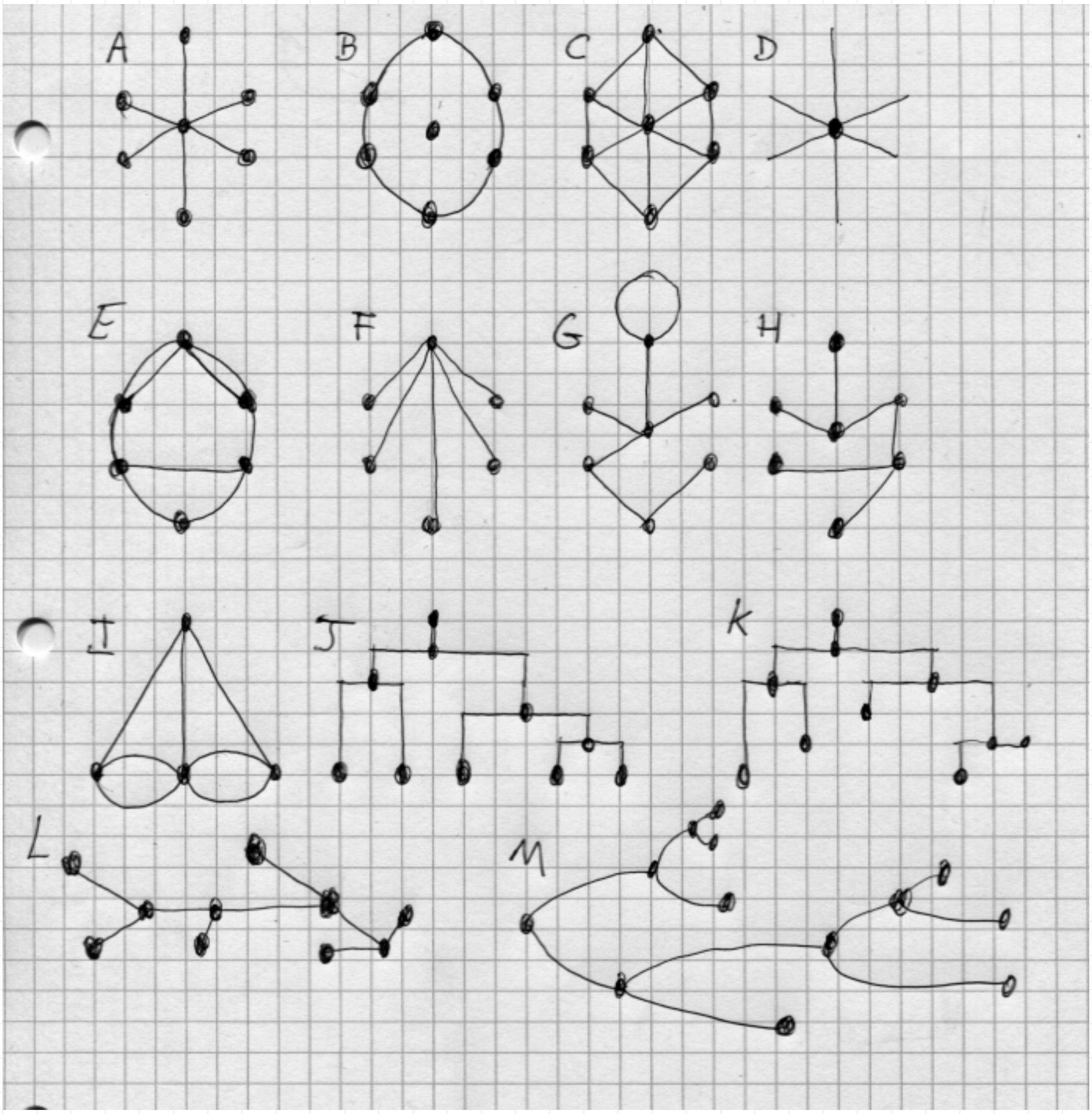
Scheda 1 – Grafi e alberi

1.1) Un **grafo** è una struttura matematica costituita da **punti** (detti anche **nodi**) e **archi** (detti anche **spigoli**), i cui estremi sono punti del grafico.

