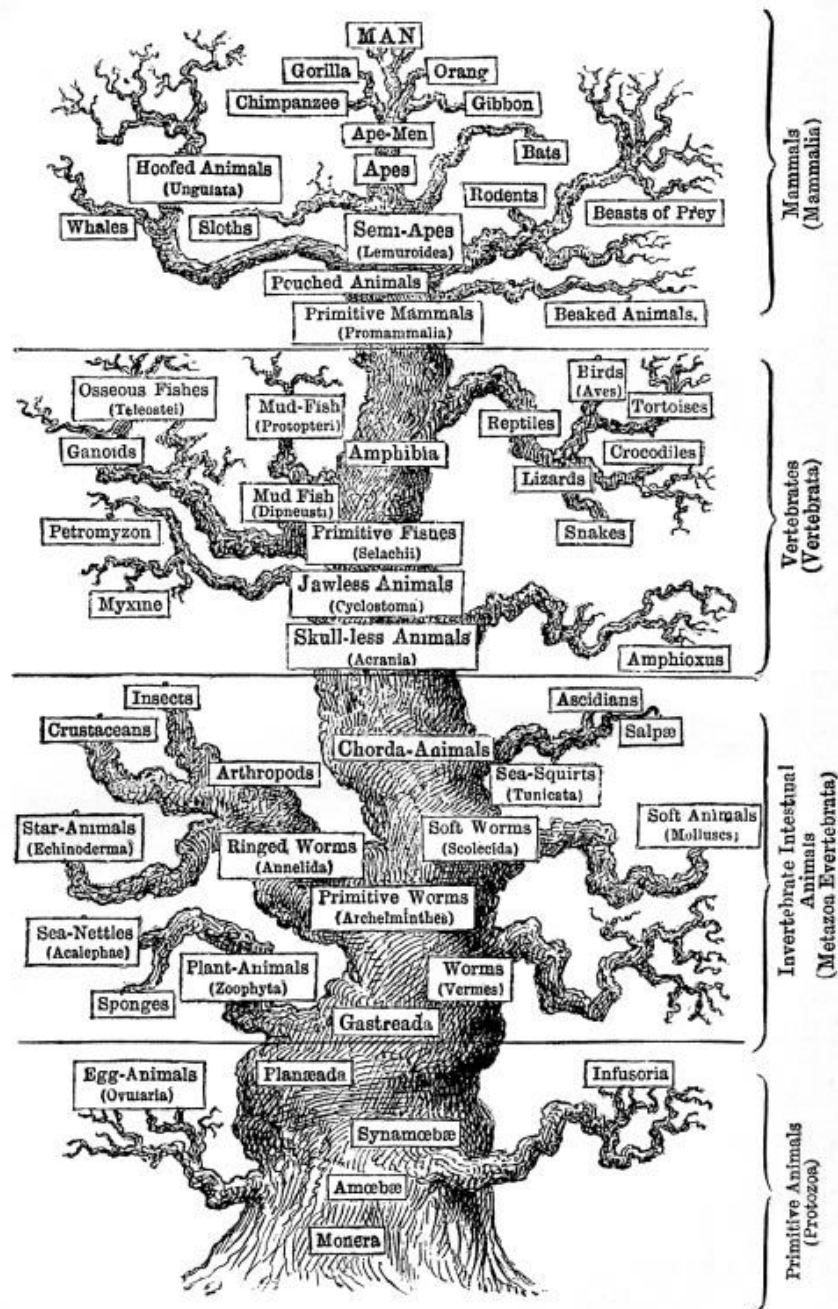


# INFERENCIA FILOGENÉTICA

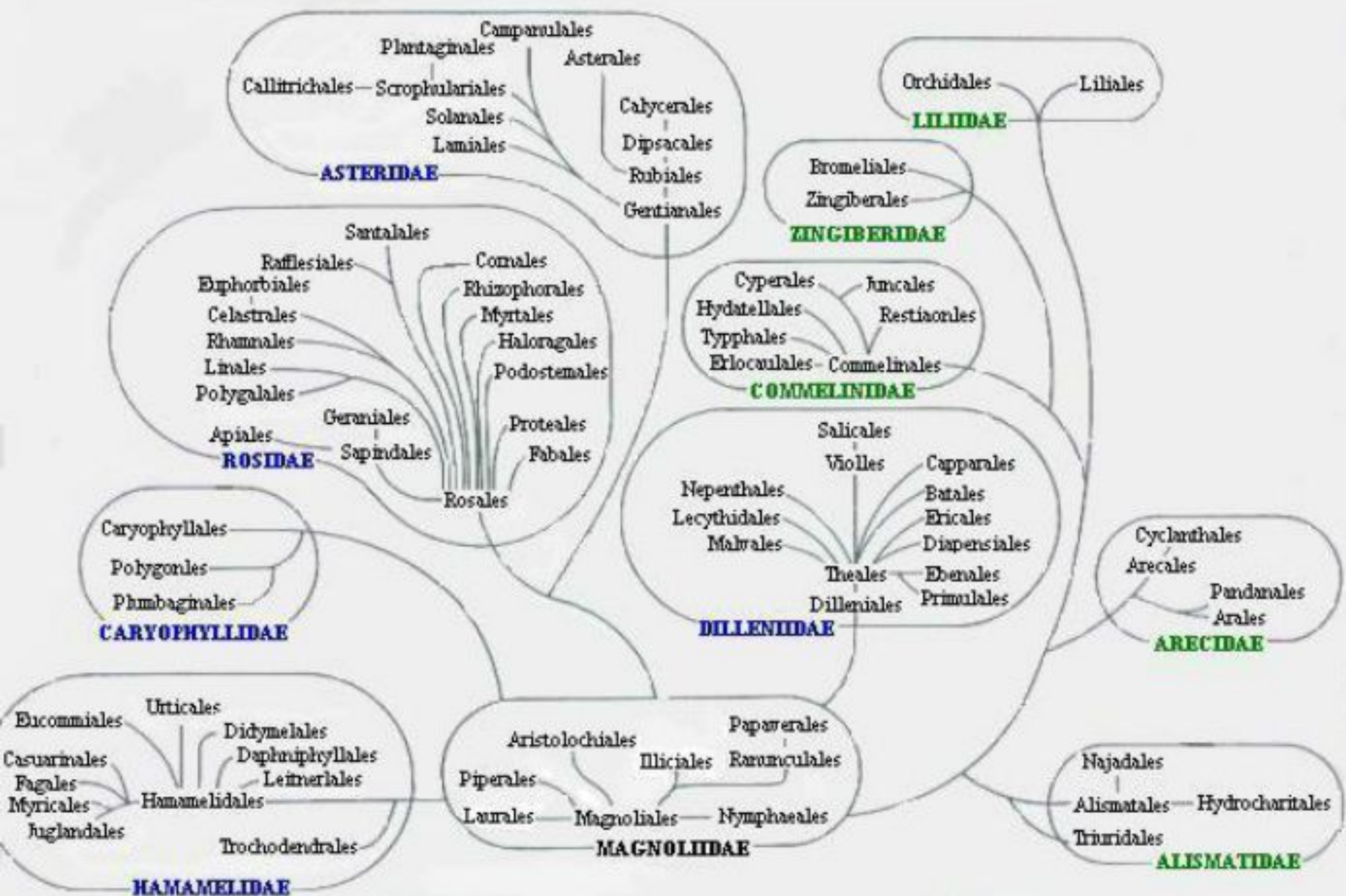
---

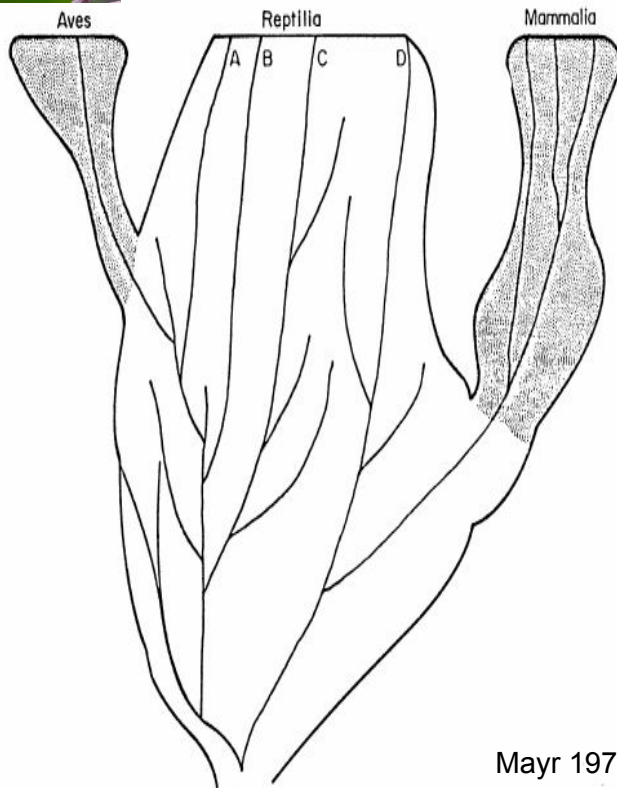
ORÍGENES, CRITERIO DE OPTIMALIDAD Y  
MÁXIMA PARSIMONIA

## PEDIGREE OF MAN.

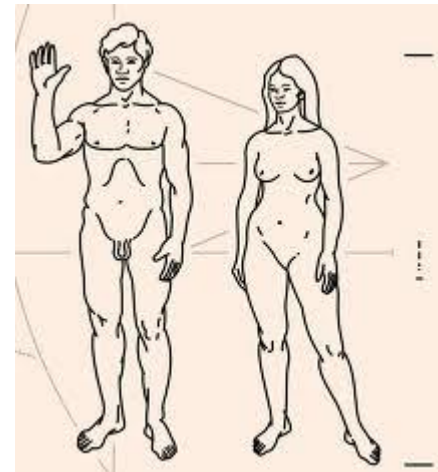


# Classification of Angiosperms by Cronquist 1981



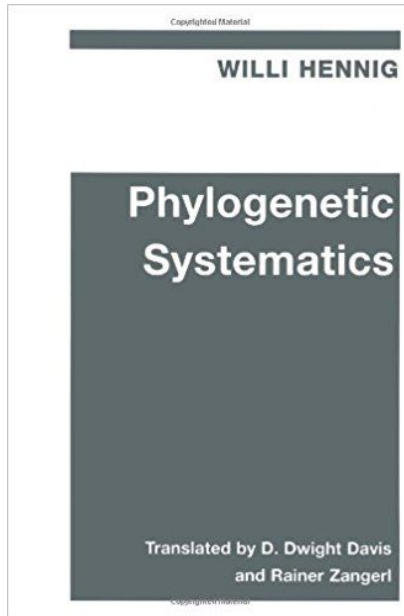


Mayr 1974



Phylum Psychozoa

# SISTEMÁTICA FILOGENÉTICA



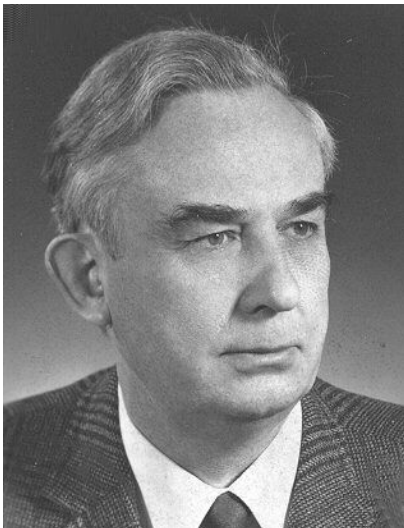
**Objetivo:** reconstruir las relaciones de parentesco entre taxones y proveer una clasificación concordante usando clados monofiléticos

## Axiomas

En la naturaleza hay un orden jerárquico

Los caracteres permiten reconstruir ese orden

Se acepta la reconstrucción con menor número de explicaciones



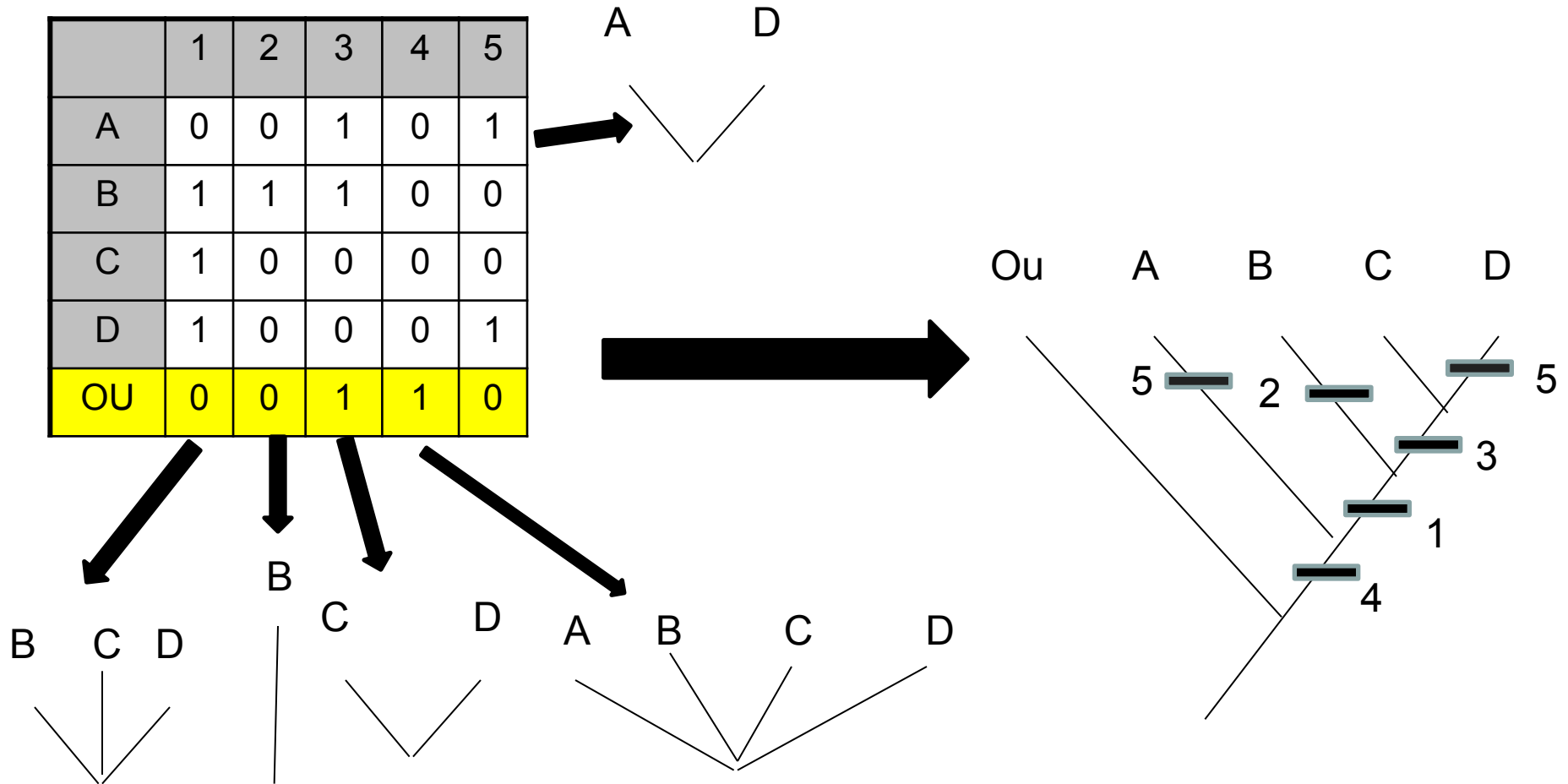
# FUNDAMENTOS DE LA SISTEMÁTICA DE ACUERDO A HENNIG

1. La única base objetiva de relación entre especies es la genealógica
2. La monofilia, como el grupo compuesto por la especie ancestral y todos sus descendientes, define los taxones objetivamente
3. Especies y taxones superiores pueden definirse como monofiléticos si y solo si comparten una novedad evolutiva única (sinapomorfía)

# ARGUMENTACIÓN HENNIGIANA

- Determinar homología primaria
- Determinar **polaridad** (Grupo ajeno)
- Identifique congruencias
- Construya el o los árboles

# ARGUMENTACIÓN HENNIGIANA





# TALLER: Argumentación Hennigiana

Infiera las relaciones filogenéticas de las monocotiledóneas con base en la siguiente matriz:

	Flores trimeras	Trepadora	Cuerpos de sílica	Cera tipo "Strelitzia"	Plantas dioicas	Fruto capsular	Venación palmeada	Ovario súpero
Algodón	0	0	0	0	0	0	0	0
Ñame	1	1	0	0	1	0	0	1
Coco	1	0	0	0	1	1	1	0
Piña	1	0	1	1	0	1	1	1
Banano	1	0	1	1	0	1	1	1

