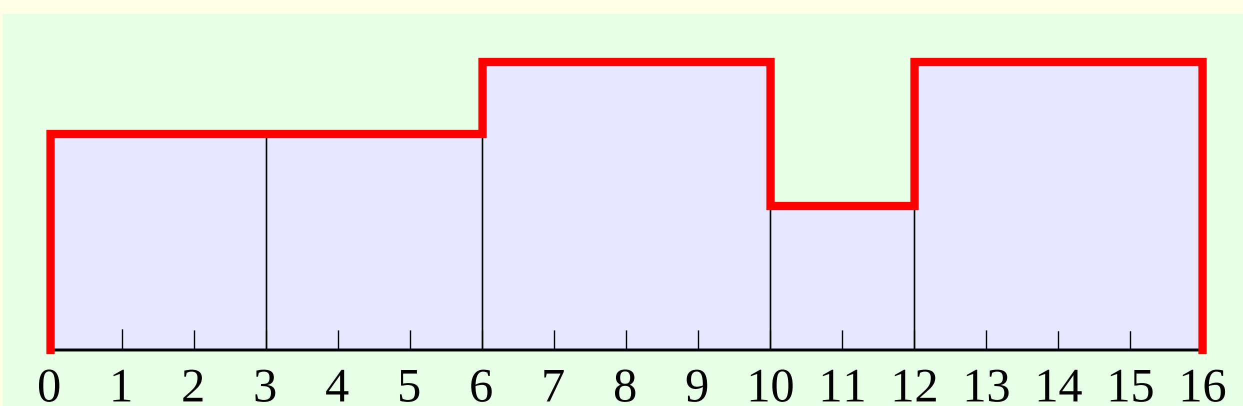
Chronotonic Representation of Rhythm - I

Kjell Gustafson, *"The graphical representation of rhythm,"* *Oxford University, 1988.*

3 views of the clave Son



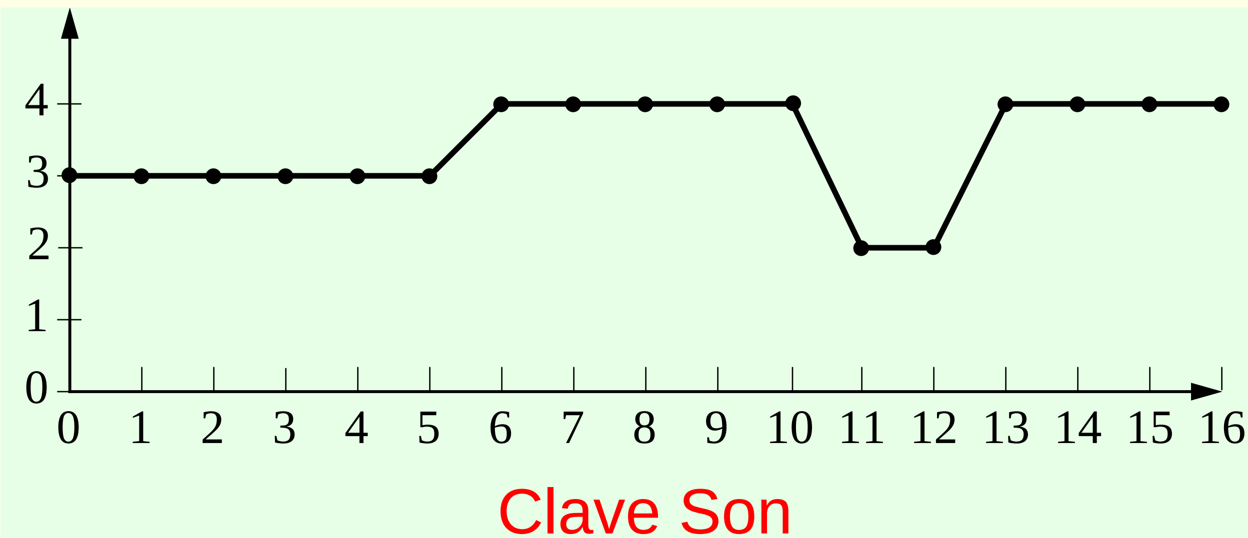


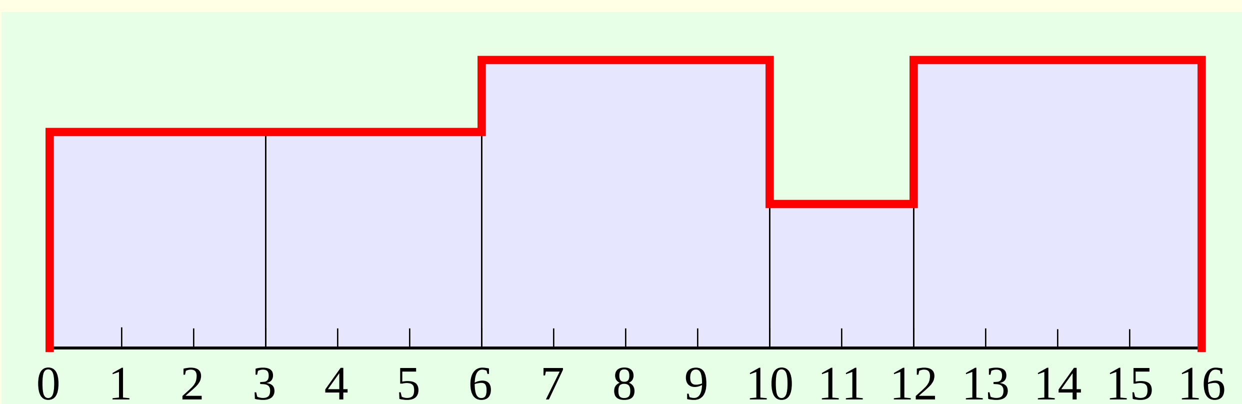


Chronotonic Representation of Rhythm - II

Ludger Hofmann-Engl, "Rhythmic similarity: A theoretical and empirical approach," Keele University, 2002.

