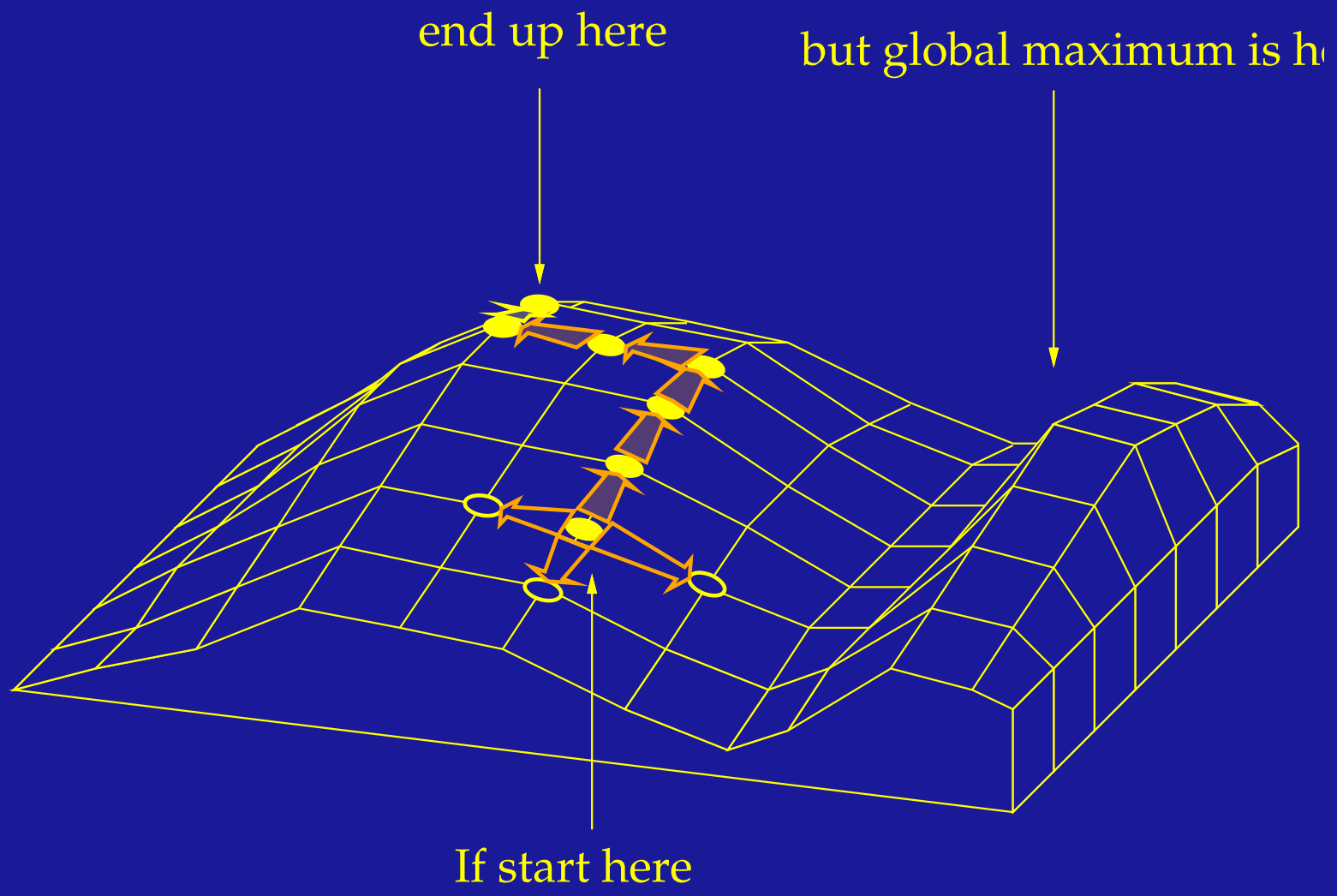
April 2004
Genome 570

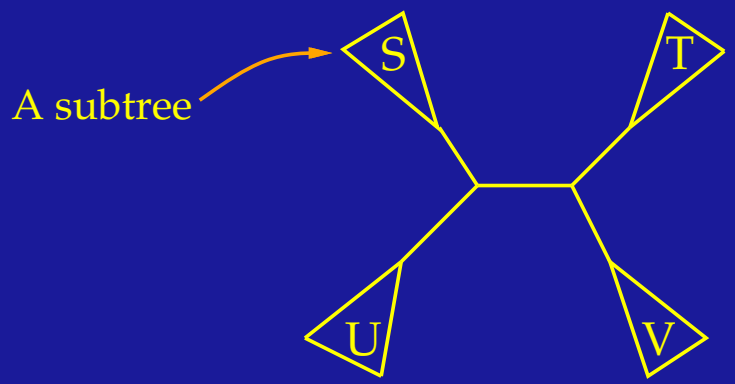
Inferring Phylogenies
Week 2: Searching for trees, ancestral states

Joe Felsenstein

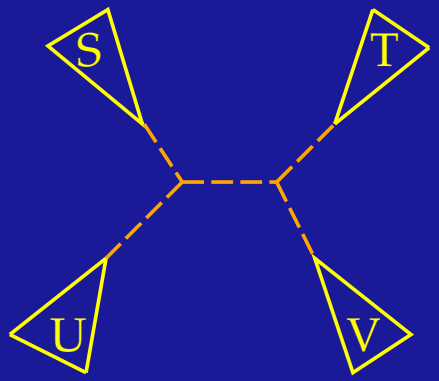
Department of Genome Sciences and Department of Biology
University of Washington, Seattle