# Abrir enlace del chat:

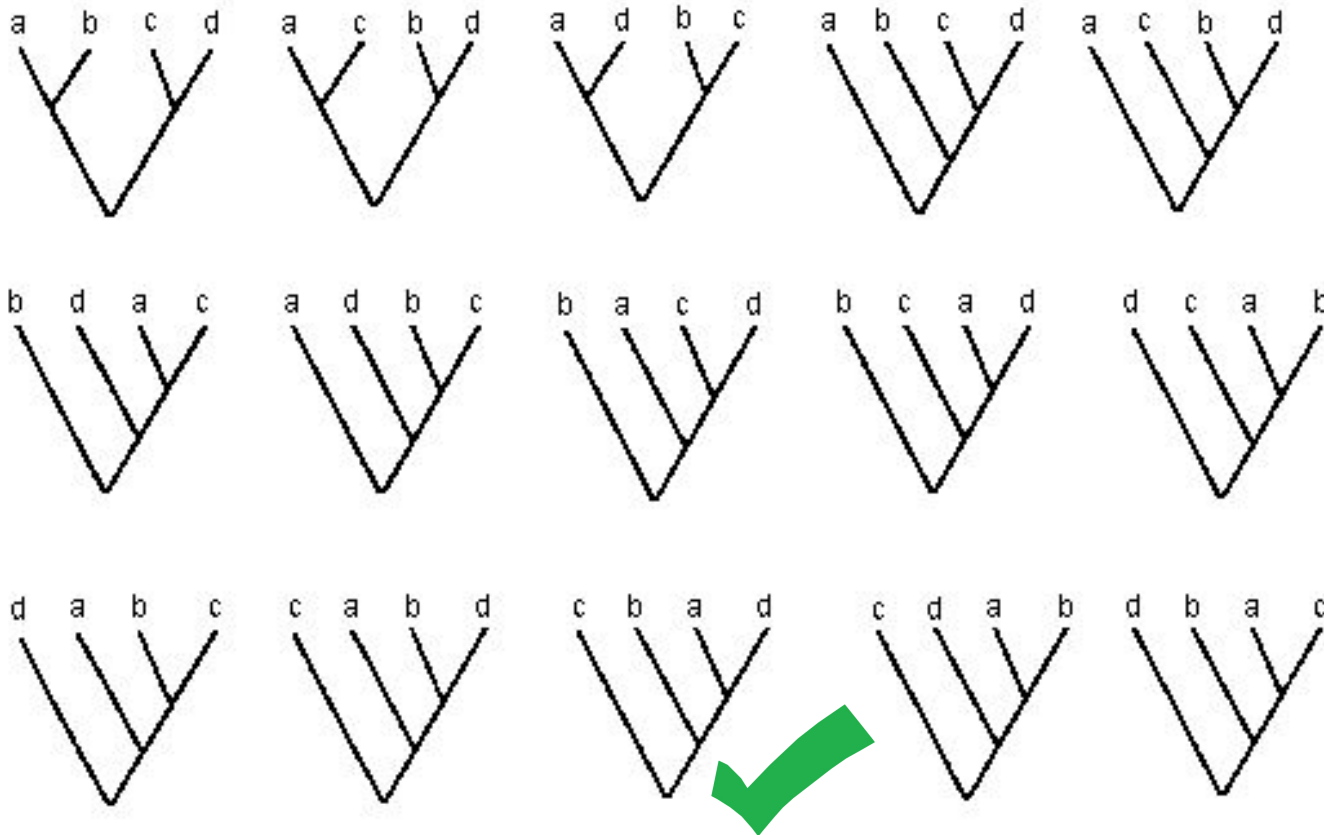
<https://jamboard.google.com/d/1aRz1gV2ikx6hH1BjWx4u50ZJR0MkOWbyRW3e8-lumKk/edit?usp=sharing>

# PROBLEMAS DE LA ARGUMENTACIÓN HENNIGIANA

- Polarización *a priori* de caracteres (grupo ajeno = ancestral)
- El método asume que **no hay homoplasia** (poco realista)
  - Caracteres inconsistentes con otros violan el modelo Hennigiano
- Imposible evitar errores o malas interpretaciones al codificar caracteres

# CRITERIO DE OPTIMALIDAD

Medida que permite decidir, con base en un conjunto de datos, cuáles árboles (hipótesis) son mejores y cuales son peores



# MÁXIMA PARSIMONIA

**Cuchilla de Occam:** la mejor hipótesis para explicar un proceso es aquella que requiere el menor número de suposiciones

**En inferencia filogenética:** el mejor árbol es aquel que explica los datos observados con la menor cantidad de homoplasia posible (menos transformaciones)

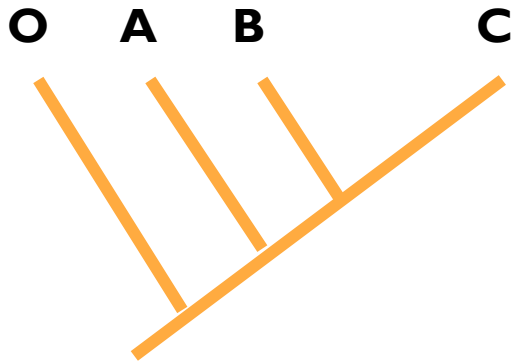
# MÁXIMA PARSIMONIA

1. Contar el **mínimo** número de cambios (pasos) de cada caracter en un árbol determinado
2. Sumar todos los números de pasos para determinar la **LONGITUD DEL ÁRBOL**
3. Repetir en los otros árboles alternativos y escoger aquel con la menor longitud como el **ÁRBOL MÁS PARSIMONIOSO**

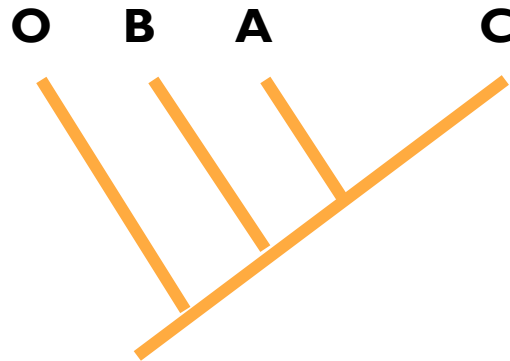
# MÁXIMA PARSIMONIA

## EJEMPLO

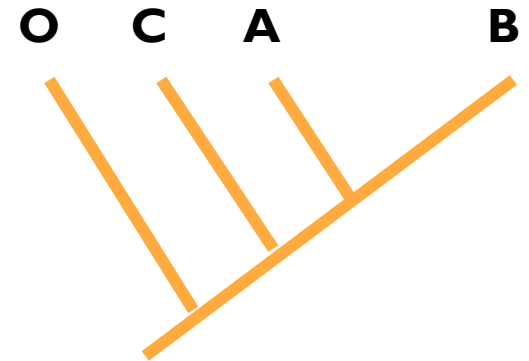
	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0