Nelle seguenti figure, i nodi sono rappresentati da punti grandi, gli archi da linee. Quali delle figure non rappresenta un grafo?



1.2) Si chiama **cappio** (o **loop**) un arco che ha estremi coincidenti in un punto, e si chiamo **archi multipli** archi che uniscono la stessa coppia di punti. Si chiama **grafo semplice** un grafo che non ha cappi né archi multipli. Quale delle figure non rappresenta è un grafo semplice?

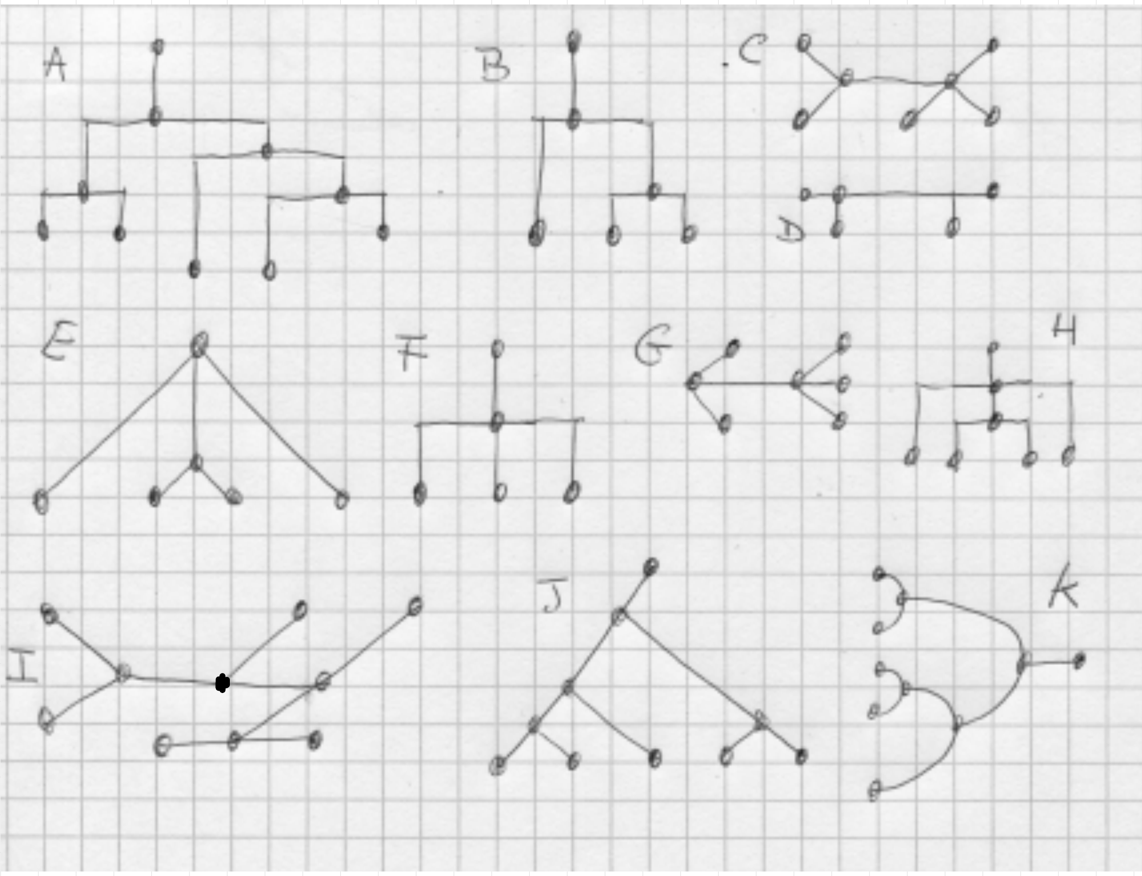
1.3) Una possibile definizione di albero è la seguente: un **albero** è un grafo semplice tale che per ogni coppia di suoi punti esiste uno ed un solo cammino che li unisce.