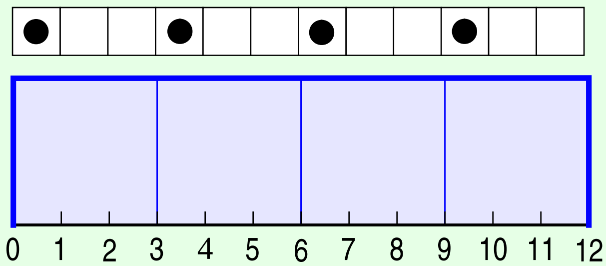
The clave Son as a chronotonic chain



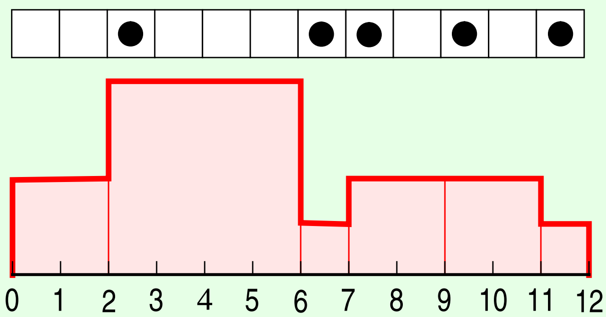


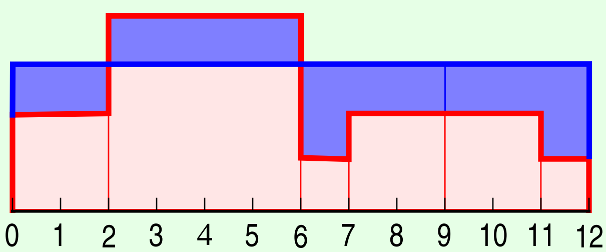
The *Chronotonic* Distance

Fandango



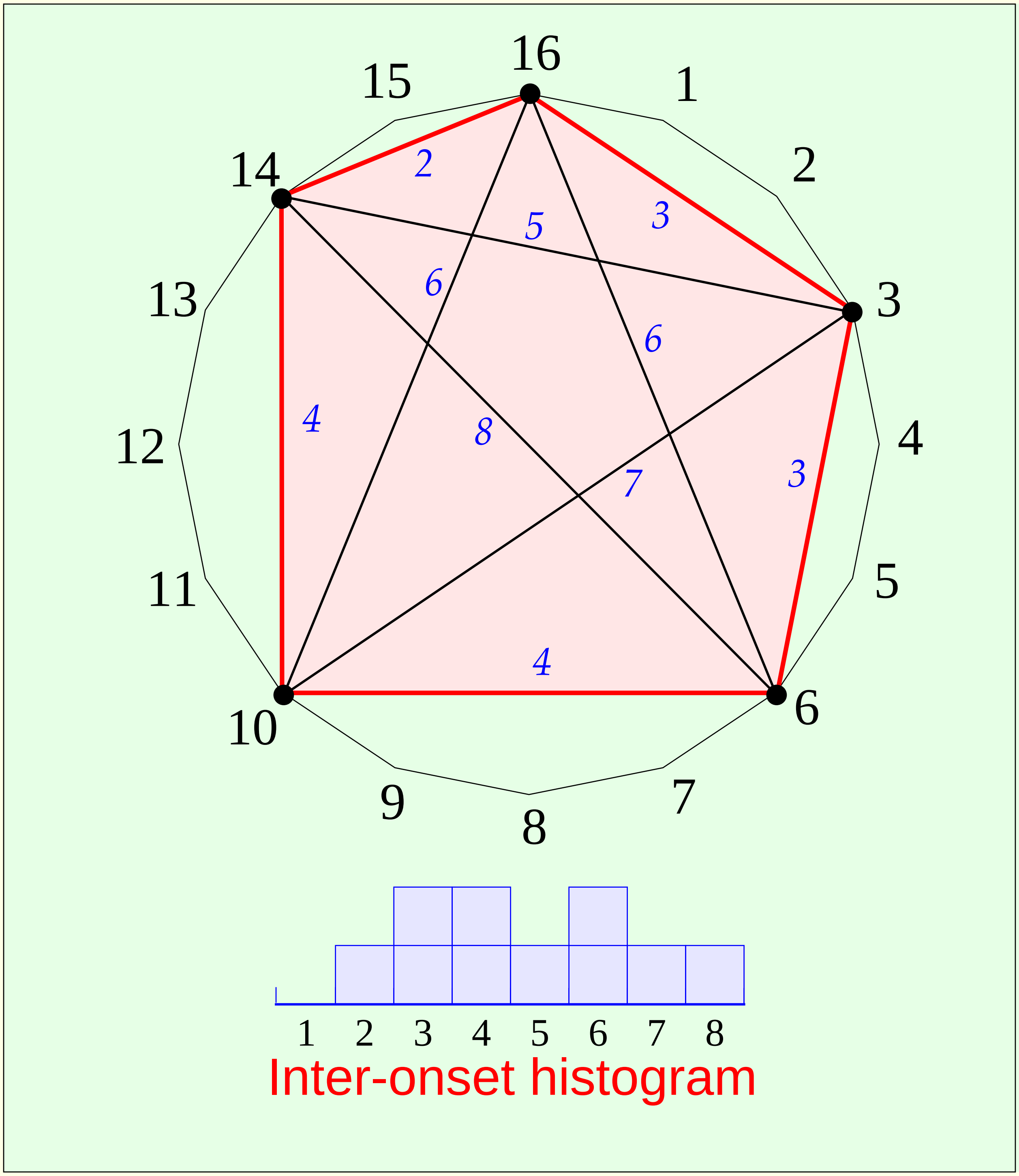
Bulería





The chronotonic distance between Fandango and Bulería is the area between the two curves shown shaded in ***dark blue***.

