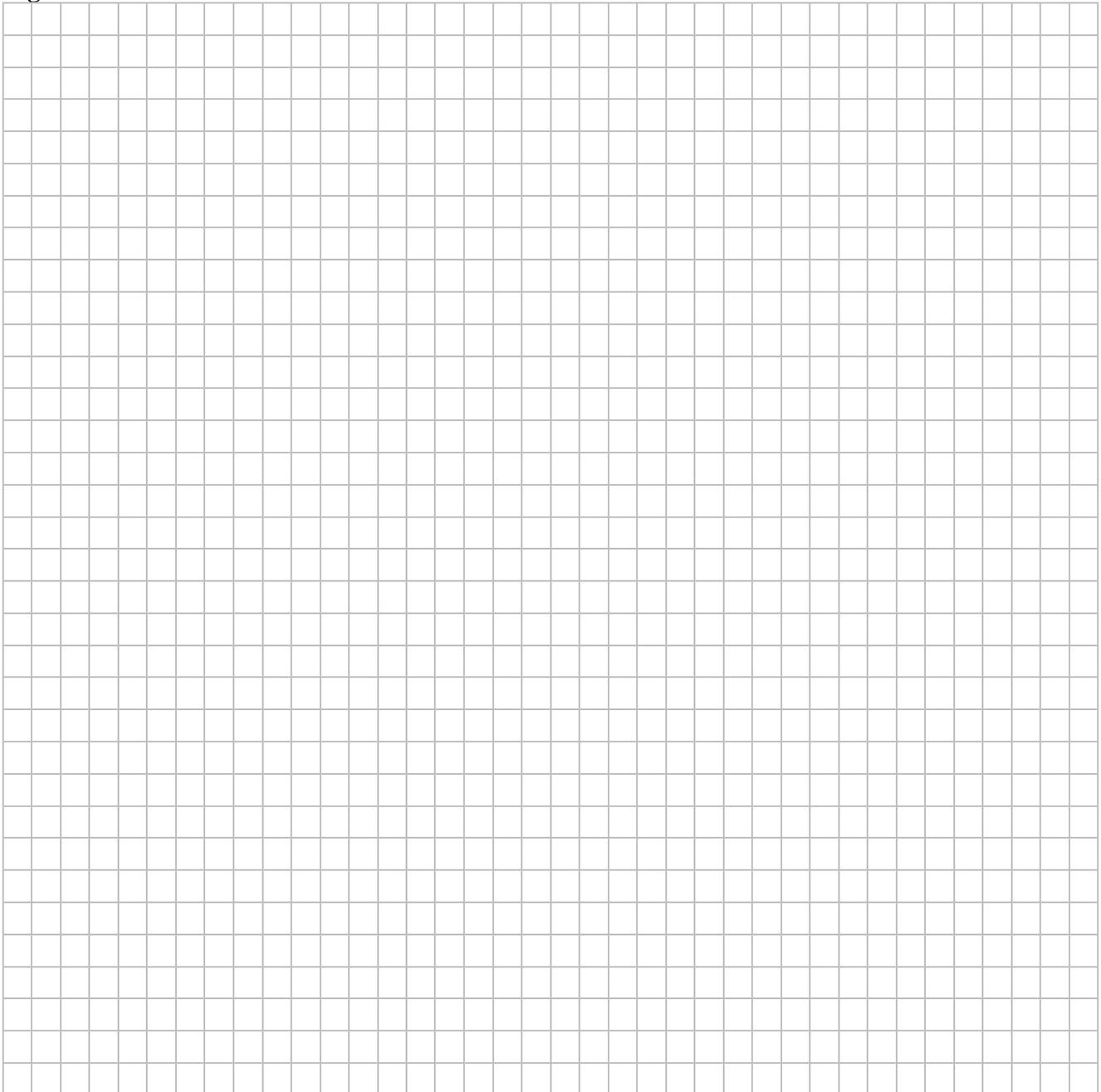


Esercizio 4

Costruire 2 alberi B+tree per indicizzare i seguenti dati: 23, 44, 56, 32, 11, 5, 67, 24, 55, 78, 99, 98, 76, 31, 53, 72, 3.