Usando l'idea che hai di albero, dai la definizione di **cammino** in un grafo:

1.4) Quale delle figure rappresenta un albero?

Scheda 2 – Alberi

2.1) I nodi di un albero che sono estremi di un solo arco si chiamano **foglie**; i nodi che non sono foglie sono **nodi interni**. Nota che ogni foglia ha uno ed un solo **genitore**, che è l'unico nodo a cui è attaccata la foglia. Quanti nodi, foglie, archi hanno gli alberi B, F, G?



albero	B	F	G
n. nodi			
n. foglie			
n. archi			

2.2) Nella definizione di albero, la lunghezza dei rami non è rilevante. Quali degli alberi in figura sono uguali tra loro?

2.3) Che relazione c'è tra il numero di nodi, di foglie, di archi? Come puoi dimostrare in generale la tua affermazione?

2.4) Un **albero con radice** è un albero con una foglia che viene denominata radice. Fissare una radice permette di definire la relazione di ascendenza (e quella di discendenza). Dimostralo, dando una corretta definizione dell'affermazione l'affermazione "il nodo A è un antenato (un ascendente) del nodo B" (puoi pensare come esempio all'albero J, in cui scegli come radice il nodo più in alto).

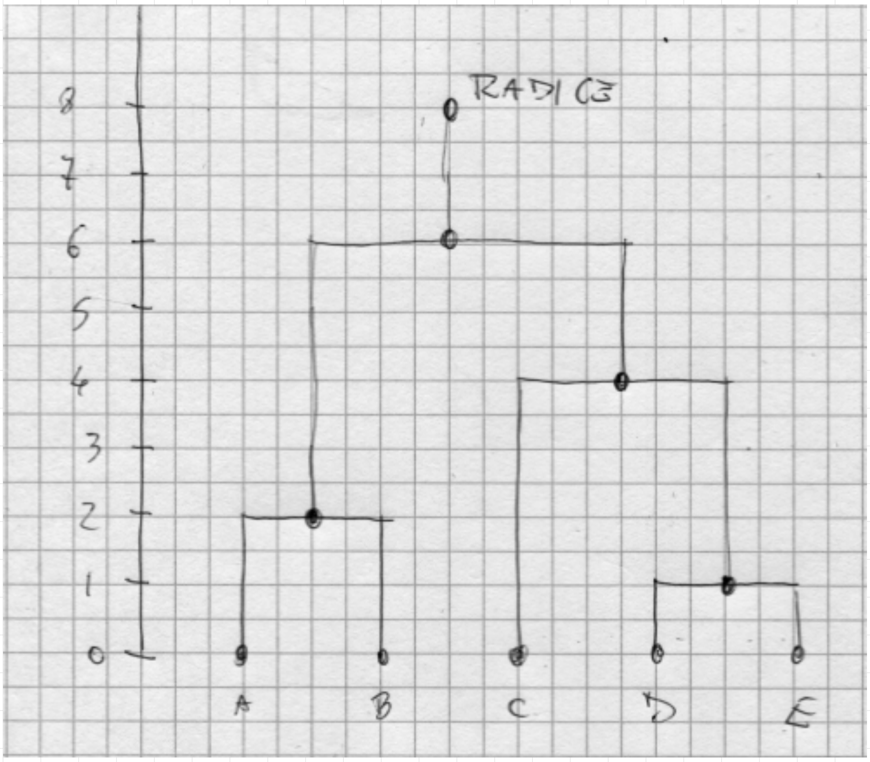
2.5) Si chiama **albero binario** un albero i cui nodi sono estremi di al più tre archi. Perché si chiama binario? (pensa al caso di un albero con radice).
Quale degli alberi precedenti non è un albero binario?