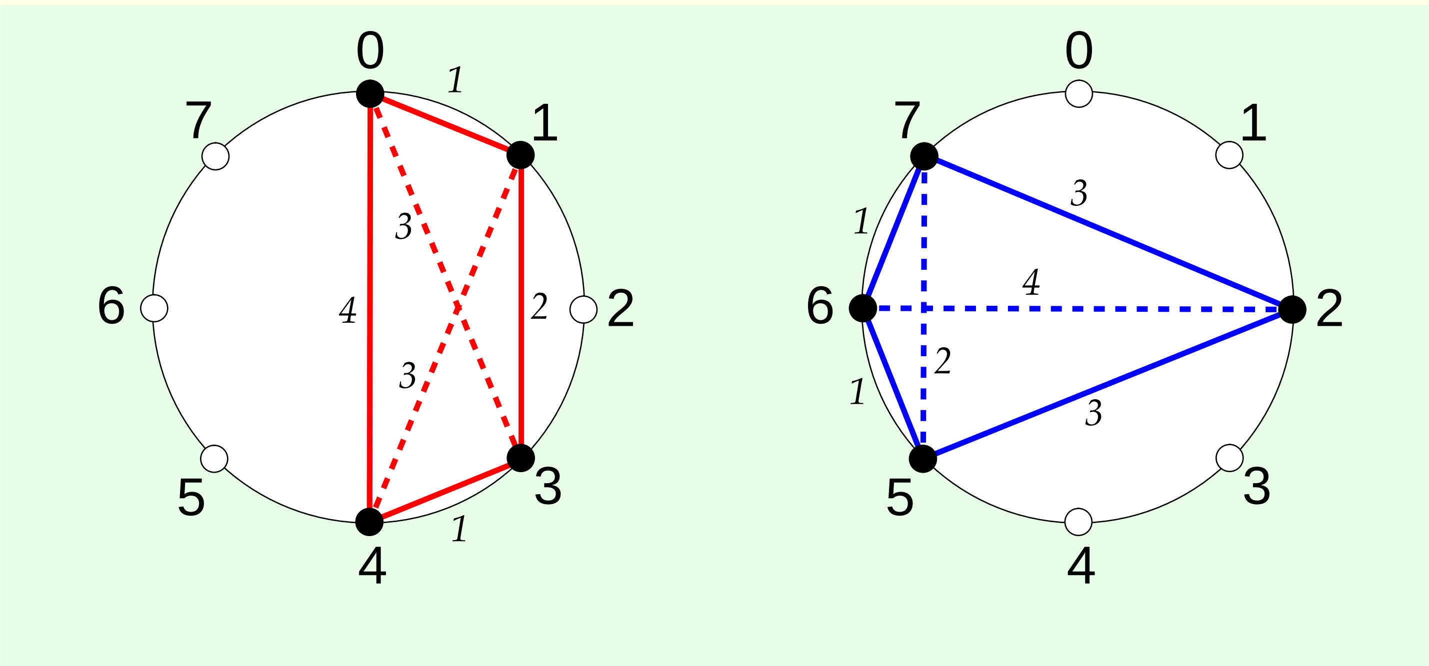
All Pairwise Interval Durations (geodesic distances) Contained in the Gahu *Clave*.

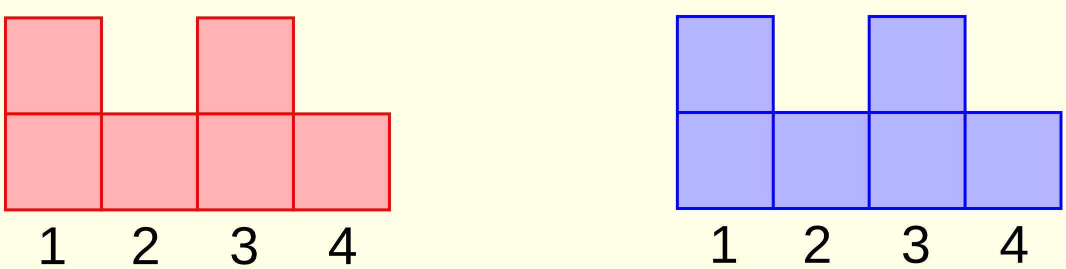


Homometric Rhythms

A. Lindo Patterson,
"Ambiguities in the X-ray analysis of crystal structures," *Physical Review*, March, 1944.

Every *n*-point subset of a regular polygon with 2*n* vertices is homometric to its complement.