Árbol 1



Árbol 2



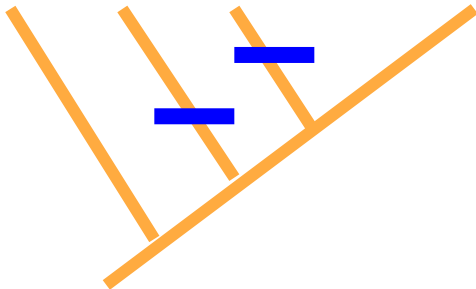
Árbol 3

# MÁXIMA PARSIMONIA

Árbol I

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0

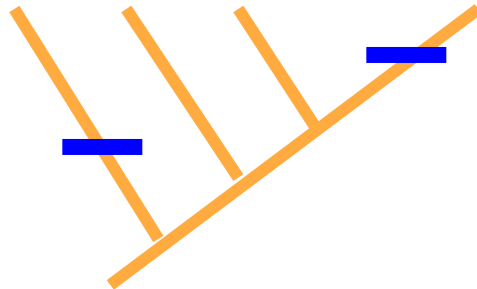
O = 0 A = 1 B = 1 C = 0



0 → 1

2 pasos

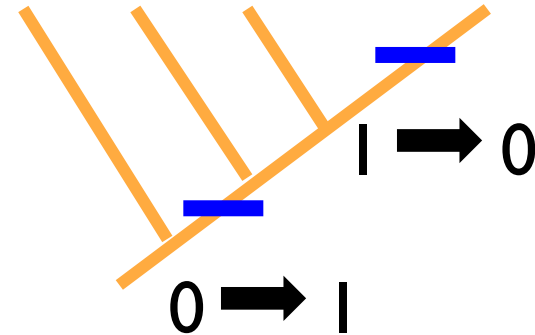
O = 0 A = 1 B = 1 C = 0



1 → 0

2 pasos

O = 0 A = 1 B = 1 C = 0



0 → 1

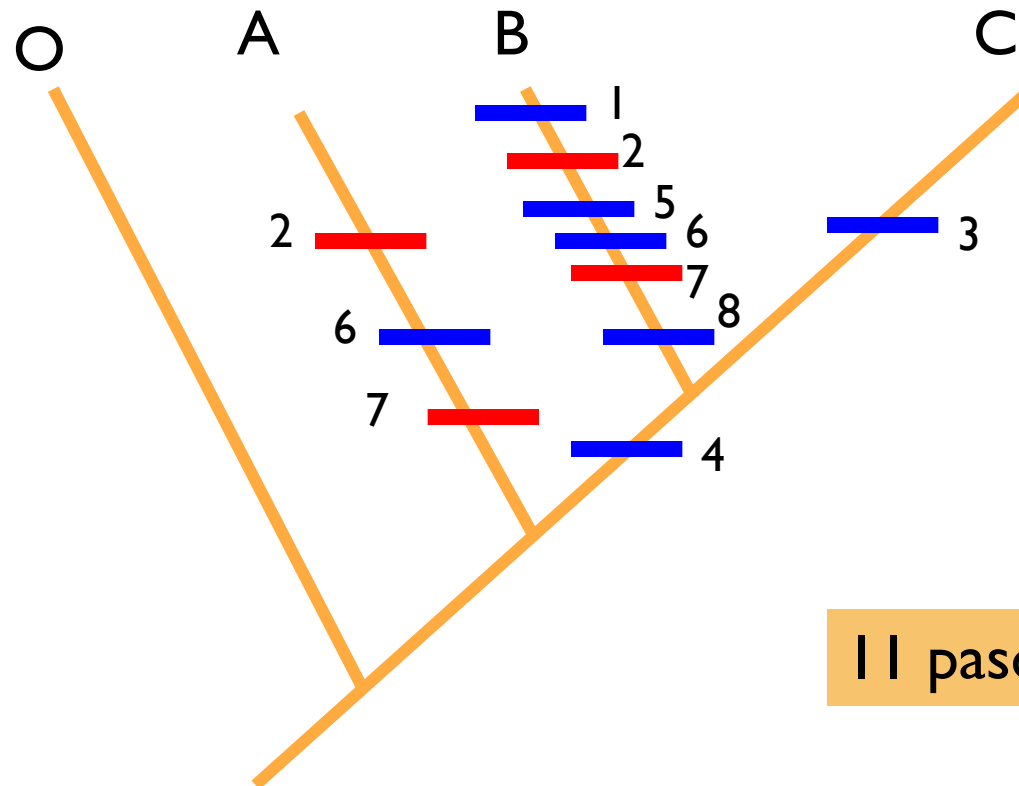
2 pasos



# MÁXIMA PARSIMONIA

Árbol I

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0

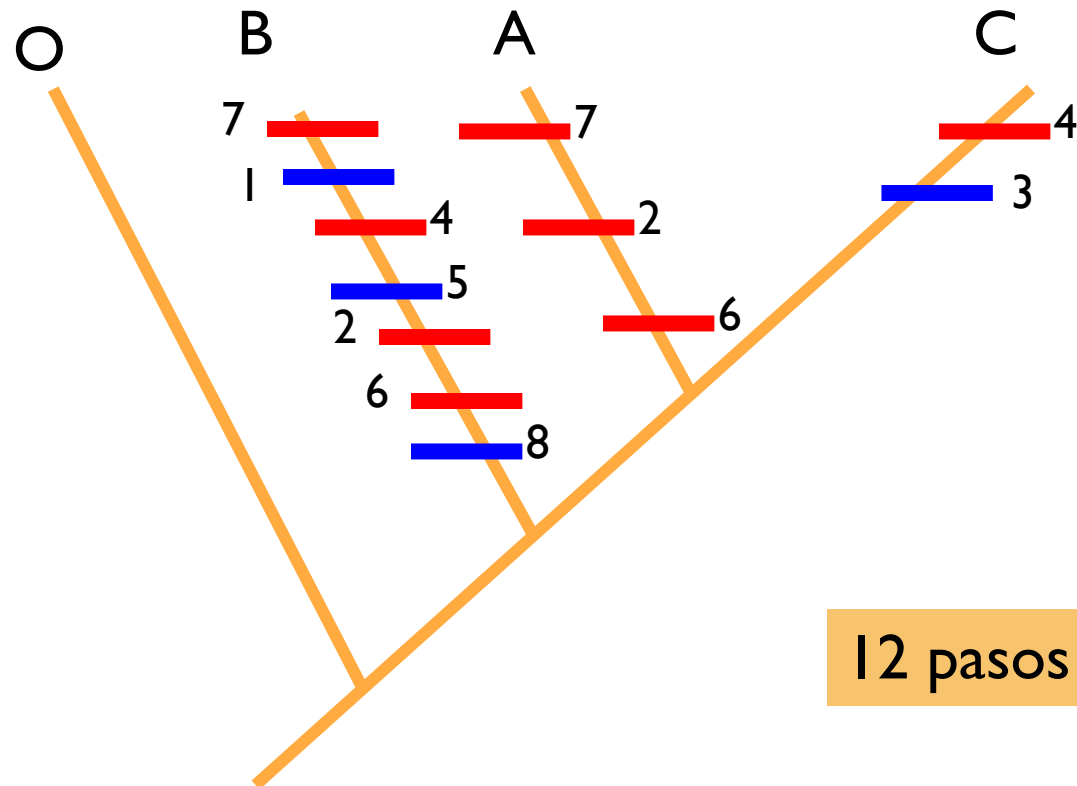


11 pasos

# MÁXIMA PARSIMONIA

Árbol 2

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0

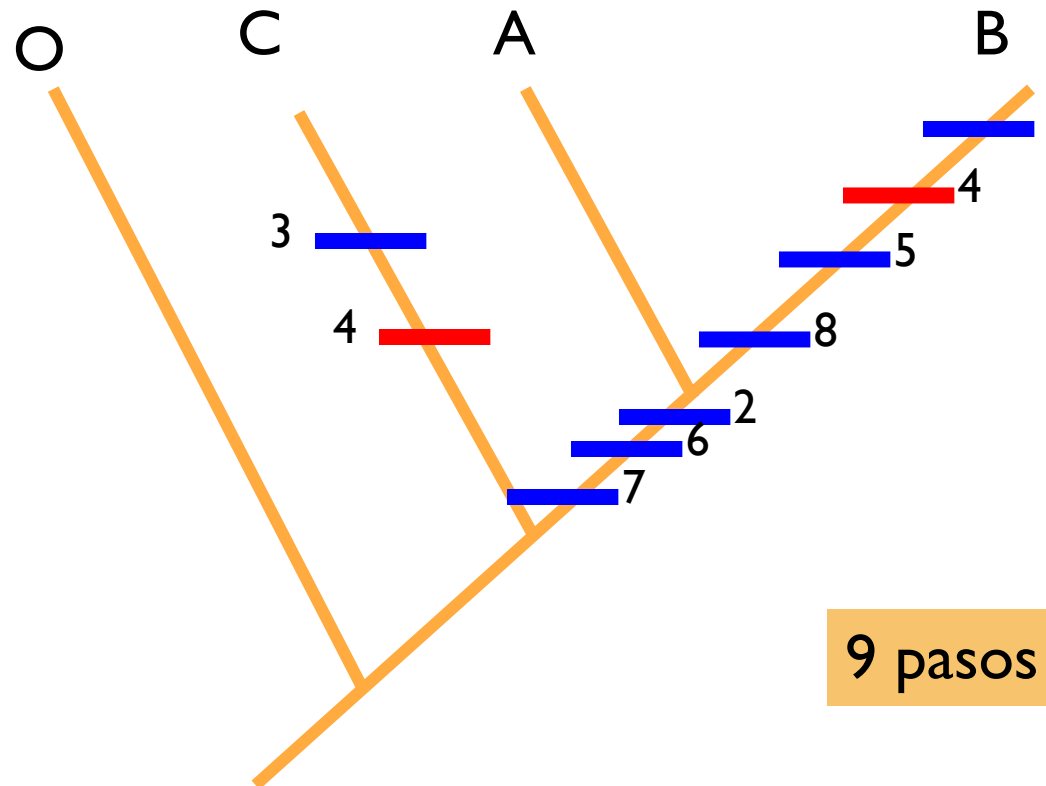


12 pasos

# MÁXIMA PARSIMONIA

Árbol 3

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0



# MÁXIMA PARSIMONIA

	1	2	3	4	5	6	7	8	
O	0	0	1	0	1	1	0	0	
A	0	1	1	0	1	0	1	0	
B	1	1	1	1	0	0	1	1	
C	0	0	0	1	1	1	0	0	Total length
Length on tree 1	1	2	1	1	1	2	2	1	11
Length on tree 2	1	2	1	2	1	2	2	1	12
Length on tree 3	1	1	1	2	1	1	1	1	9

Árbol más  
parsimonioso

# MÁXIMA PARSIMONIA

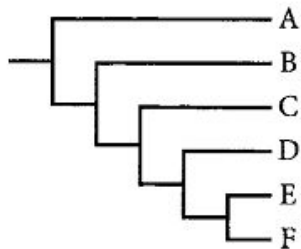
- HOMOPLASIA VS. CONSISTENCIA

Taxa	Consistent					Homoplastic				
A	A	G	T	G	G	G	C	G	A	T
B	A	G	T	G	G	G	T	T	C	G
C	A	G	T	G	C	A	T	G	G	A
D	A	G	T	T	T	A	C	T	C	T
E	A	A	C	T	A	G	C	A	A	C
F	A	G	C	C	A	G	T	A	G	G
States	1	2	2	3	4	2	2	3	3	4
Changes	0	1	1	2	3	2	3	3	4	4
CI		1.0	1.0	1.0	1.0	0.5	0.33	0.66	0.5	0.75

Índice de consistencia (ci)

$$ci = L_{\min} / L_{\text{obs}}$$

$$L_{\min} = \# \text{ de estados} - 1$$



# MÁXIMA PARSIMONIA

- **Otros índices**

**Índice de Homoplasia (hi) =  $1 - CI$**

**Índice de Retención (ri) =  $(L_{\max} - L_{\text{obs}}) / (L_{\max} - L_{\min})$**

- **Índices a través del árbol**

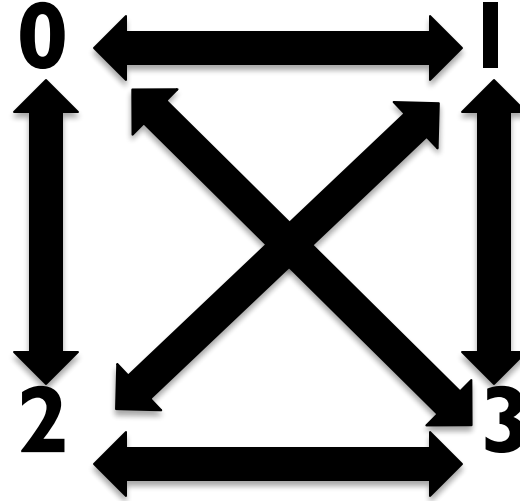
$$CI = \Sigma L_{\min} / \Sigma L_{\text{obs}}$$

$$RI = (\Sigma L_{\max} - \Sigma L_{\text{obs}}) / (\Sigma L_{\max} - \Sigma L_{\min})$$