Scheda 3 – Alberi ultrametrici

Se si attribuisce una lunghezza a ogni arco di un albero, si può definire la distanza tra due nodi qualunque come la lunghezza del cammino che li unisce. Nella ricostruzione dell'evoluzione delle specie e dell'evoluzione delle lingue la distanza tra i nodi è espressa dal tempo. Il problema centrale di queste ricostruzioni è determinare l'albero, cioè trovare la posizione dei nodi e la lunghezza degli archi, conoscendo la sola distanza tra foglie.

C'è una classe di fenomeni per cui questa ricostruzione è particolarmente semplice, ed è quella rappresentabile mediante alberi detti **ultrametrici**, di cui ora ci occuperemo.

3.1) Considera l'albero in figura, e supponi che la distanza tra un nodo e il suo genitore sia data dalla differenza delle ordinate (le differenza in ascissa non conta).

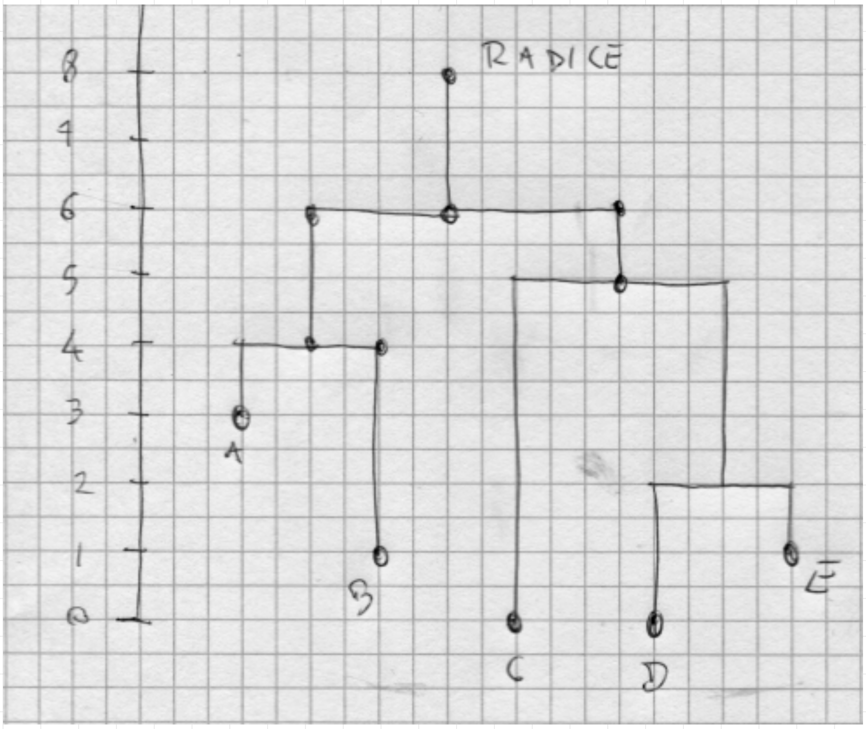


Ricava le tabella (o matrice) delle distanze tra le foglie, esclusa la radice.

	A	B	C	D	E
A	0	4			
B	.	0			
C	.	.	0		
D	.	.	.	0	
E	.	.	.	.	0

(non riempire la parte inferiore della tabella perché si ottiene dalla parte superiore... come?)

3.2) Come nel caso precedente, ma considera quest'altro albero.



Ricava le tabella (o matrice) delle distanze tra e foglie, esclusa la radice.

	A	B	C	D	E
A	0	4			
B	.	0			
C	.	.	0		
D	.	.	.	0	
E	.	.	.	.	0

Un albero binario è **ultrametrico** se tutti i rami che uniscono la radice alle foglie hanno la stessa lunghezza. Per questa proprietà, gli alberi ultrametrici sono particolarmente adatti a descrivere fenomeni in cui, dopo la biforcazione, la distanza tra i rami cresce uniformemente nel tempo. L'albero del punto 3.1 è ultrametrico, quello del punto 3.2 no.