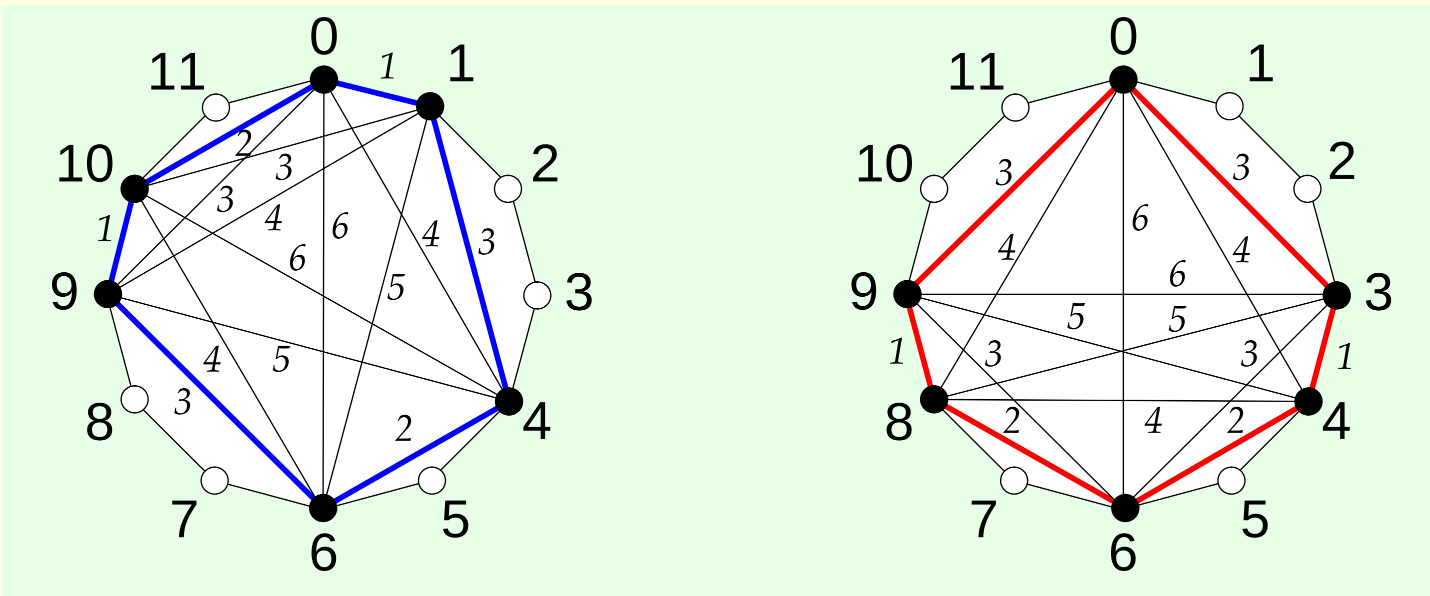


Interval histograms

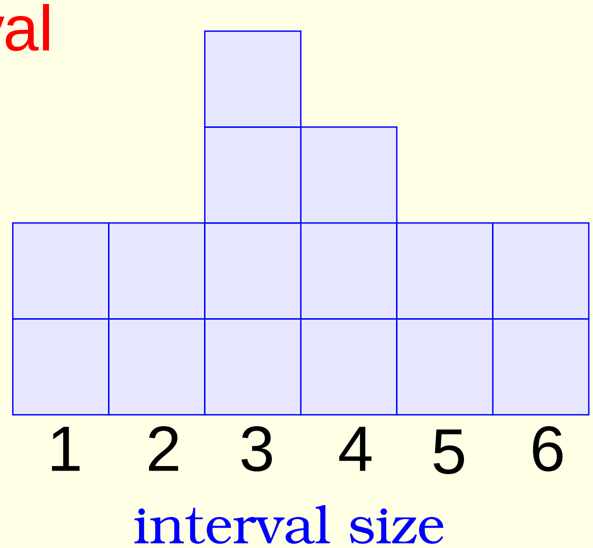
The Hexachordal Theorem

Theorem: Two *complementary* hexachords have the same *interval content*.

First observed empirically: Arnold Schoenberg, ~ 1908.



pitch interval histogram



The Hexachordal Theorem: Music-Theory Proofs

Theorem: Two *complementary* hexachords have the same *interval content*.

First observed empirically: Arnold Schoenberg, 1908.