email: joe@gs.washington.edu

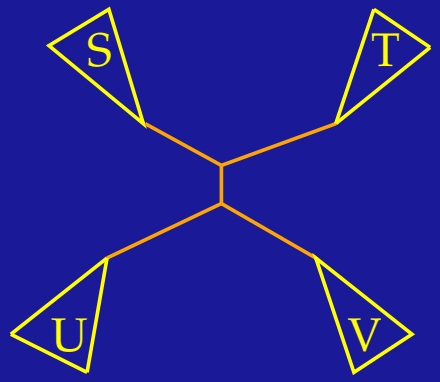

end up here
but global maximum is h
If start here

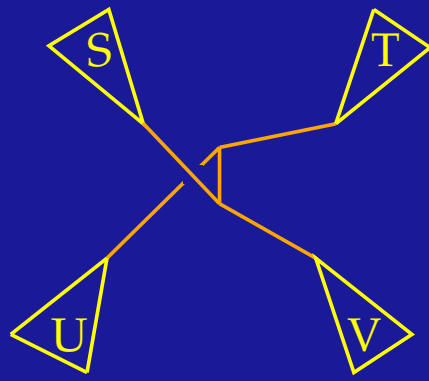


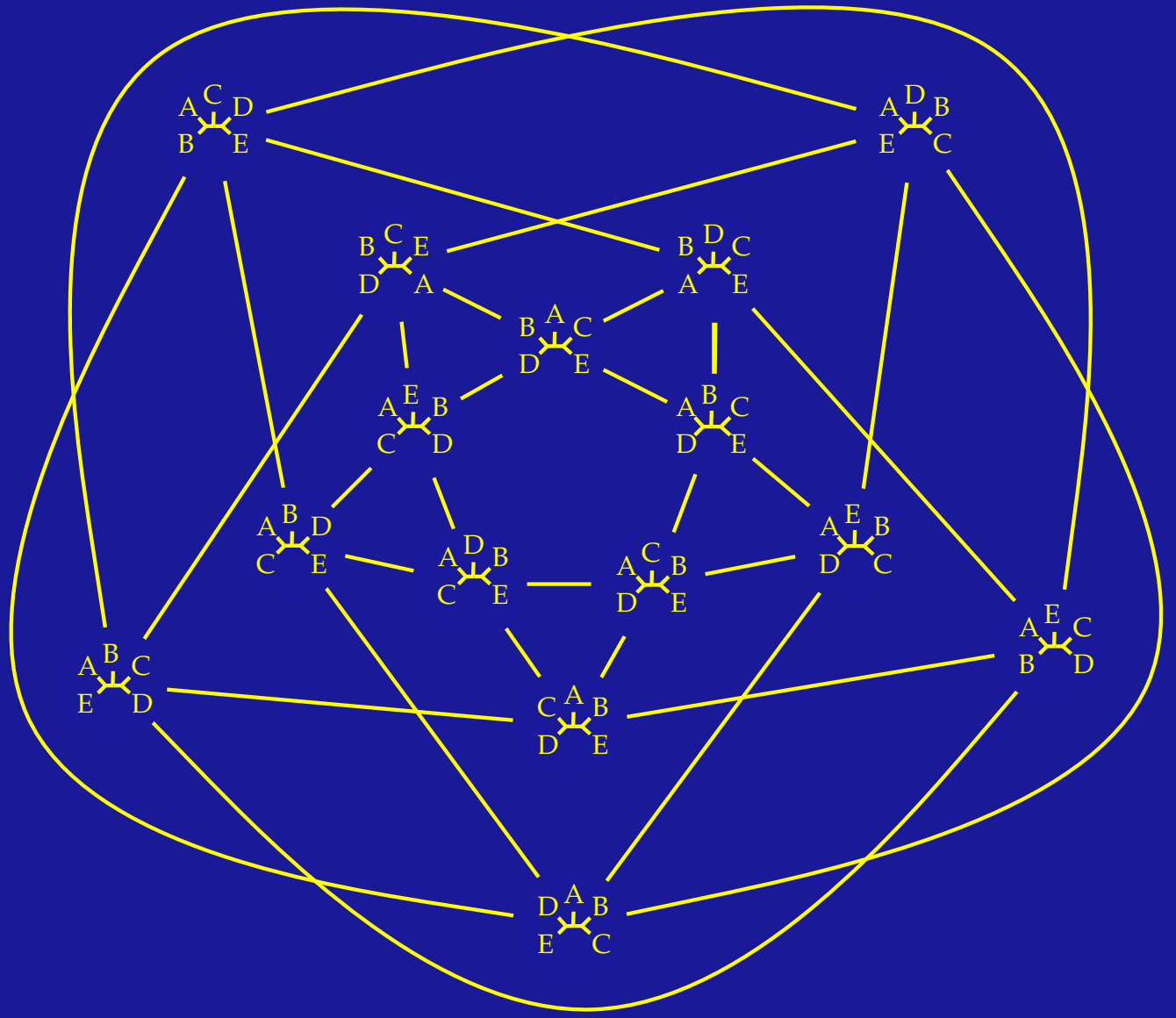
is rearranged by dissolving the connections to an interior brar

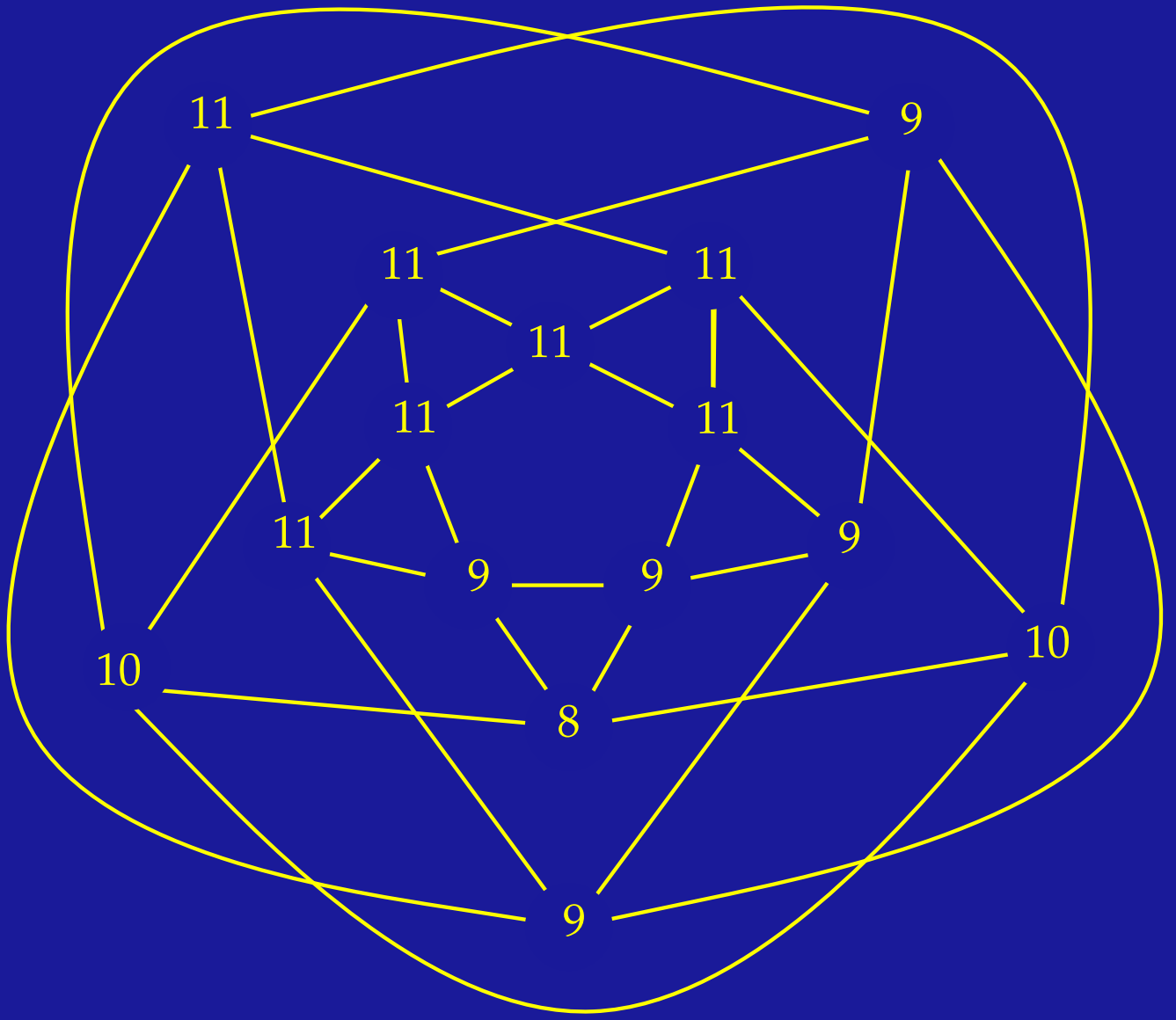


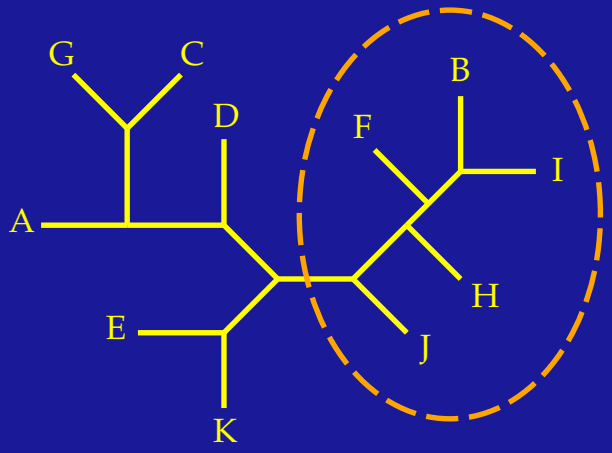
and reforming them in one of the two possible alternative wa



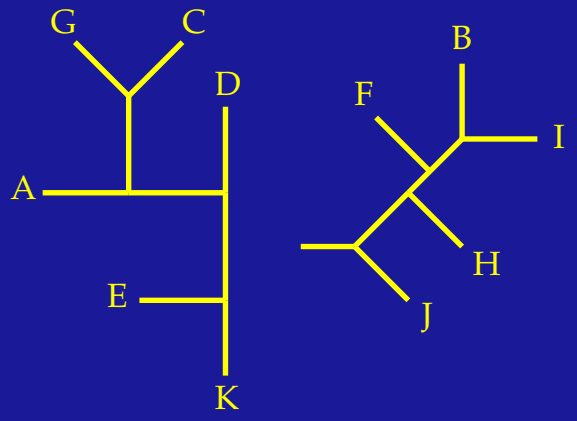



A C D
B E
A D B
E C
B C E
D A
B D C
A E
B A C
D E
A E B
C D
A B C
D E
A B D
C E
A D B
C E
A C B
D E
A E B
D C
A E C
B D
A B C
E D
C A B
D E
D A B
E C