# MÁXIMA PARSIMONIA

## Variaciones de optimización de caracteres en árboles

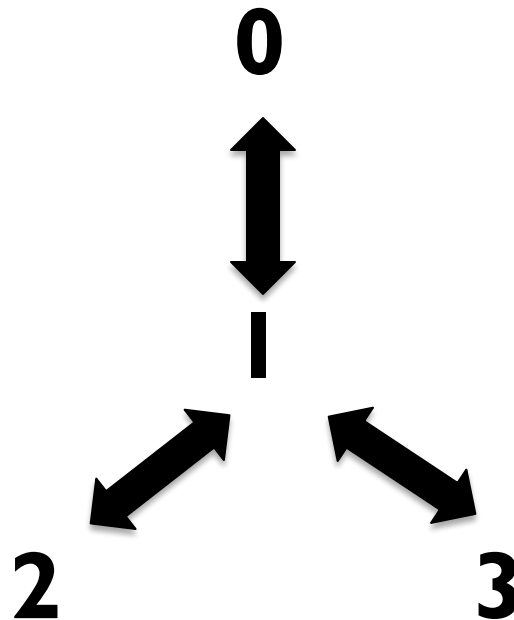
- Caracteres no ordenados o no aditivos (pesos iguales) = **Parsimonia de Fitch**



# MÁXIMA PARSIMONIA

## Variaciones de optimización de caracteres en árboles

- Caracteres ordenados o aditivos (pesos diferentes de caracter) = **Parsimonia de Wagner**



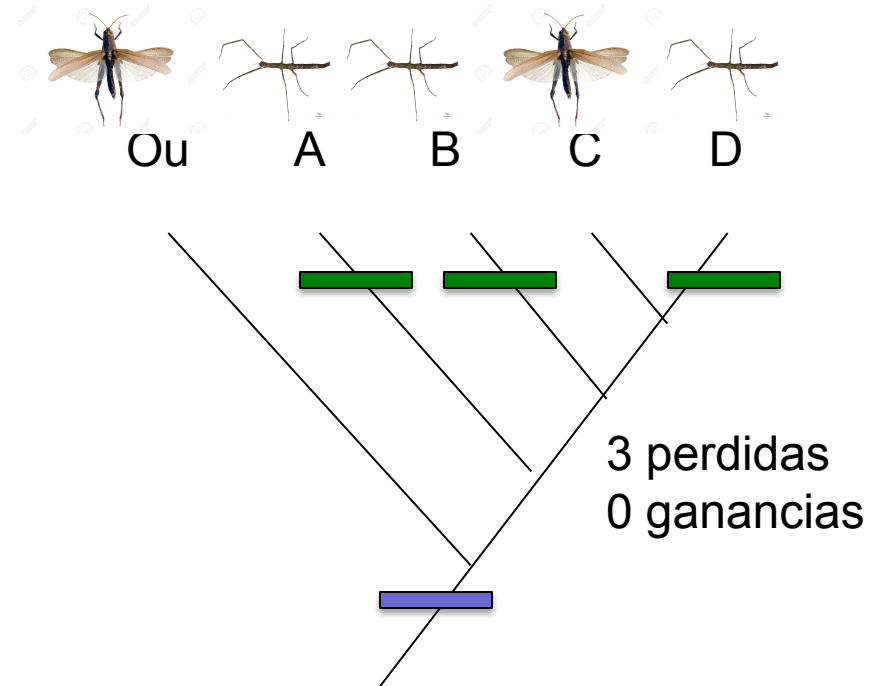
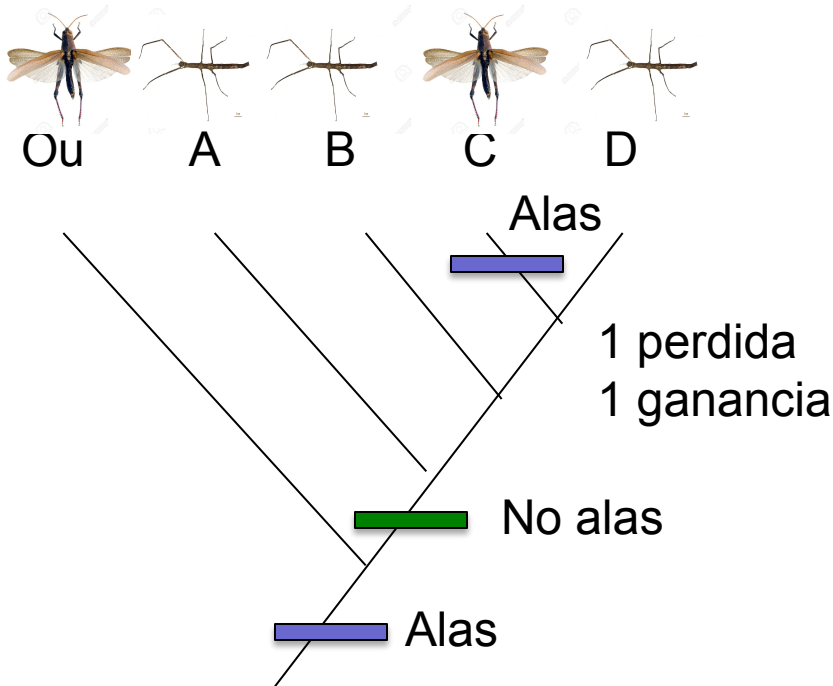


# MÁXIMA PARSIMONIA

# Variaciones de optimización de caracteres en árboles

# Parsimonia de Dollo

- Una vez un carácter complejo se puede perder muchas veces pero no podrá evolucionar de nuevo



# MÁXIMA PARSIMONIA

## Variaciones de optimización de caracteres en árboles

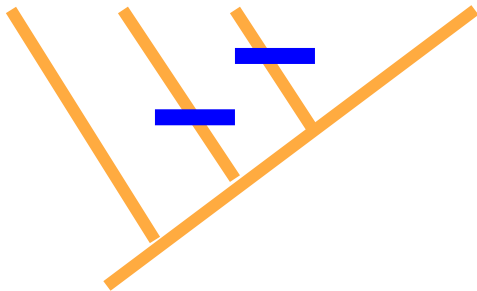
- PESAJE DE CARACTERES
  - **Pesaje a priori (Parsimonia generalizada)**
    - Asignación de peso a criterio del investigador
    - Ej.: transversiones vs. transiciones, codones, caracteres diagnósticos
  - **Pesos a posteriori**
    - Pesaje a posteriori después de un análisis de pesos iguales (homólogos pesan más)

# MÁXIMA PARSIMONIA

	1	2	3	4	5	6	7	8		
O	0	0	1	0	1	1	0	0		
A	0	1	1	0	1	0	1	0		
B	1	1	1	1	0	0	1	1		
C	0	0	0	1	1	1	0	0		
Peso	1	1	1	5	1	1	1	1	Longi-tu d	Costo total
Costo árbol 1	1	2	1	5	1	2	2	1	11	15
Costo árbol 2	1	2	1	10	1	2	2	1	12	20
Costo árbol 3	1	1	1	10	1	1	1	1	9	17

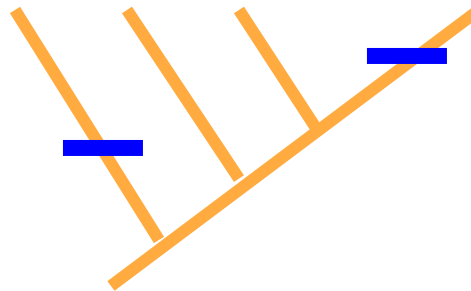
# MÁXIMA PARSIMONIA

## ACCTTRAN y DELTRAN



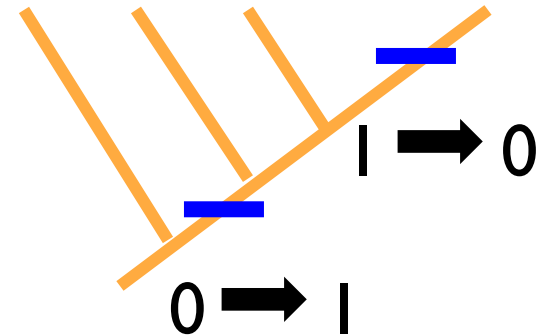
0 → 1

## 2 pasos



1 → 0

## 2 pasos

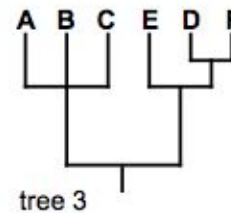
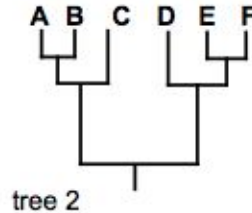
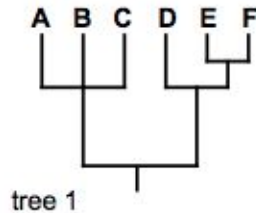