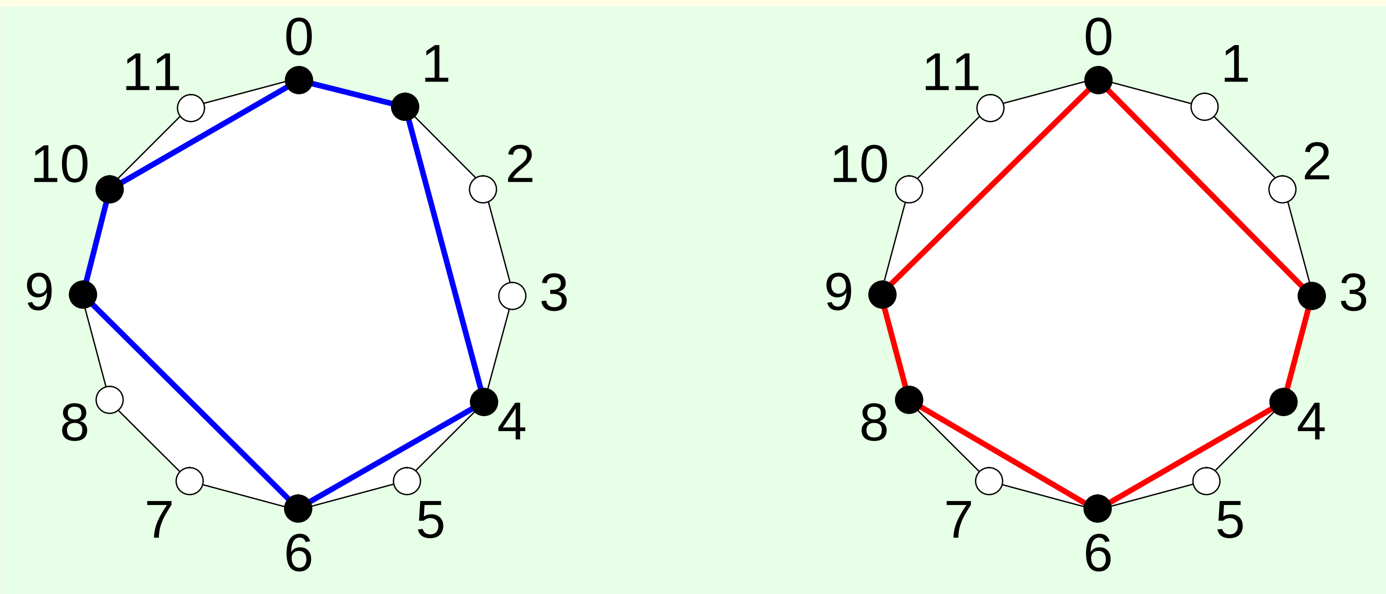
Proofs:

1. Milton Babbitt and David Lewin - 1959, *topology*

2. David Lewin - 1960, *group theory*

3. Eric Regener - 1974, *elementary algebra*

4. Emmanuel Amiot - 2006, *discrete fourier transform*



The Hexachordal Theorem: Crystallography Proofs

First observed experimentally: Linus Pauling and M. D. Shappell, 1930.

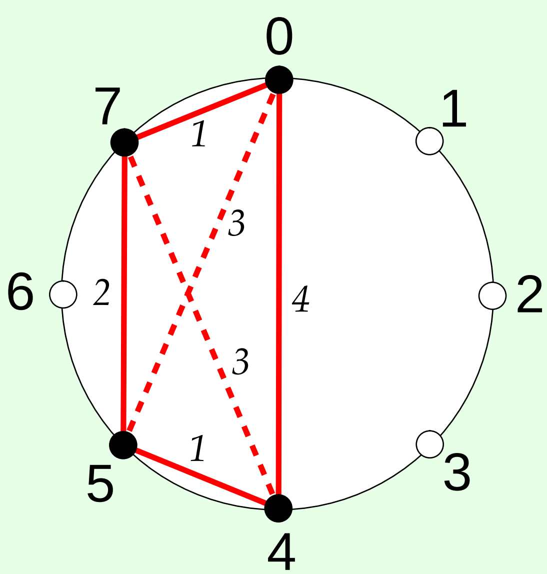
Proofs:

1. Lindo Patterson - 1944, *claimed proof not published*

2. Martin Buerger - 1976, *image algebra*

3. Juan Iglesias - 1981, *elementary induction*

4. Steven Blau - 1999, *elementary induction*



The Interval-content Theorem of Iglesias

Juan E. Iglesias,
"On Patterson's cyclotomic sets and how to count them," *Zeitschrift für Kristallographie*, 1981.

Theorem: Let p of the N vertices of a regular polygon inscribed on a circle be black dots, and the remaining $q = N - p$ vertices be white dots. Let n_{ww}, n_{bb}, and n_{bw} denote the multiplicity of the distances ***of a specified length*** between *white*-*white*, *black*-*black*, and *black*-*white*, vertices, respectively.

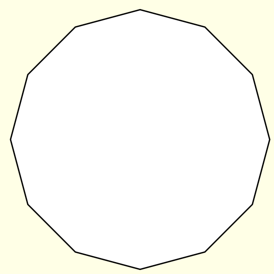
Then the following relations hold:

$$p = n_{bb} + (1/2)n_{bw}$$

$$q = n_{ww} + (1/2)n_{bw}$$

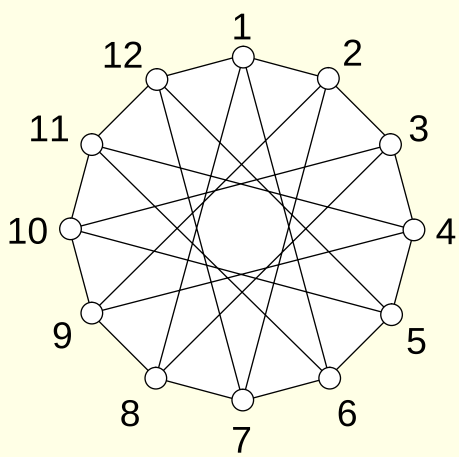
Lemma: Any given duration value *d* occurs with multiplicity *N*.

(1) If $d = 1$ or $d = N-1$ the multiplicity equals the number of sides of an *N-vertex regular* polygon.



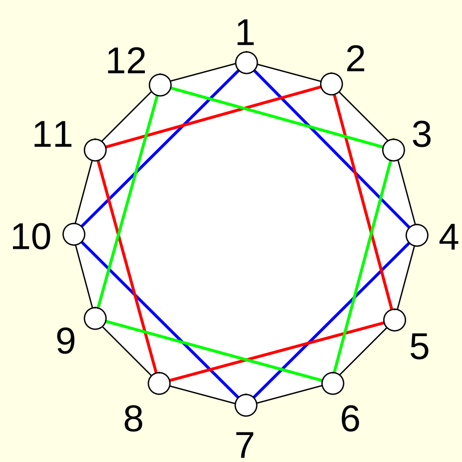
(2) If $1 < d < N-1$, and *d* and *N* are *relatively prime*, the multiplicity equals the number of sides of an *n-vertex regular star*-polygon.