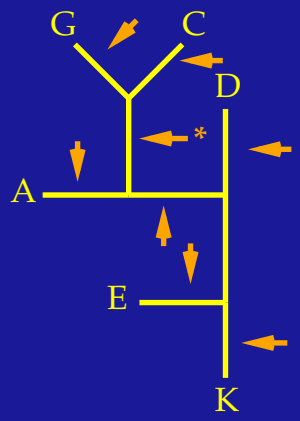

11
9
11
11
11
11
11
9
9
9
10
10
8
9



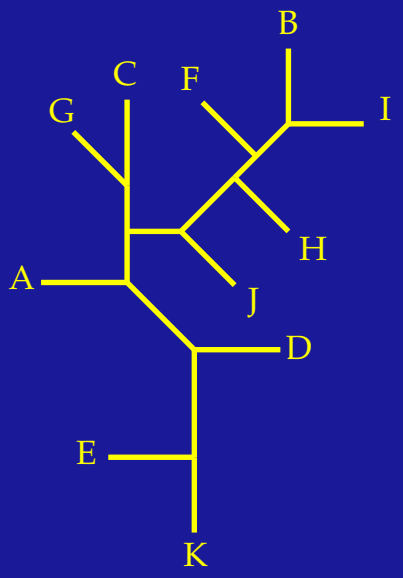
Break a branch, remove a subtree

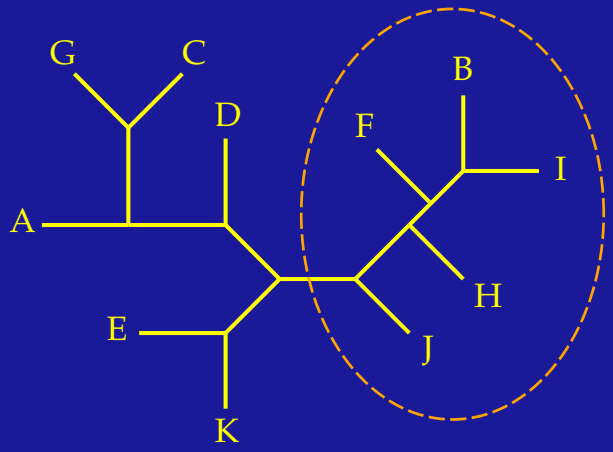


Add it in, attaching it to one (*)
of the other branches

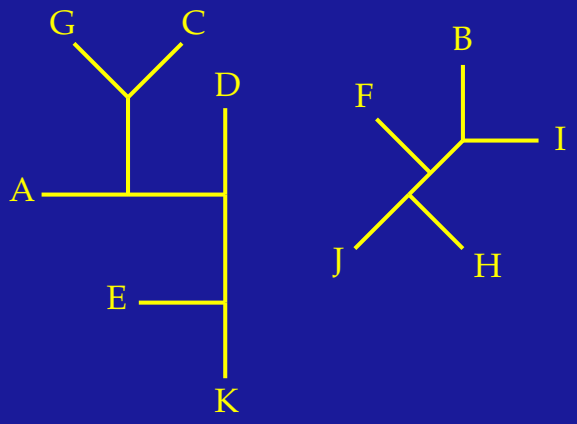


Here is the result:

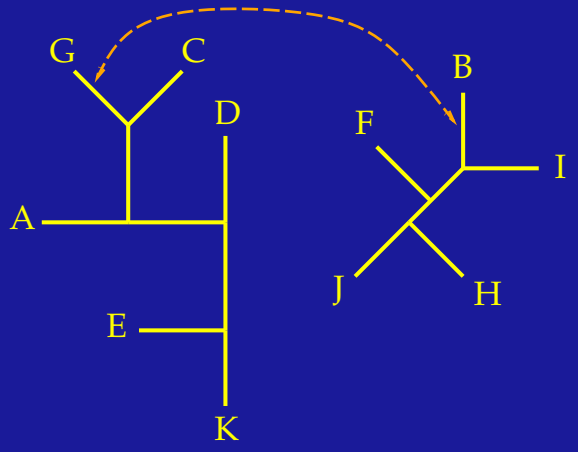




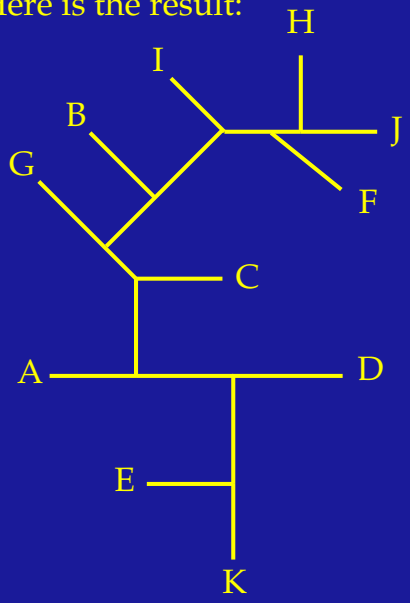
Break a branch, separate the subtree

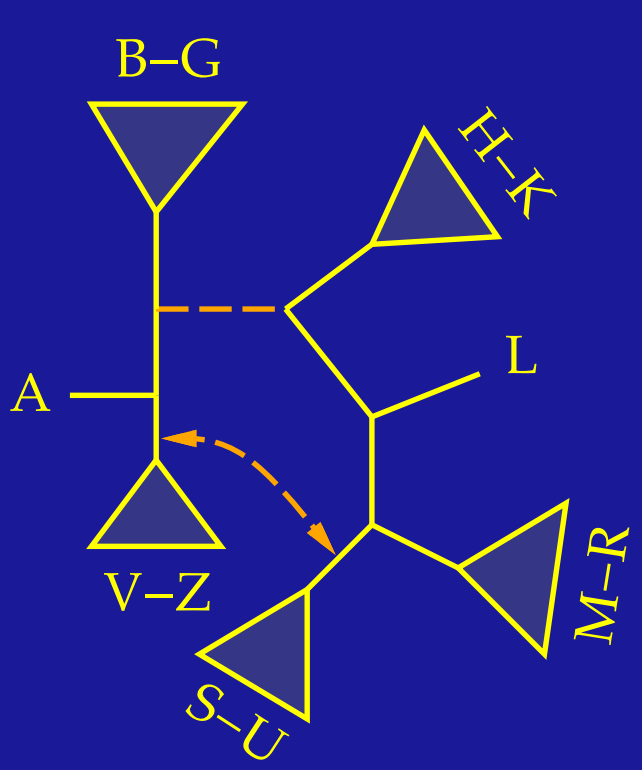


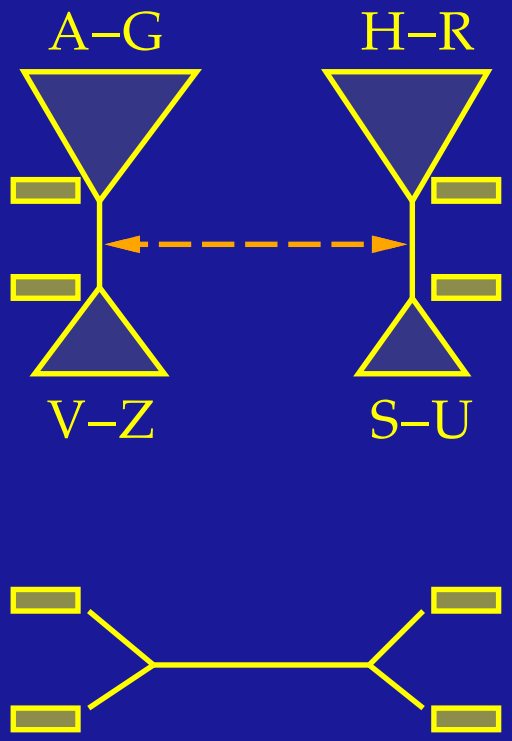
Connect a branch of one
to a branch of the other