## 2 pasos

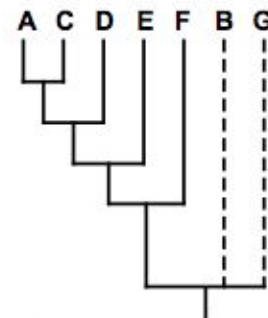
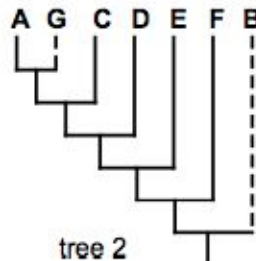
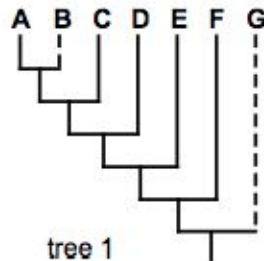
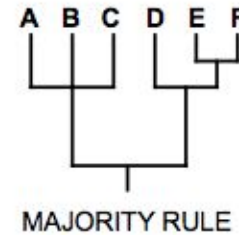
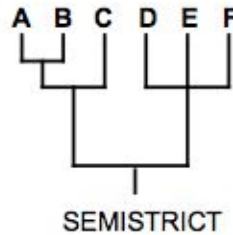
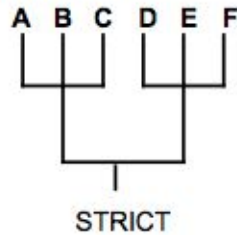
- ACCTTRAN: Transformación acelerada (favorece reversiones)
- DELTRAN: Transformación retrasada (favorece los paralelismos)

# MÁXIMA PARSIMONIA

¿Qué pasa si hay más de un árbol más parsimonioso?



Consensus trees



# MÁXIMA PARSIMONIA

## Problemas

- Longitud de ramas no se toma en cuenta (se ignora la tasa de evolución de los caracteres en cada rama)
- Atracción de ramas largas (¡¡entre más caracteres, peor!!)
- Pesaje de caracteres es necesario (aún si son pesos iguales)
  - No hay métodos formales para decidir pesos

# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

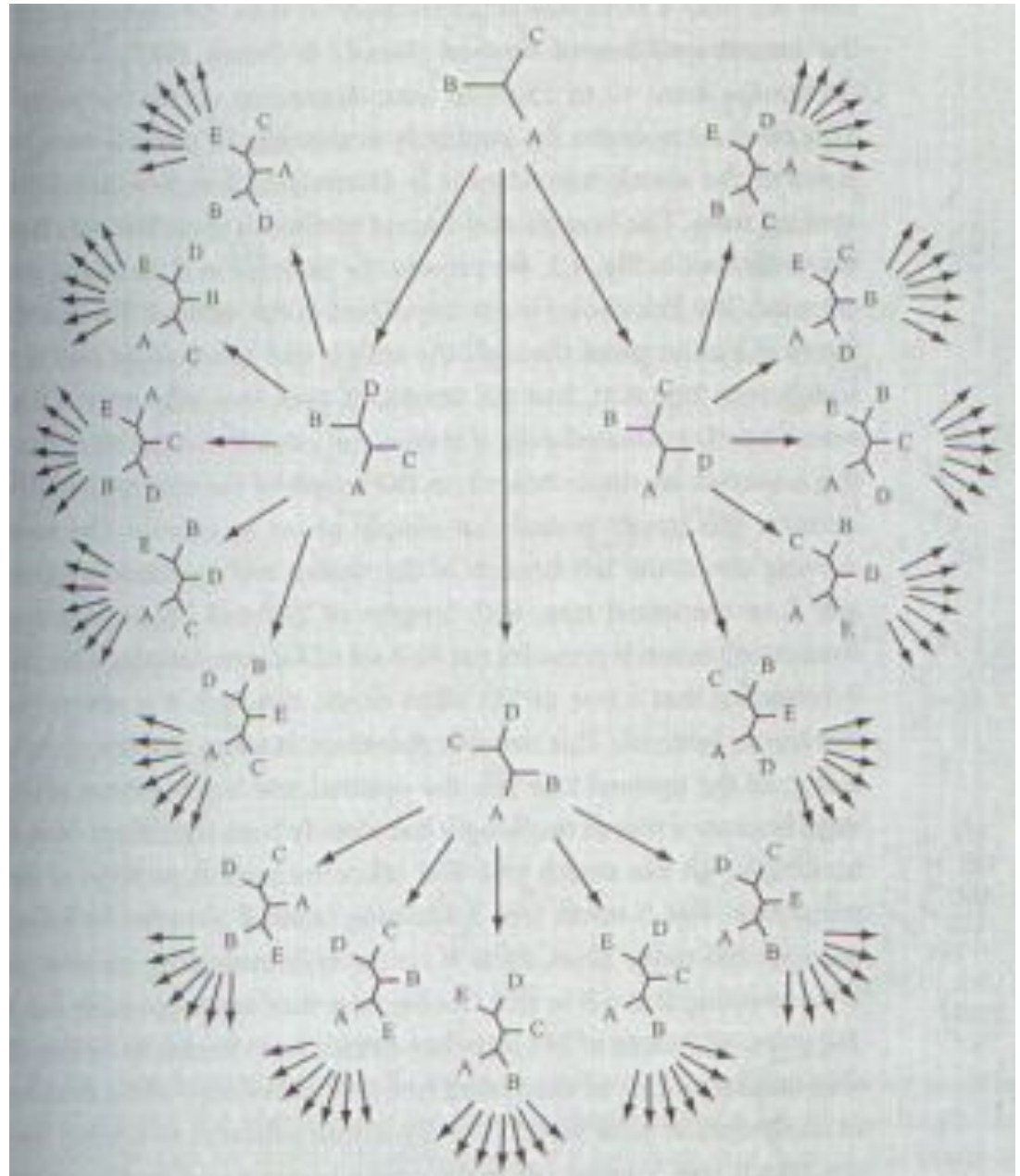
El problema de encontrar  
árboles óptimos...

Taxa	Árboles resueltos
1	--
2	1
3	1
4	3
5	15
6	105
7	945
8	10395
9	135135
10	2027025
11	34459425
12	654729075
13	13749310575
14	316234143225
15	7905853580625
16	213458046676875
17	6190283353629370
18	191898783962510000
19	6332659870762850000
20	221643095476699000000
62	6,66409461 x 10 E 98
63	> 10 E 100

# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos exactos:

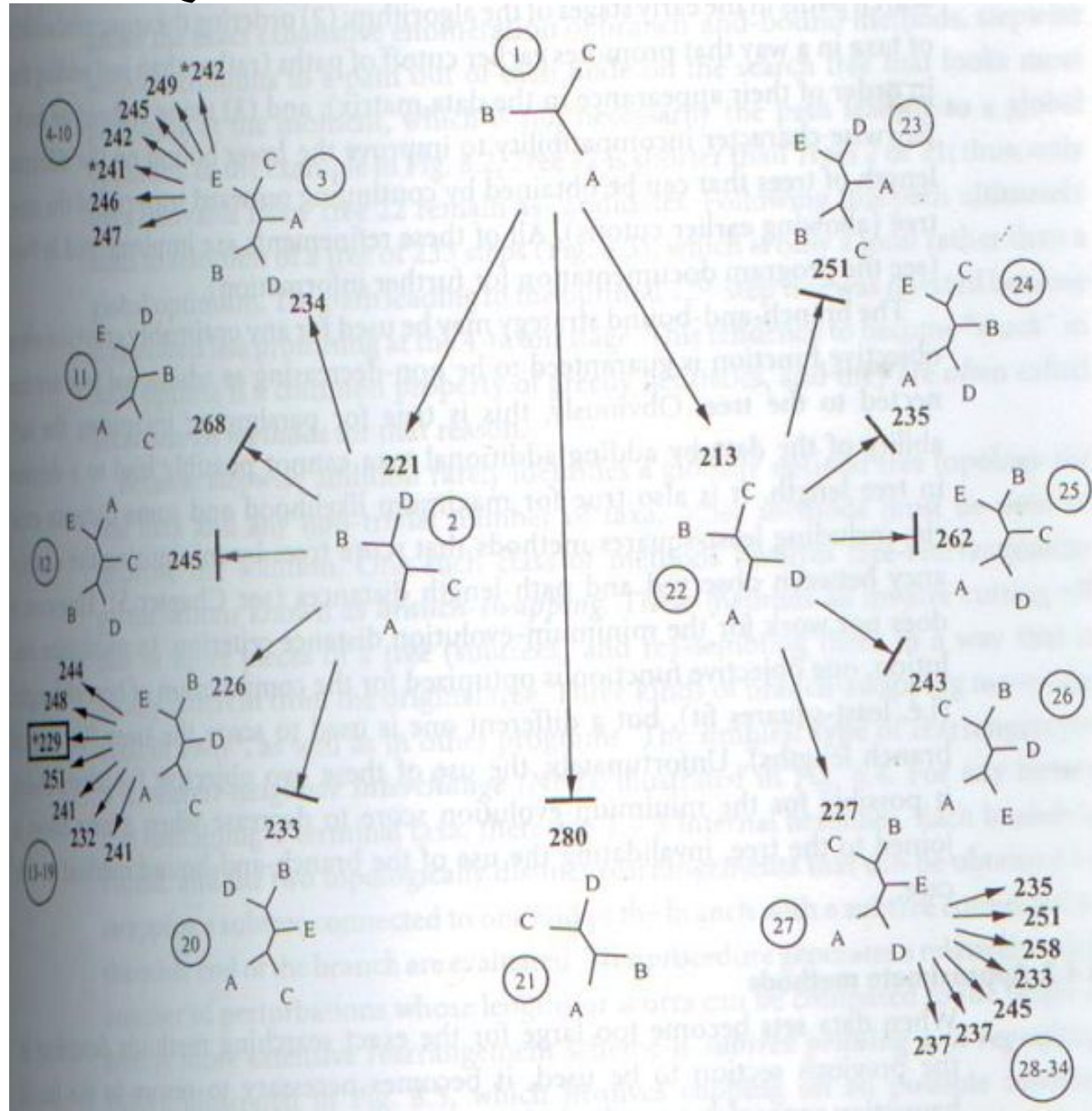
### I. Búsqueda exhaustiva





# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos exactos: 2. Branch & Bound