$N = 12$
$d = 5$



(3) If *d* and *N* are *not relatively prime* then the multiplicity equals the total number of sides of a group of convex polygons. There are *g.c.d.*(*d*,*N*) polygons with *N/g.c.d*(*d*,*N*) sides each.

$N = 12$
$d = 3$

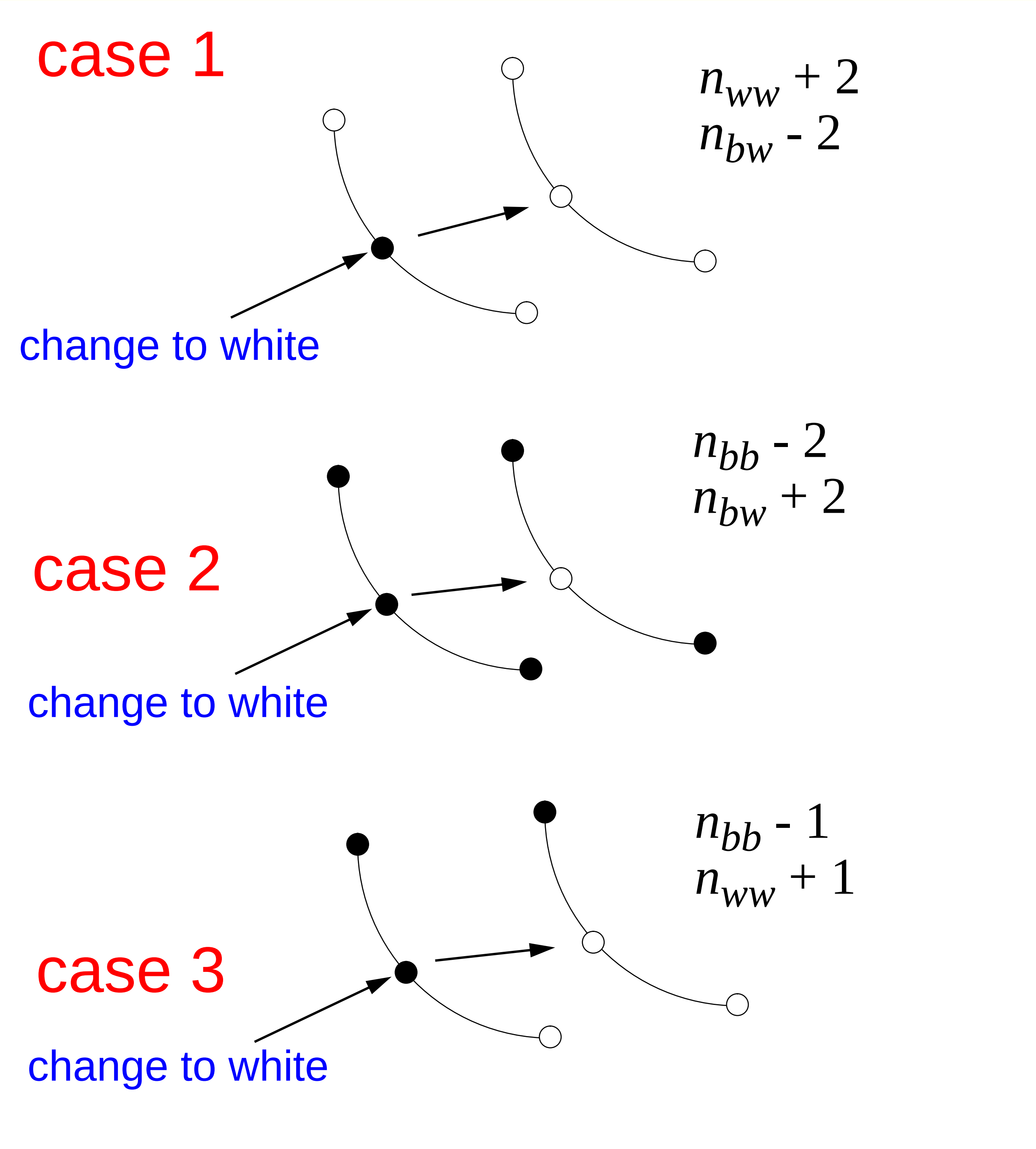


Proof of Iglesias' theorem:

For each duration value d

$$p = n_{bb} + (1/2)n_{bw}$$

$$q = n_{ww} + (1/2)n_{bw}$$



Iglesias' Proof of Patterson's Theorems

Theorem 1: If two different black sets form a homometric pair, then their corresponding complementary white sets also form a homometric pair.

Proof: If the black sets are homometric they must have the same number of points.
Then, for each duration value d

$$p = n_{bb} + (1/2)n_{bw} = n^*_{bb} + (1/2)n^*_{bw}$$

$$q = n_{ww} + (1/2)n_{bw} = n^*_{ww} + (1/2)n^*_{bw}$$

and thus

$$p - q = n_{bb} - n_{ww} = n^*_{bb} - n^*_{ww}$$

Since the black sets are homometric $n_{bb} = n^*_{bb}$
and thus $n_{ww} = n^*_{ww}$

Theorem 2: If $p = q$ the two sets are homometric.