3.3) Considera la seguente tabella delle distanze tra le foglie di un albero ultrametrico e disegna il corrispondente albero. Dov'è la radice?

	A	B	C	D
A	0	2	8	6
B	2	0	8	6
C	8	8	0	8
D	6	6	8	0

3.4) Esiste un albero con tre foglie tale che $|AB| = |AC| = 4$ e $|BC| = 2$?
Esiste un albero con tre foglie tale che $|AB| = |AC| = 2$ e $|BC| = 4$?
Esiste una proprietà per le distanze di tre foglie qualunque che caratterizza gli alberi ultrametrici. Qual è?

3.5) Scrivi un algoritmo per ricostruire l'albero a partire da una matrice delle distanze che soddisfa la proprietà di ultrametricità. Probabilmente quello che enuncerai sarà simile all'algoritmo UPGMA (Unweighted Pair Group Method with Arithmetic Mean), scoperto da Sokal e Michener nel 1958.

Scheda 4 – Un albero delle lingue

4.1) Supponi di avere una matrice delle distanze che non soddisfa esattamente la proprietà di ultrametricità (per inesattezza del modello, o per inesattezza delle misure). Che fai?

4.2) Ho preso varie traduzioni di una fiaba di *Pollicino*, ho calcolato le frequenze relative delle singole lettere dell'alfabeto, e ho calcolato come distanza tra due lingue la somma delle differenze delle percentuali delle singole lettere nelle varianti della fiaba.
Questa è la tabella che ho ottenuto:

	ita	fra	spa	por	eng	ger	nld
ita	0	31	26	30	47	51	59
fra	31	0	30	35	45	40	50
spa	26	30	0	20	45	47	57
por	30	35	20	0	48	52	59
eng	47	45	45	48	0	37	43
ger	51	40	47	37	37	0	32
nld	59	50	47	43	43	32	0

Prova a costruire il corrispondente albero delle lingue, anche se la matrice delle distanze non è ultrametrica.

Scheda 5 – Approfondimenti: alberi e distanze

5.1) Disegna un triangolo di lati lunghi 3, 4, 5.

Perché non esiste un triangolo di lati 3, 4, 8?

5.2) Un albero binario ha tre foglie A, B, C; supponi che le distanze tra loro siano

$$|AB| = 3, \quad |AC| = 4, \quad |CB| = 5.$$

Disegna l'albero (nota che non è ultrametrico, dunque non mettere la radice).

Perché non esiste un albero tale che

$$|AB| = 3, \quad |AC| = 4, \quad |CB| = 8?$$

5.3) Più in generale, se $|AB| = c$, $|AC| = b$, e $|BC| = a$, determina la posizione del nodo centrale, e discuti la sua esistenza al variare di a, b, c.

5.4) Ricostruisci l'albero binario che ha quattro foglie A, B, C, D, con le seguenti distanze tra loro

	A	B	C	D
A	0	3	5	5
B	3	0	4	6
C	5	4	0	8
D	6	6	8	0

5.5) Considera l'albero che hai ricostruito e la corrispondente matrice di distanze, e le seguenti possibili modifiche delle distanze tra le foglie. Esiste ancora l'albero?

- se $|AD| = 7$ l'albero esiste / non esiste
- se $|AD| = 3$ l'albero esiste / non esiste
- se $|BC| = 2$ l'albero esiste / non esiste
- se $|BD| = 6$ l'albero esiste / non esiste
- se $|AB| = 4$ l'albero esiste / non esiste

5.6) Trova qual è la proprietà delle distanze tra 4 punti che assicura che appartengano ad un albero.

5.7) Dai un algoritmo per determinare l'albero. Prova a farti un'idea di quante operazioni (moltiplicazioni) tra numeri sono necessarie per ricostruire un albero con n foglie. Probabilmente quello che troverai sarà un algoritmo funzinante, ma decisamente lento. L'algoritmo più usato è il **neighbor joining**, scoperto da Saitou e Nei nel 1987.