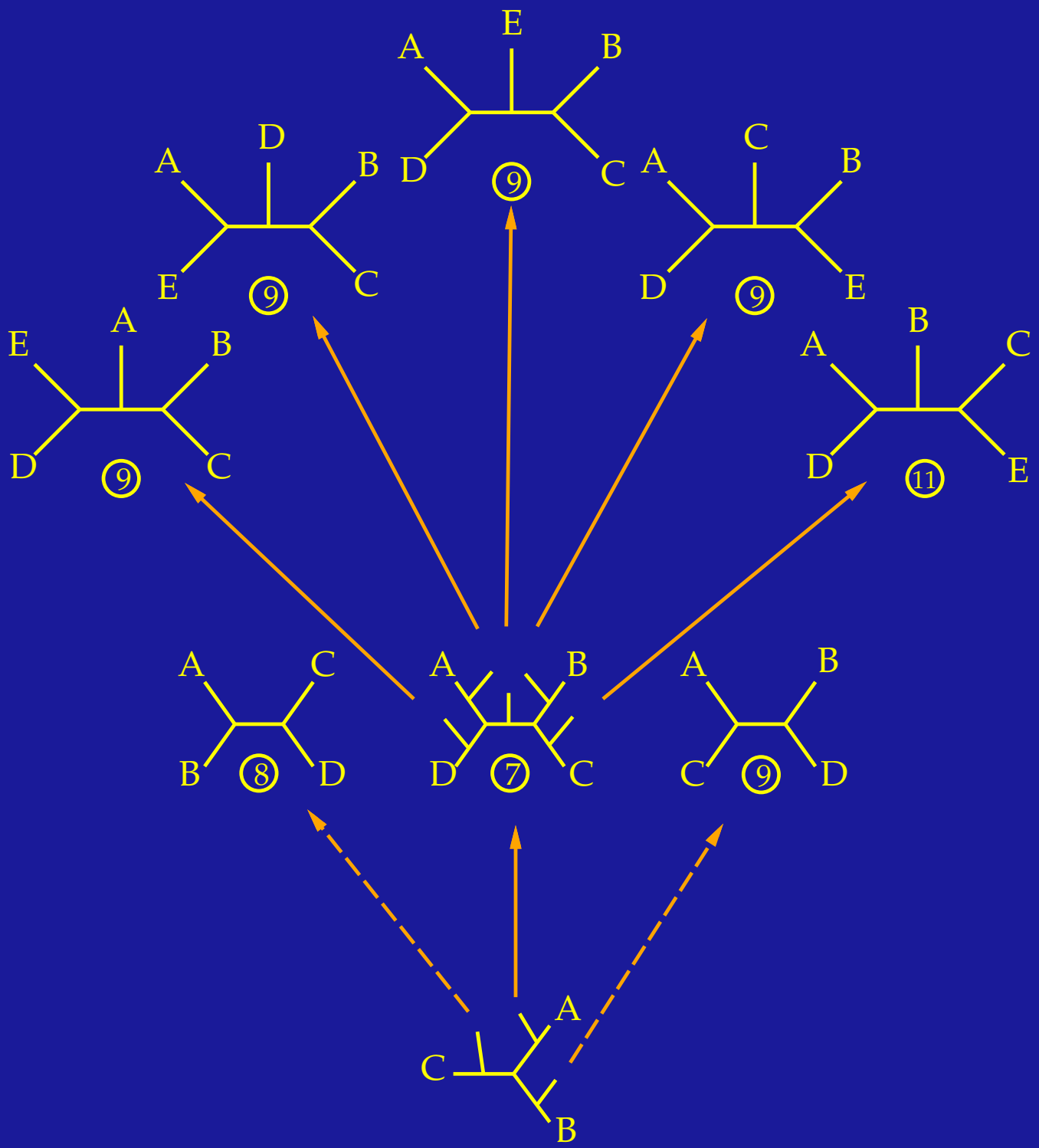


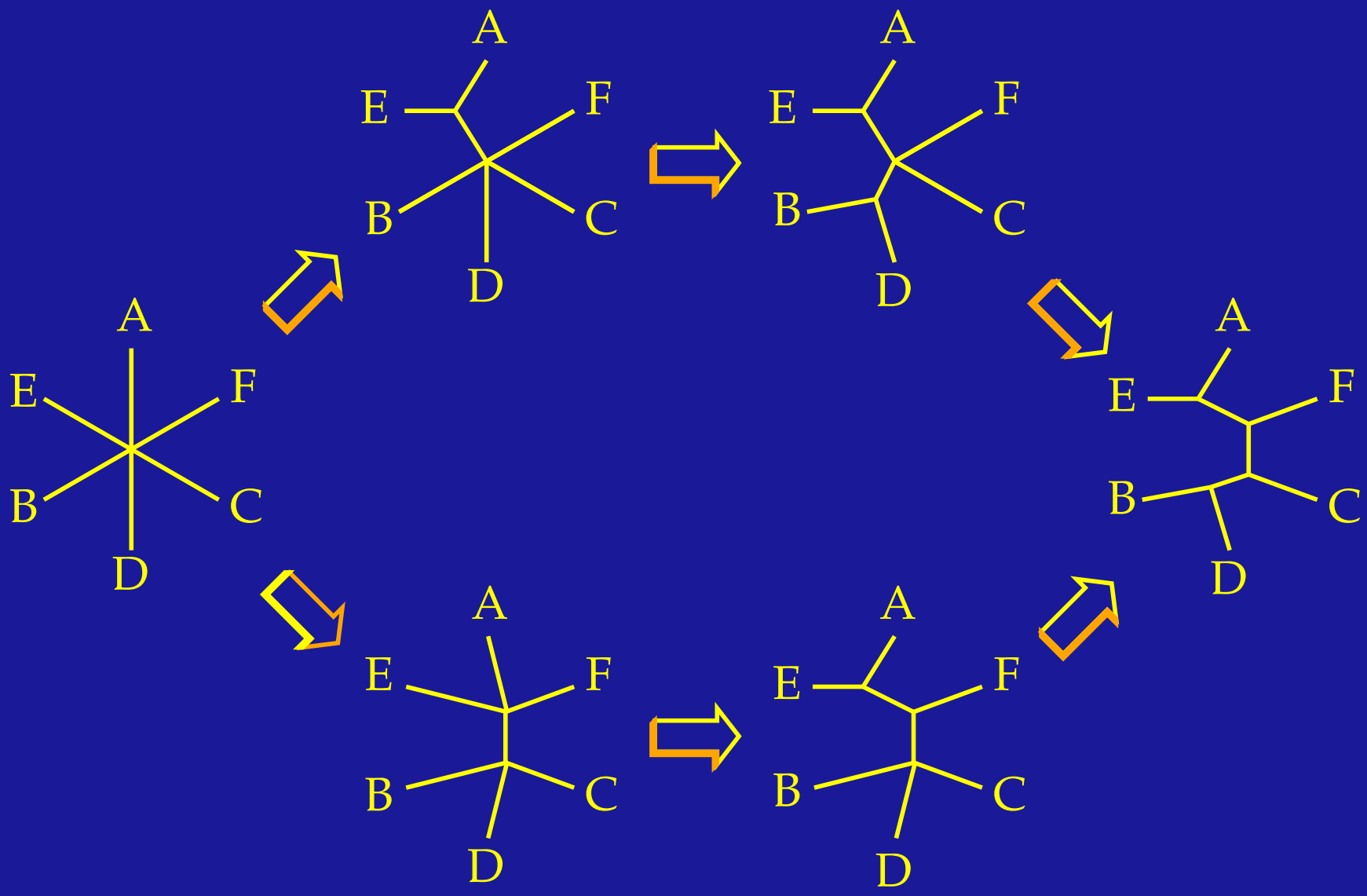
Here is the result:

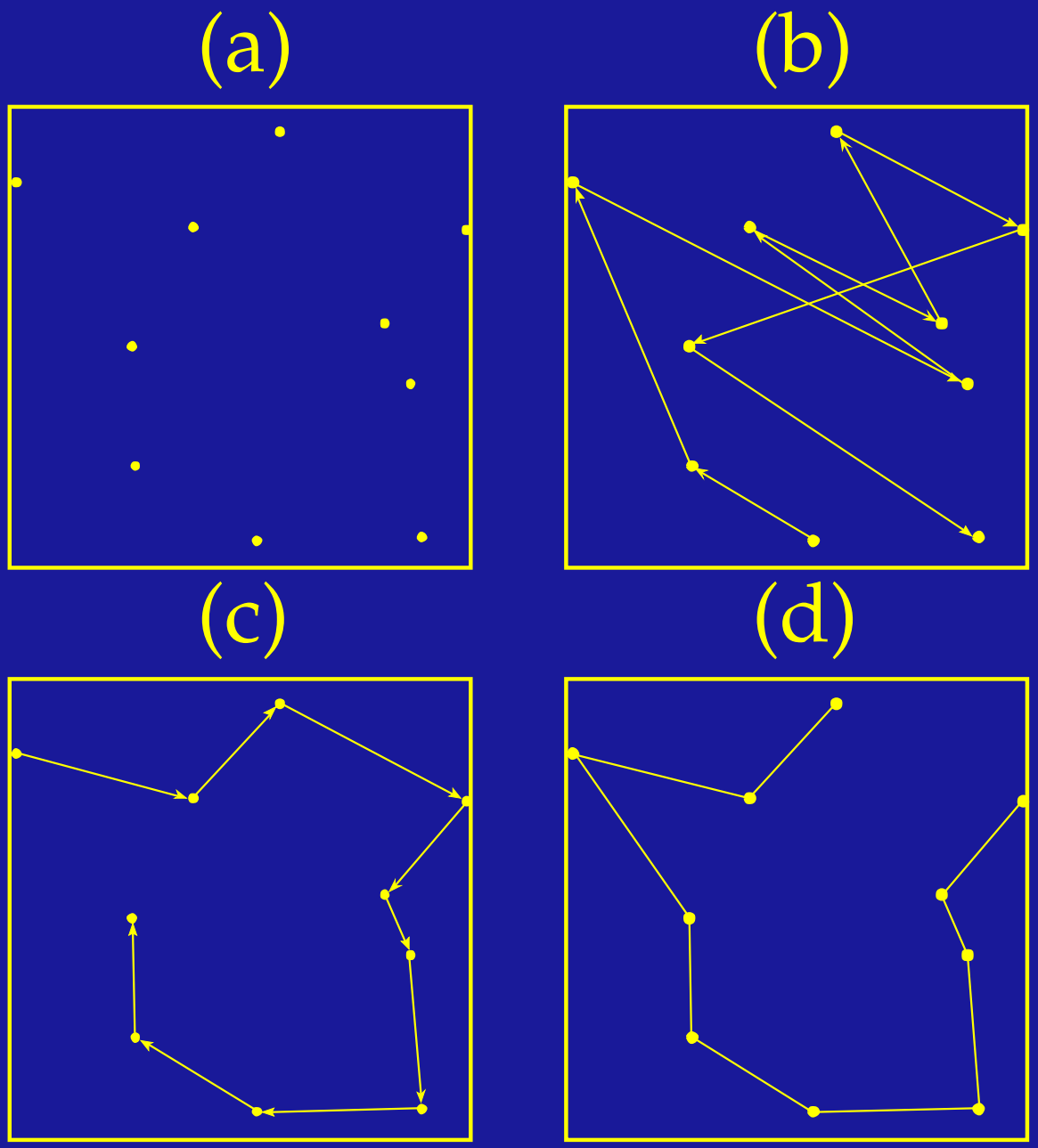


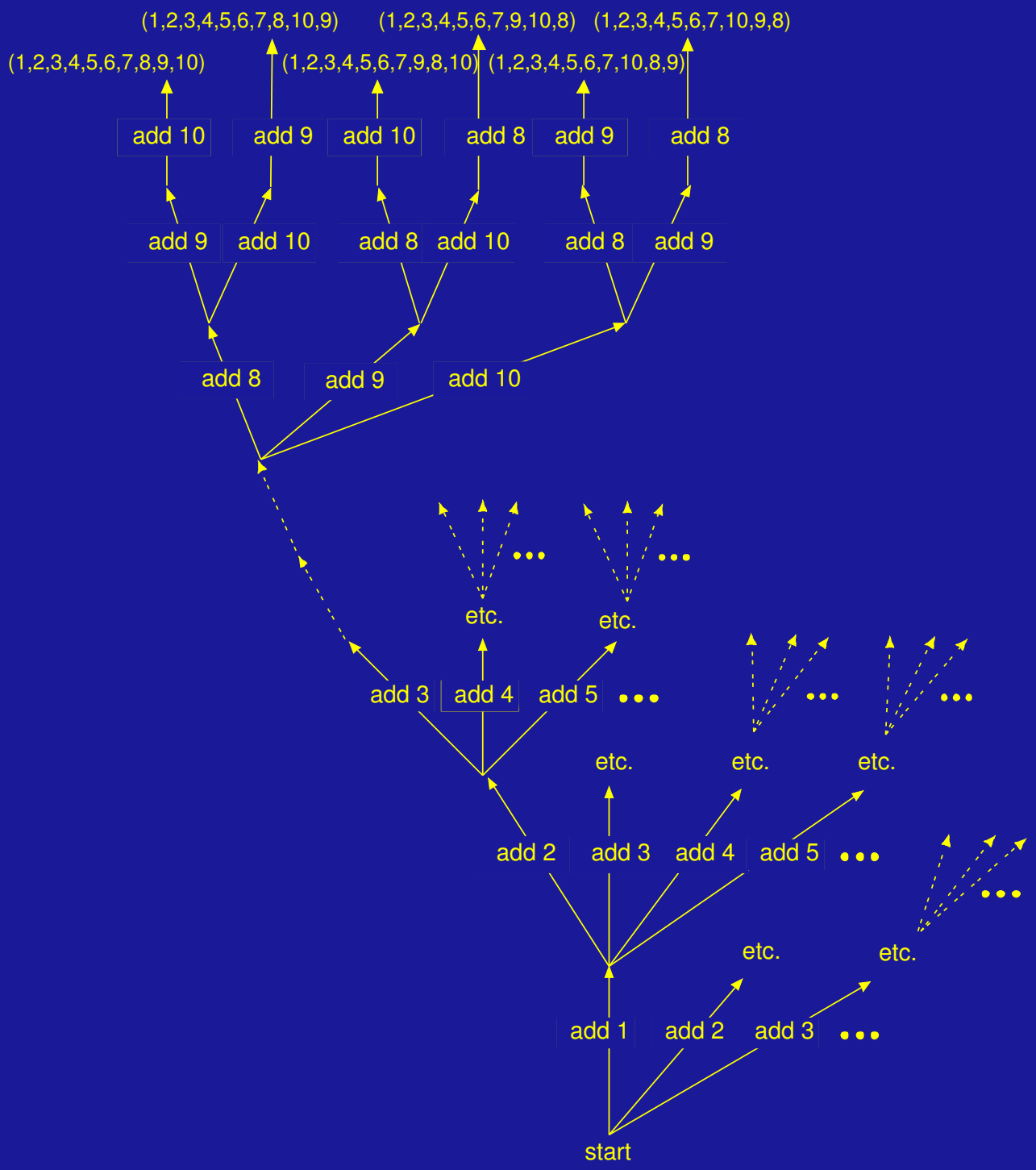

B–G
H–K
L
A
V–Z
M–R
S–U


A–G
H–R
V–Z
S–U


A E B D C 9
A D B E C 9
A C B D E 9
E A B D C 9
A B C D E 11
A C B D 8
A B D C 7
A B C D 9
A C B

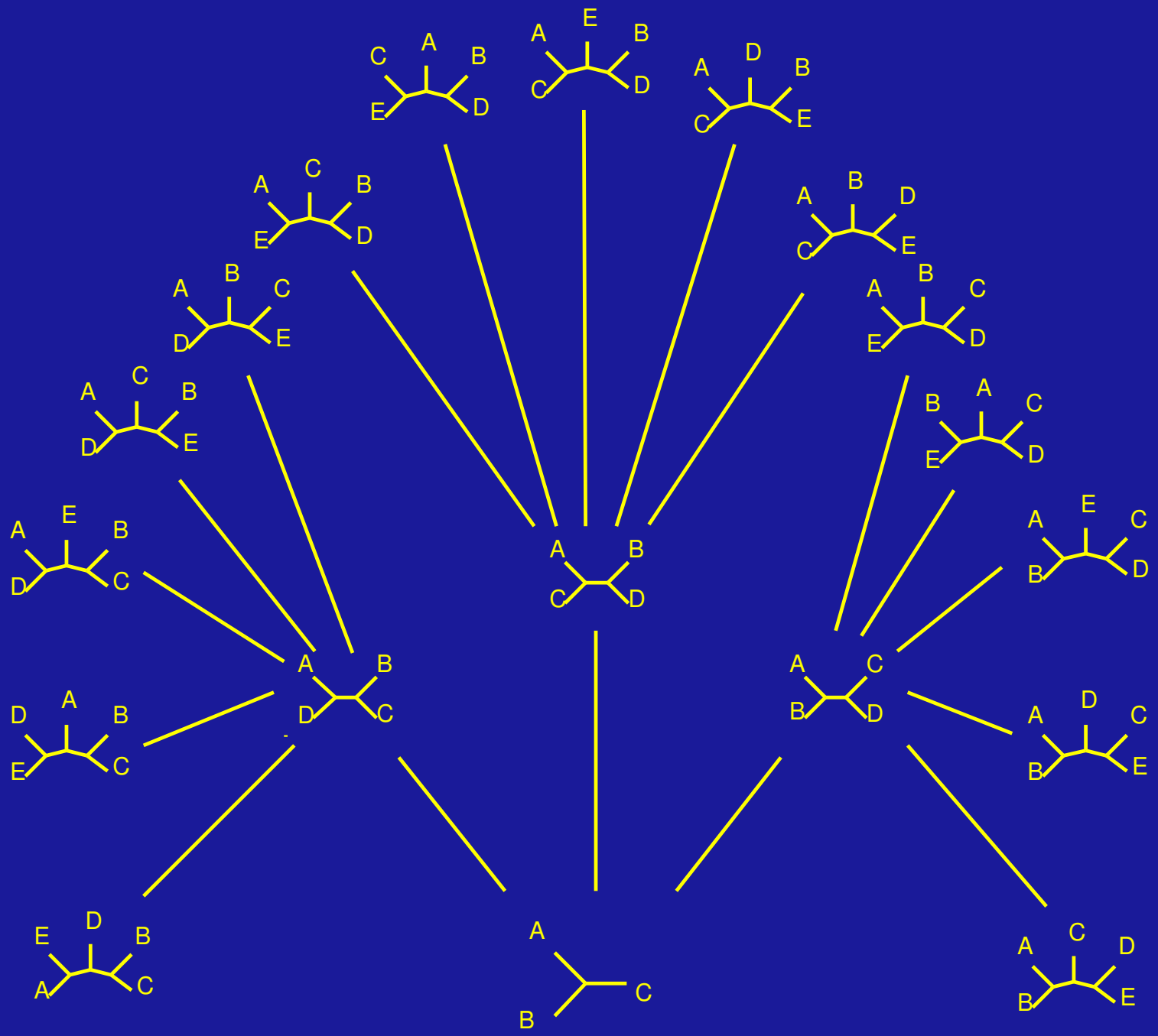

A
E
F
B
C
D

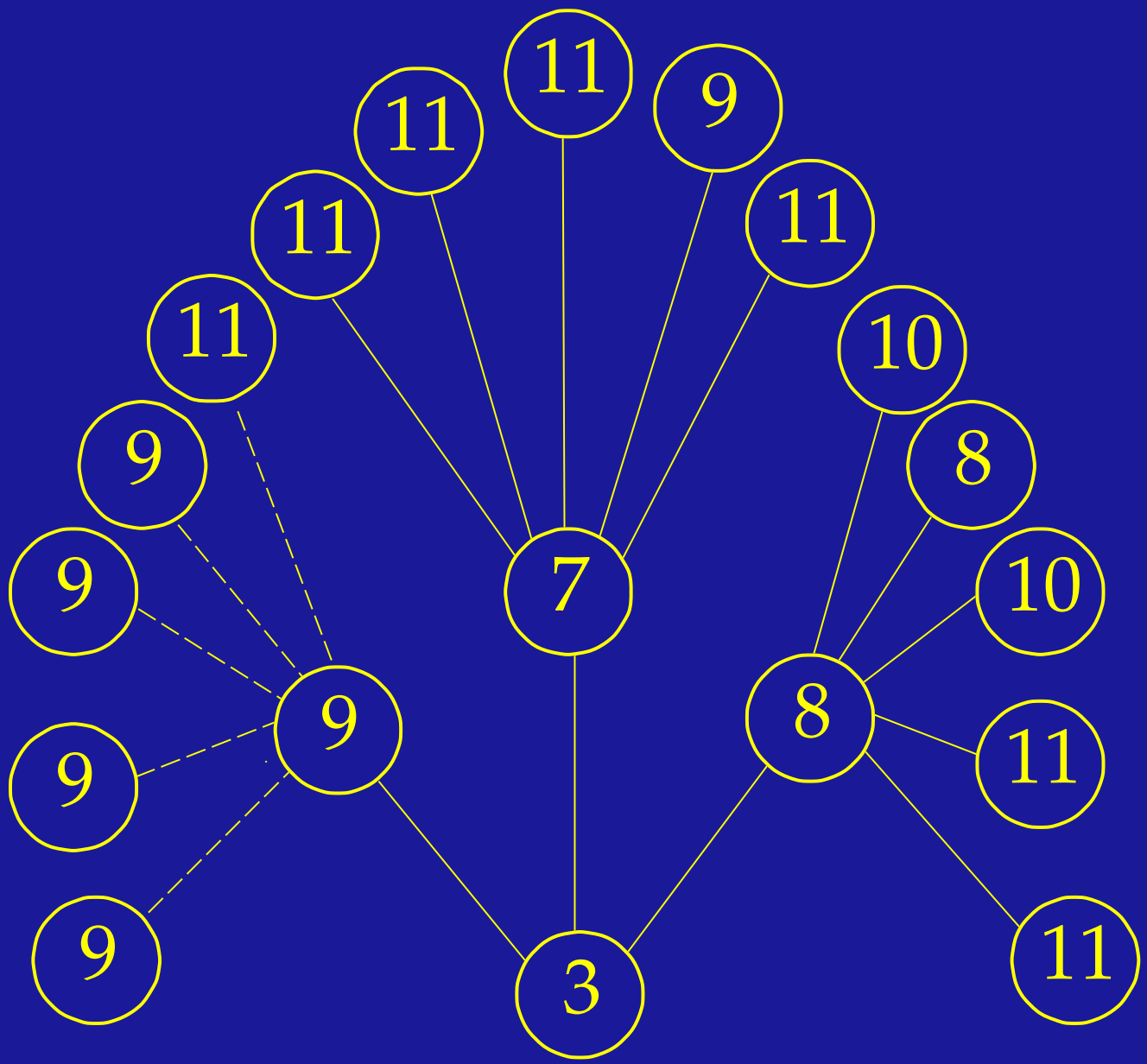

(a)
(b)
(c)
(d)


(1,2,3,4,5,6,7,8,9,10)
(1,2,3,4,5,6,7,8,10,9)
(1,2,3,4,5,6,7,9,8,10)
(1,2,3,4,5,6,7,9,10,8)
(1,2,3,4,5,6,7,10,8,9)
(1,2,3,4,5,6,7,10,9,8)
add 10
add 9
add 10
add 8
add 9
add 8
add 9
add 10
add 8
add 10
add 8
add 9
add 8
add 9
add 10
etc.
etc.
add 3
add 4
add 5
etc.
etc.
etc.
add 2
add 3
add 4
add 5
etc.
etc.
add 1
add 2
add 3
start

Results for this case:

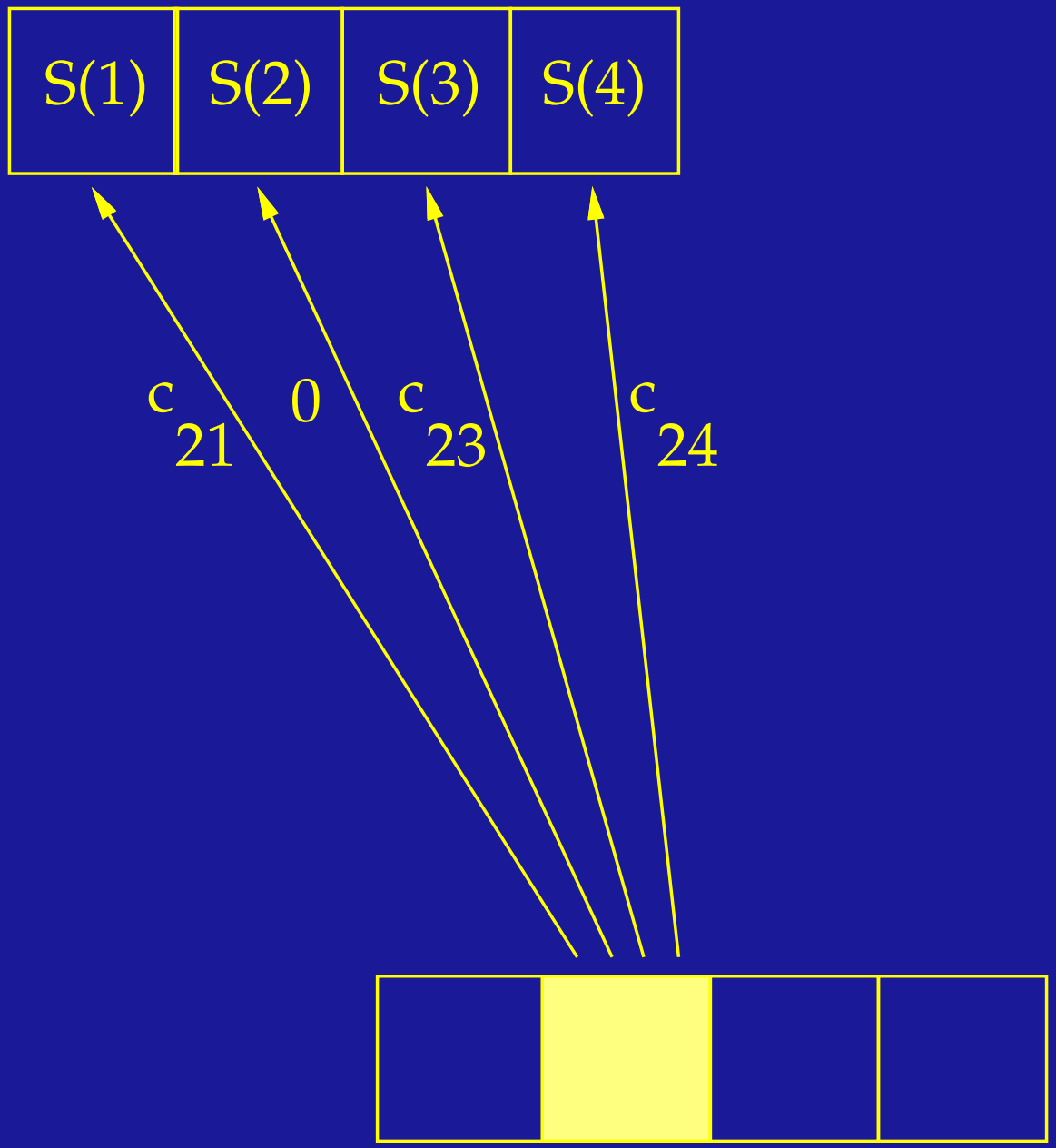
Algorithm	length	time
Greedy search from all points	2.802660	(fast)
Exhaustive enumeration	2.781230	10.85 sec
Branch and bound	2.781230	0.46 sec

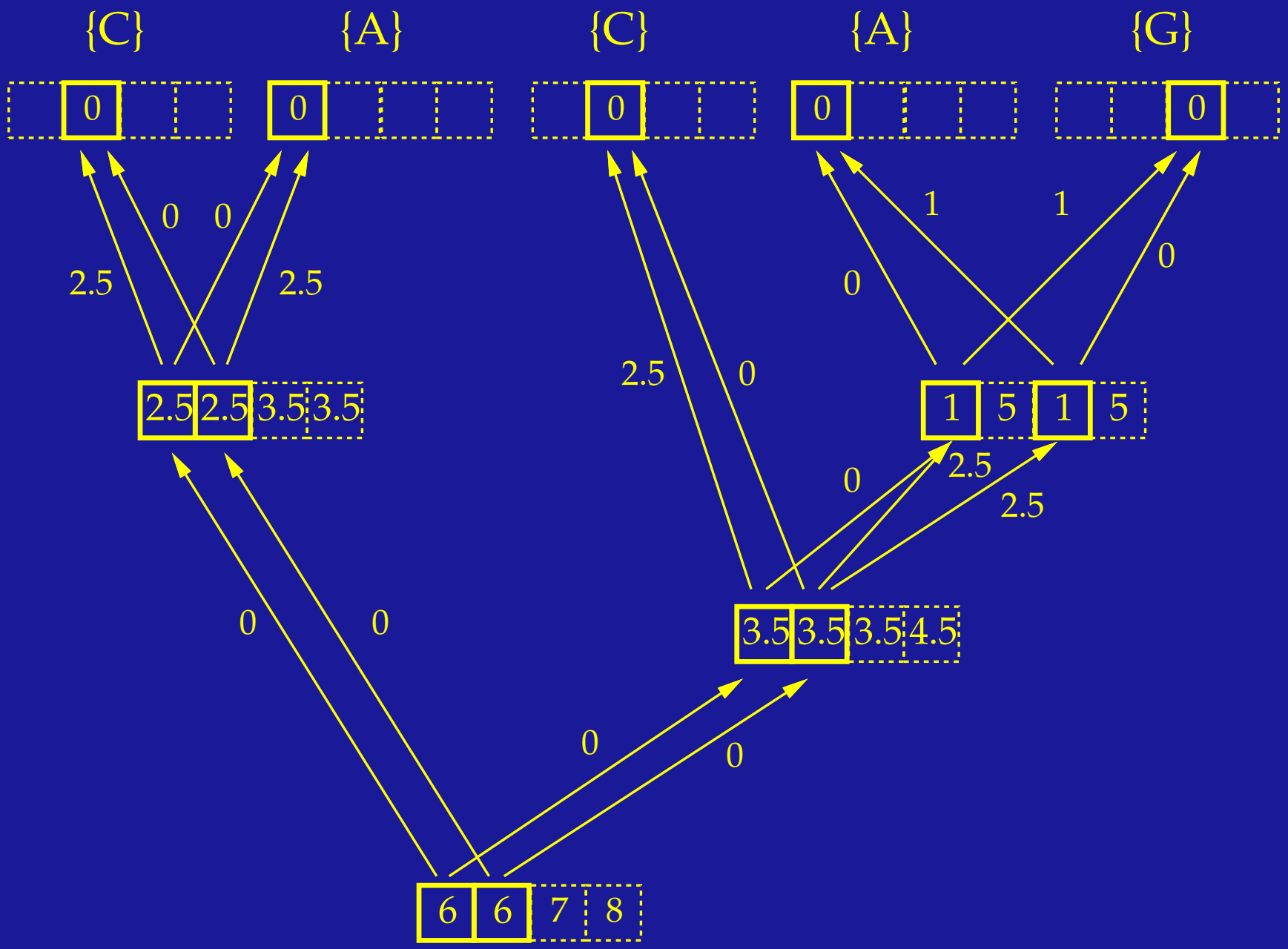

C A B E D
A E B C D
A D B C E
A C B E D
A B D C E
A B C D E
A C B D E
A E B D C
B A C E D
A E C B D
A D C B E
D A B E C
A D B E C
E D B A C
A C D B E

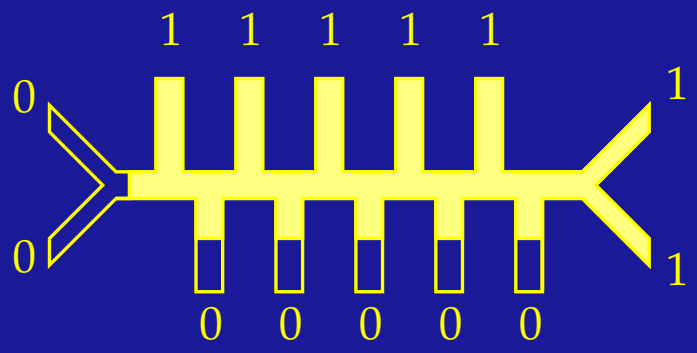

11
11
9
11
11
11
10
9
8
9
7
10
9
8
9
11
9
11

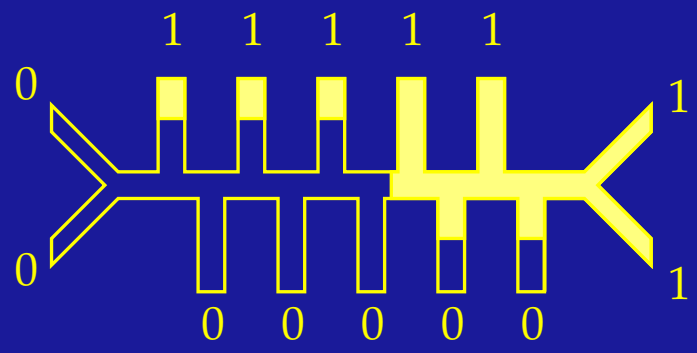
Methods for calculating a lower bound on tree score

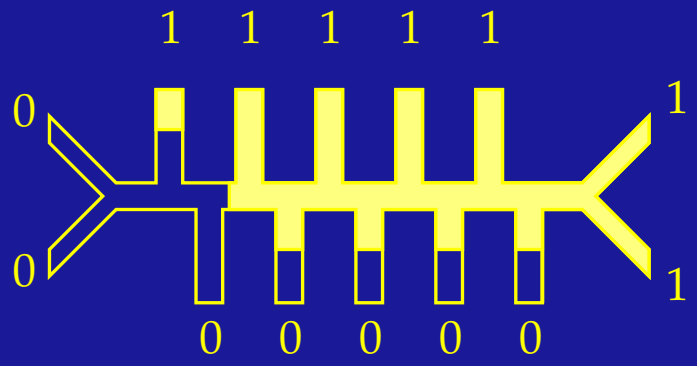
- The score of the partial tree is a lower bound (since adding more species cannot decrease the number of steps)
- Also can add the number of characters that do not show variation on the species added so far, but will once added (actually, the number of new states that will appear once all species are added – if A and G are there already, will C also appear?)
- Can also take all disjoint pairs of characters that will become incompatible once added, but aren't incompatible now (this is due to Dave Swofford. Each brings in one more step.)