Proof: If $p = q$ then

$$n_{bb} + (1/2)n_{bw} = n_{ww} + (1/2)n_{bw}$$

and thus

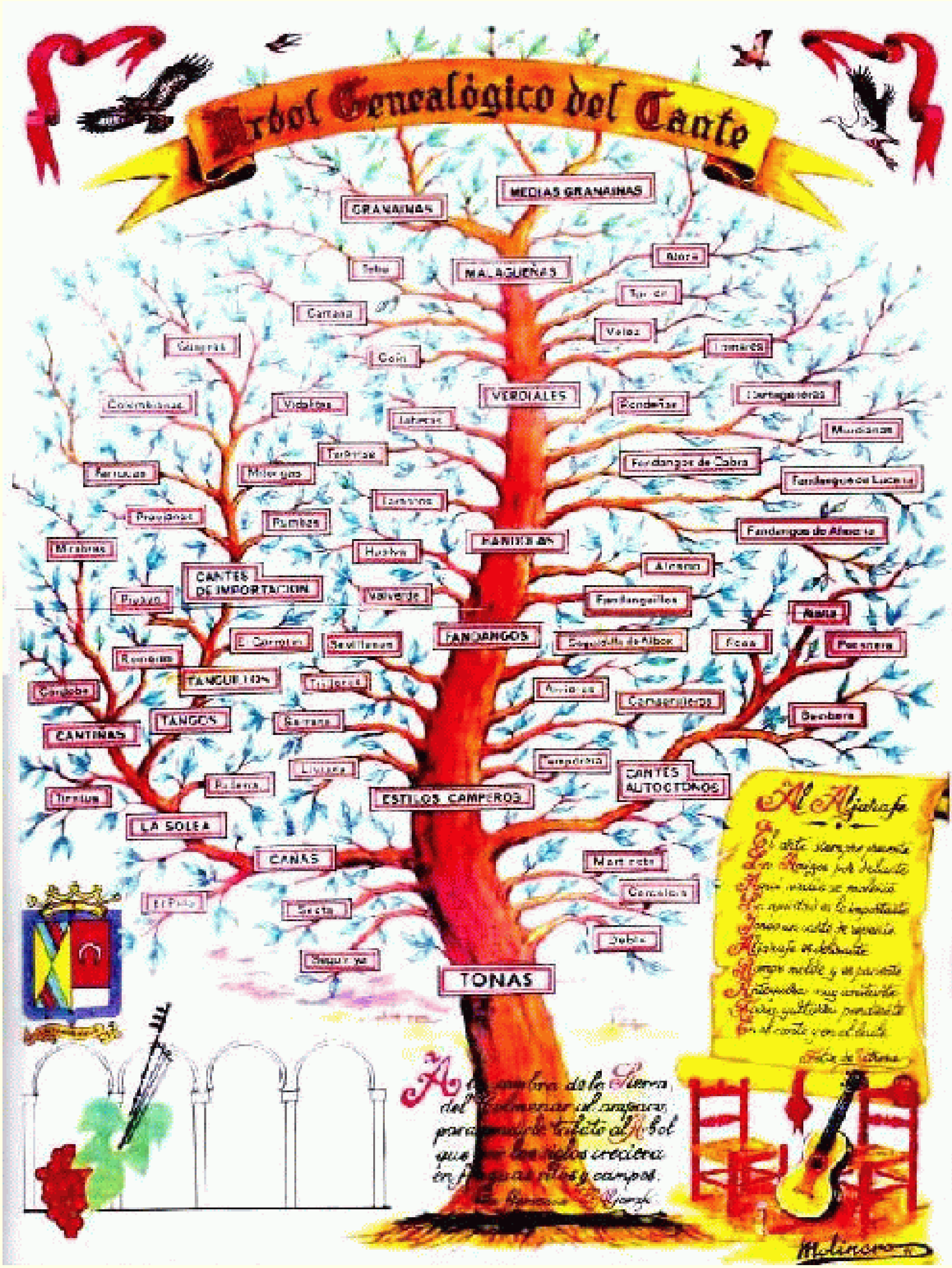
$$n_{bb} = n_{ww}$$

Flamenco Music from *Southern Spain*

Characterized by rhythmic cycles called *compás*, often marked by accented clapping patterns.