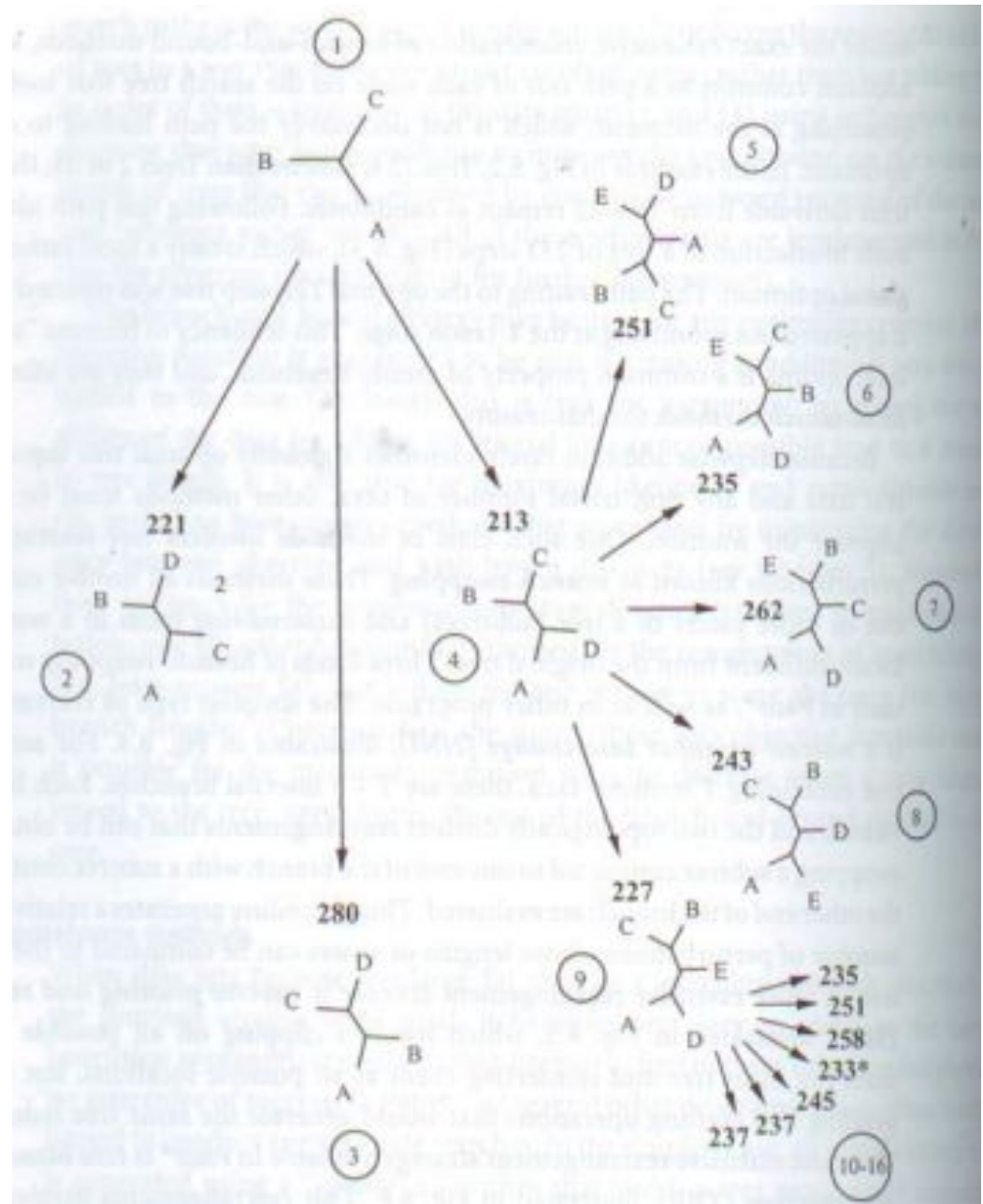


# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos heurísticos:

### I. Buscar árbol inicial:

- Adición paso a paso (Stepwise addition)
- Aleatorio

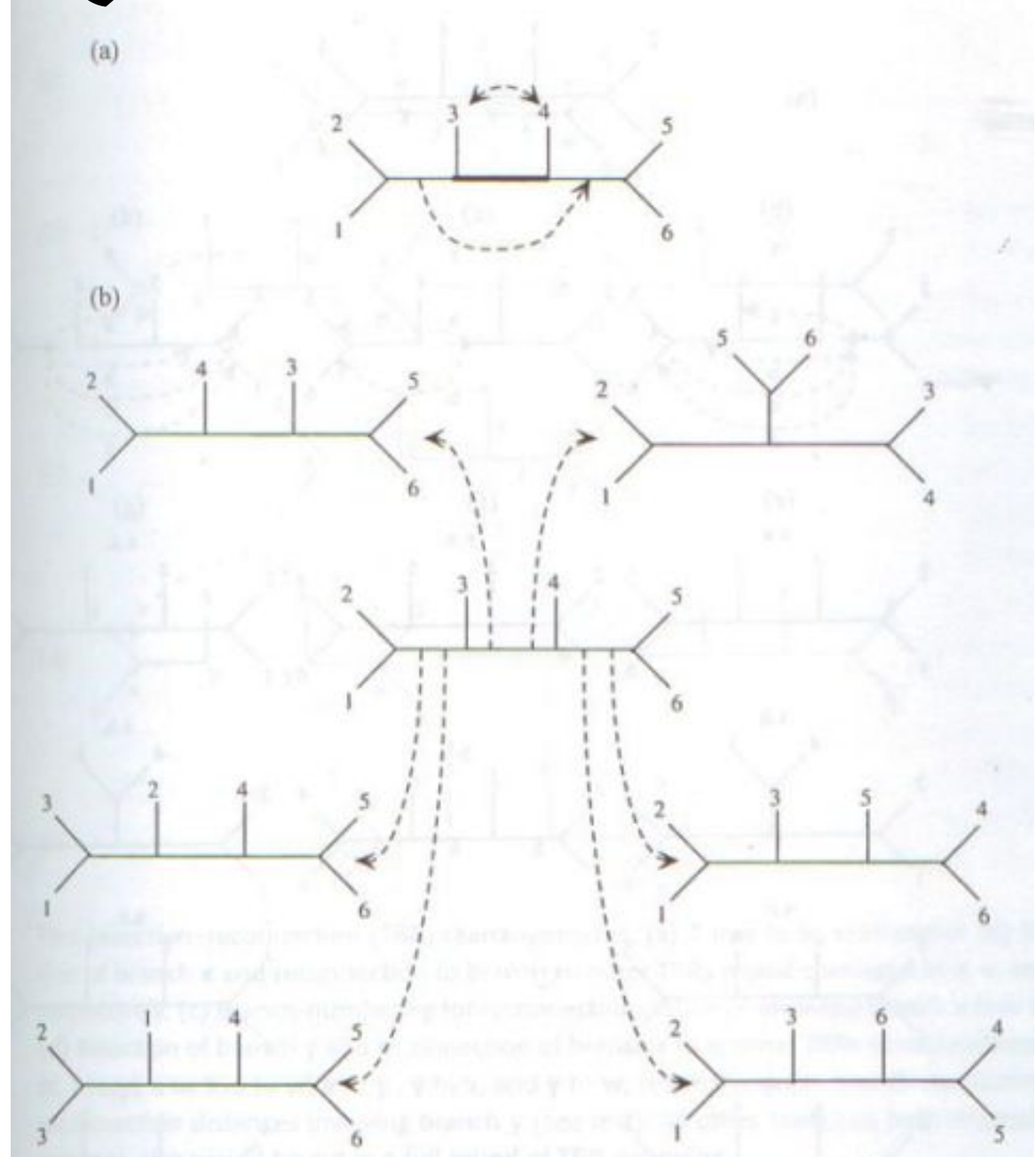


# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos heurísticos:

### 2. Perturbar árbol inicial:

- Nearest Neighbor Interchange (NNI)



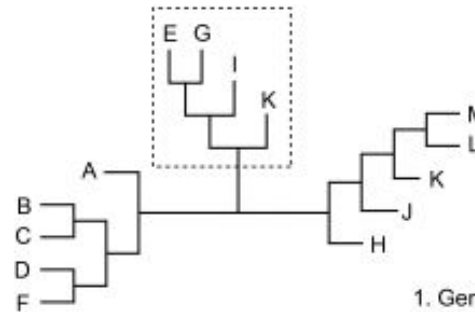
# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos heurísticos:

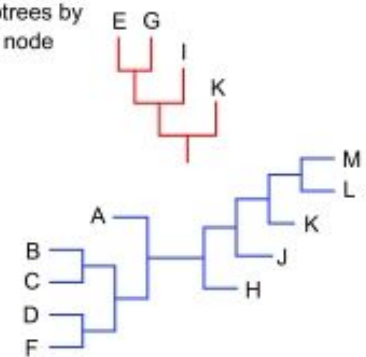
### 2. Perturbar árbol inicial:

- Subtree Pruning & Regrafting (SPR)

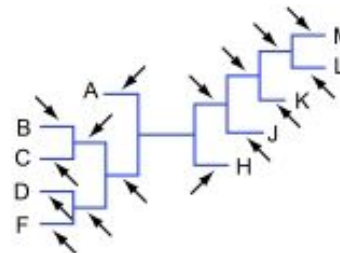
0. Starting tree



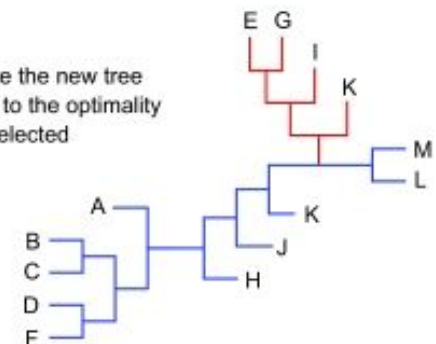
1. Generate two subtrees by breaking an internal node



2. Try to insert the red subtree at each node of the blue subtree



3. Evaluate the new tree according to the optimality criterion selected

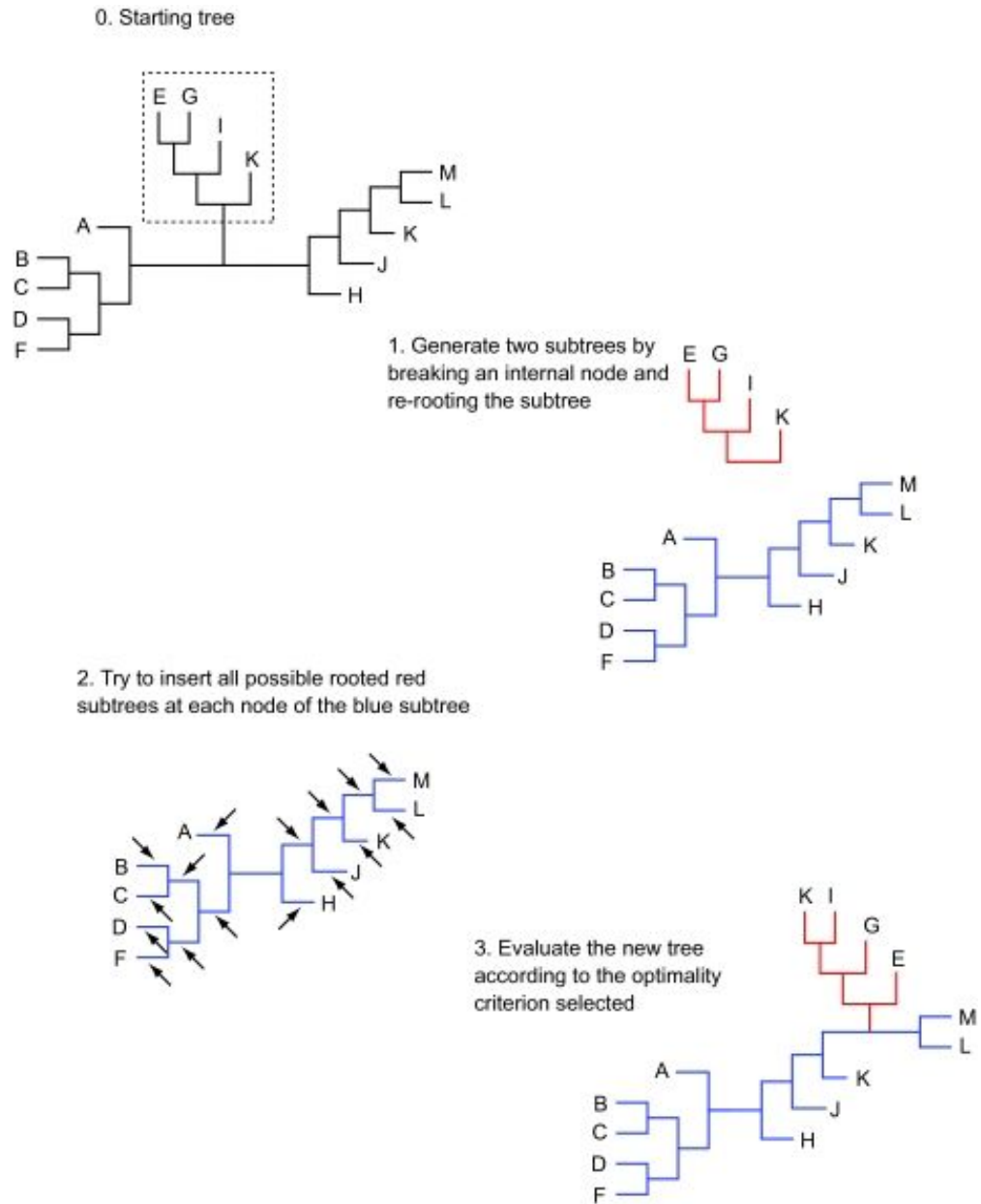


# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos heurísticos:

### 2. Perturbar árbol inicial:

- Tree Bisection & Reconnection (TBR)

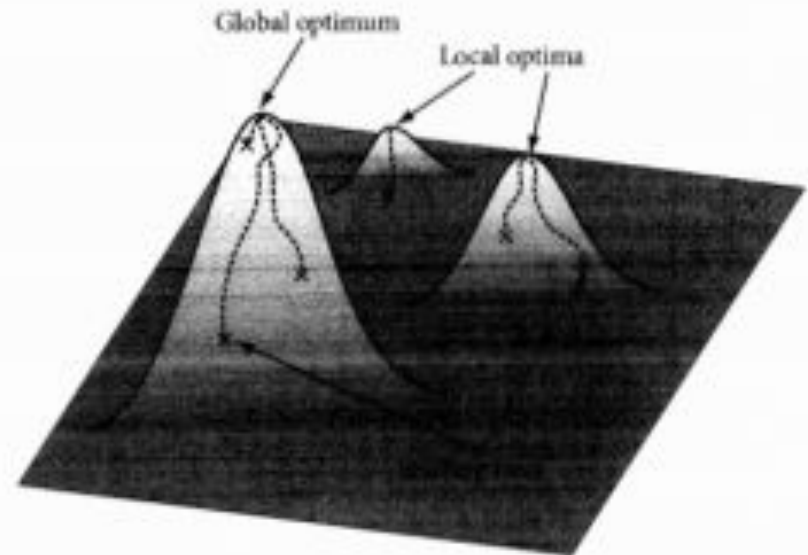
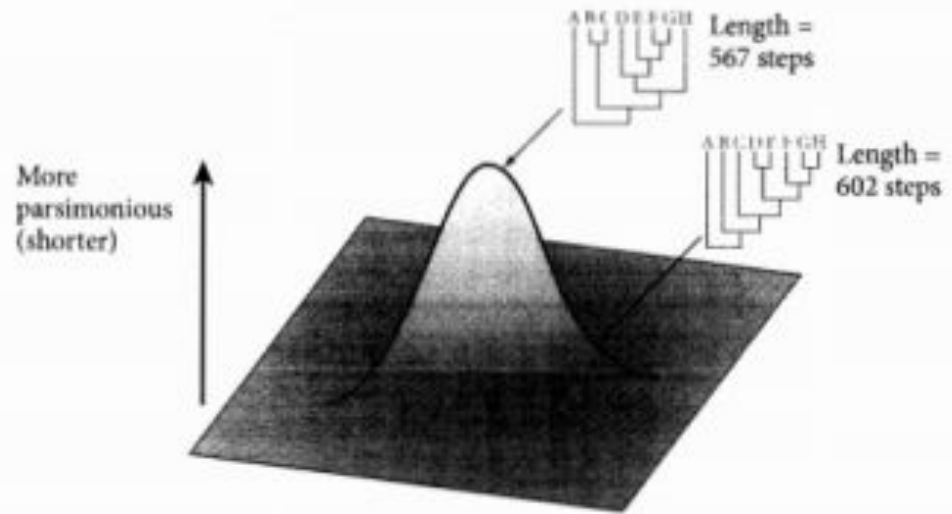


# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos heurísticos:

3. Visitar óptimos locales para tener óptimo global:

- Réplicas
- Stepwise-random-addition



# ESTRATEGIAS DE BÚSQUEDA DE ÁRBOLES ÓPTIMOS

## Métodos heurísticos:

Para más de 100 terminales:

- Nueva Tecnología: Parsimonia RATCHET y Tree-Drifting
  - Sacrifican búsquedas intensivas en islas para poder visitar más islas en el espacio de árboles (escapar de óptimos locales).
  - 2 pasos:
    - Búsquedas en subset de datos con nuevos pesos (para explorar islas)
    - Volver a pesos originales y escoger mejores árboles