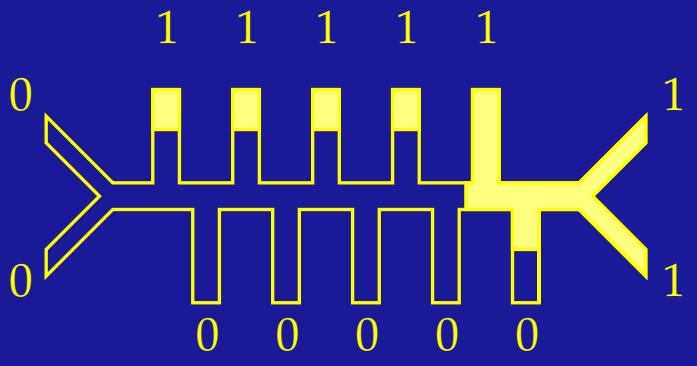

S(1)
S(2)
S(3)
S(4)
c 21
0
c 23
c 24

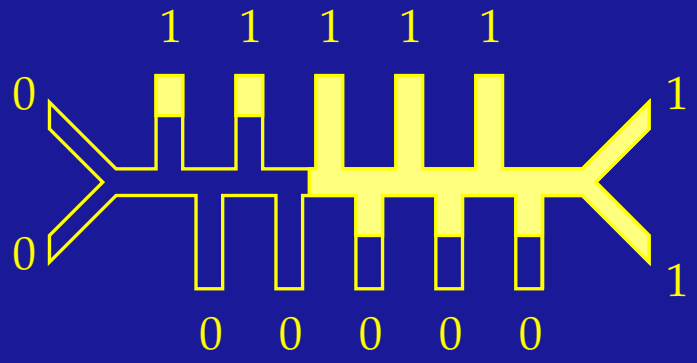

{C}
{A}
{C}
{A}
{G}
0
0
0
0
0
2.5
0
0
2.5
2.5
0
1
1
0
2.5
2.5
3.5
3.5
1
5
1
5
0
2.5
2.5
0
0
3.5
3.5
3.5
4.5
0
0
6
6
7
8

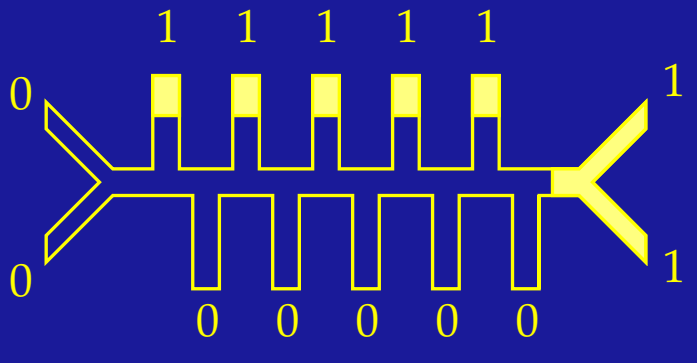

1 1 1 1 1
0
0
1
1
0 0 0 0 0

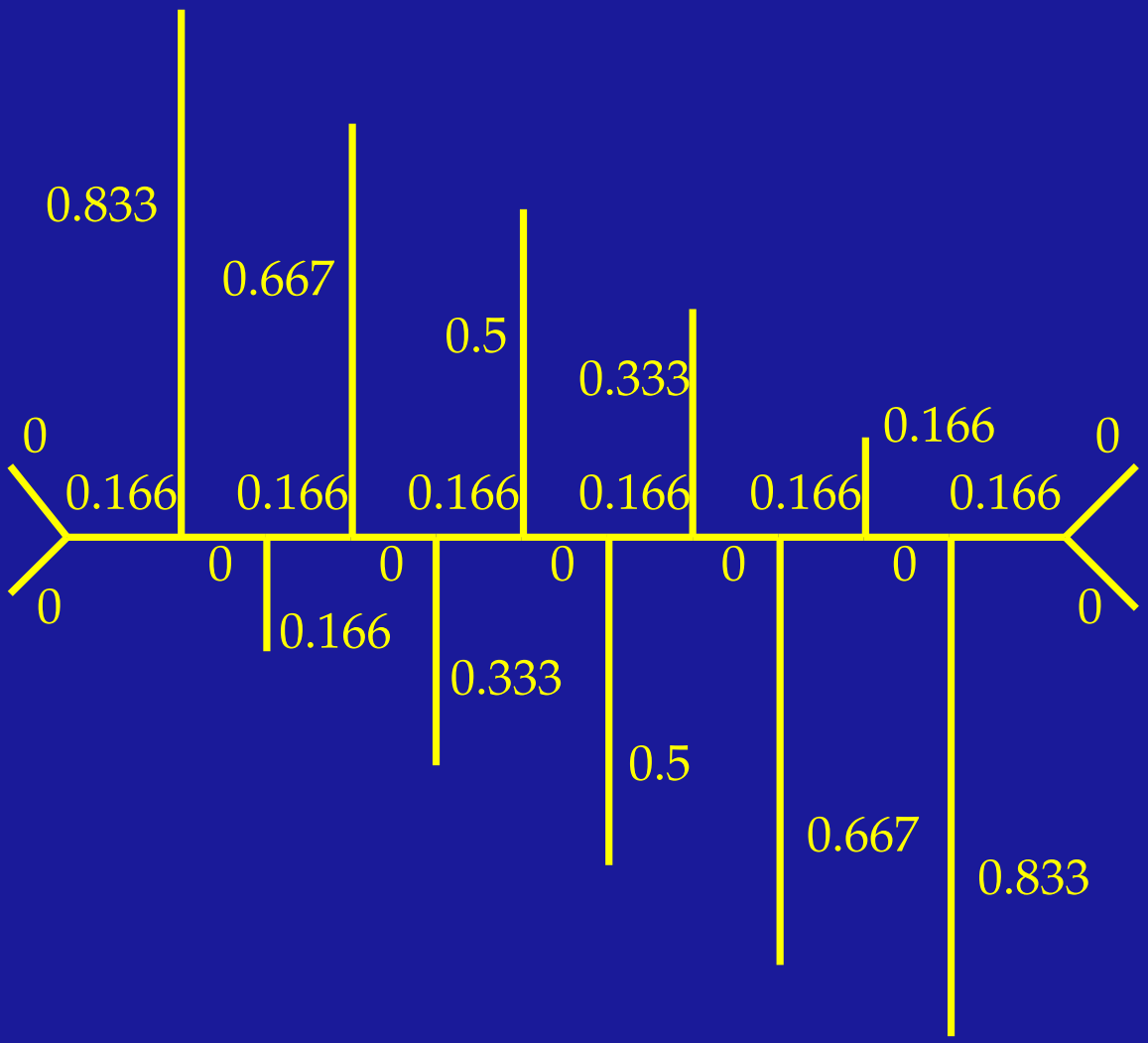

1 1 1 1 1
0
0
1
1
0 0 0 0 0


1 1 1 1 1
0
0
1
1
0 0 0 0 0


1 1 1 1 1
0
0
1
1
0 0 0 0 0


1 1 1 1 1
0
0
1
1
0 0 0 0 0


1 1 1 1 1
0
0
1
1
0 0 0 0 0


0.833
0.667
0.5
0.333
0.166
0
0.166
0.166
0.166
0.166
0.166
0.166
0
0
0
0
0
0
0
0
0.166
0.333
0.5
0.667
0.833

This freeware-friendly presentation prepared with

- Linux (operating system)
- PDFLaTeX (mathematical typesetting and PDF preparation)
- Idraw (drawing program to modify plots and draw figures)
- Adobe Acrobat Reader (to display the PDF in full-screen mode)