Guajira	●			●			●		●		●	

From the Restaurant “Al Aljarate” in Madrid



The Music of Andalucia

The Five 12 -pulse Flamenco Meters

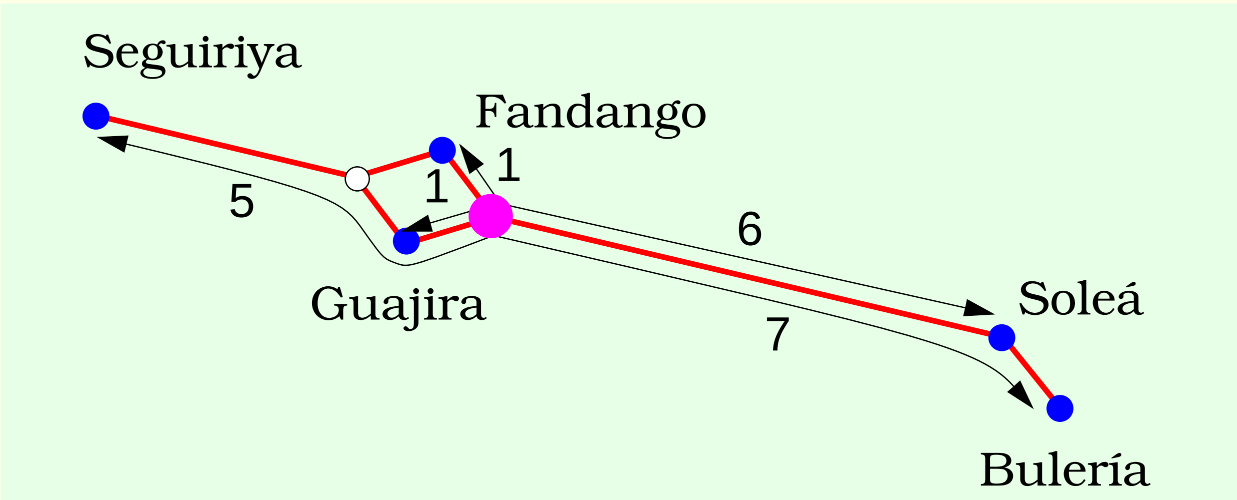
The *SplitsTree* Obtained from the *Directed-swap* Distance Matrix

- *Bulería and Soleá form a very distinct cluster.*
  (only Bulería and Soleá have anacrusis)

- *Fandango and Guajira form a center cluster.*

- *Seguiriya is a singleton cluster.*

Reconstructing an *“Ancestral”* Rhythm from the *Flamenco* Meters with the *Directed-swap* Distance

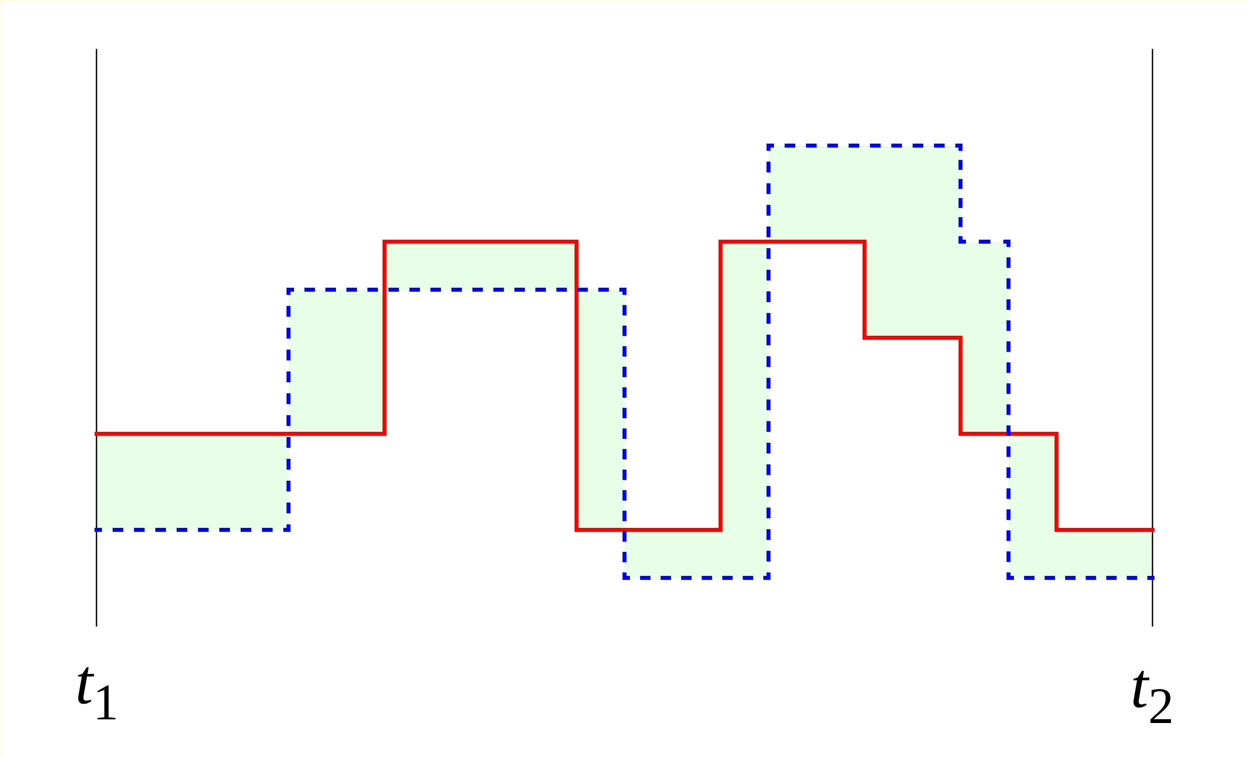


Fandango	●			●			●			●		
Soleá			●			●		●		●		●
Bulería			●				●	●		●		●
Seguiriya	●		●		●			●			●	
Guajira	●			●			●		●		●	
ancestral rhythm	●			●			●			●	●	

Measuring the Similarity of Melody.

D. O'Maidín, "A geometrical algorithm for melodic difference," 1998.

- ✓ *Models a melody by a rectilinear monotonic pitch-duration function of time.*
- ✓ *Computes the area-difference between two melodies.*



Computing Area-difference Between Two Melodies.

G. Aloupis, T. Fevens, S. Langerman, T. Matsui, A. Mesa, Y. Nuñez, D. Rappaport and G. Toussaint
WADS 2003

✓ *O(n) algorithm for fixed* θ.

✓ $O(n^2 \log n)$ *for unrestricted rigid motions.*