I 2 B+tree devono essere di ordine $g=2$, con foglie di capacità 4 e fattore di caricamento pari a 0.5, il primo albero deve avere altezza minima, il secondo altezza massima.

Svolgimento



FORMULE DI RIEPILOGO PER IL CALCOLO DEI COSTI

- 1 **Dimensione di una relazione:** $NP = \lceil NR \times \text{len}(t) / (D \times u) \rceil$
- 1 **Costo di ordinamento (Sort Merg Z vie):** $2 \times NP \times \lceil \log_2 NP \rceil$
- 1 **Scansione sequenziale tabelle**
- in media a $(NP+1)/2$ blocchi
 - nel caso peggiore a NP blocchi
 - in caso di non esistenza del record si accede a NP blocchi
- 1 **Scansione binaria tabelle**
- in media a $\lfloor \log_2 NP \rfloor$ blocchi
 - nel caso peggiore a $\lfloor \log_2 NP \rfloor + 1$ blocchi
 - In caso di non esistenza del record si accede a $\lfloor \log_2 NP \rfloor + 1$ blocchi
- 1 **Accesso a tabelle con indice**
- **Numero di foglie del B+-Tree:** $NL = \lceil (NK \cdot \text{len}(k) + NR \cdot \text{len}(p)) / (D \cdot u) \rceil$
 - **altezza del B+-Tree:** $\lceil \log_{2g+1} (NN + 1) \rceil \leq h \leq \left\lceil 1 + \log_{g+1} \left(\frac{NN + 1}{2} \right) \right\rceil$, $NN = \min(NL, NK)$
 -
 - **ordine del B+-Tree:** $g = \left\lfloor \frac{D - \text{len}(p)}{2(\text{len}(k) + \text{len}(p))} \right\rfloor$
 -
 - **Indice Clustered:** $h - 1 + \lceil EK / NK \cdot NL \rceil + \lceil EK / NK \cdot NP \rceil$
 - **Indice Unclustered:** $h - 1 + \lceil EK / NK \cdot NL \rceil + EK \cdot \Phi(NR / NK, NP)$
 - **Formula di Cardenas** $\Phi(ER, NP) = NP \times \left(1 - (1 - 1/NP)^{ER} \right) \leq \min\{ER, NP\}$
- 1 **Proiezione**
- **Mantenendo la chiave della relazione:** NP
 - **Perdendo la chiave della relazione:** assimilabile al costo di ordinamento
- 1 **Raggruppamento: assimilabile al costo di ordinamento**
- 1 **Nested loop:**
- **senza predicato di selezione** $NP_R + NR_R \times NP_S$
 - **con predicato di selezione** $C_a(R) + (\text{sel}(\text{pred}) \times NR_R) \times C_a(S)$
- 1 **Sort merge**
- **Tabelle già ordinate** $NP_R + NP_S$
 - **Tabelle da ordinare** $\text{Sort}(R) + \text{Sort}(S) + NP_R + NP_S$
- 1 **Fattori di selettività**
- **Predicato "="** $f(A=v) = 1/NK_A$
 - **Predicato "IN"** $f(A \in \text{set}) = \text{card}(\text{set})/NK_A$
 - **Predicato "<"** $f(A < v) = \frac{v - \min(A)}{\max(A) - \min(A)} \times \frac{NK_A - 1}{NK_A}$
 - **per attributi con molti valori:** $f(A < v) = \frac{v - \min(A)}{\max(A) - \min(A)}$
 - **Predicato "between"** $f(A \in [v1, v2]) = \frac{v2 - v1}{\max(A) - \min(A)} \times \frac{NK_A - 1}{NK_A} + \frac{1}{NK_A}$
 - **per attributi con molti valori:** $f(A \in [v1, v2]) = \frac{v2 - v1}{\max(A) - \min(A)}$
 - **f(p1 AND p2)** $= f_{p1} \times f_{p2}$
 - **f(p1 OR p2)** $= f_{p1} + f_{p2} - f_{p1} \times f_{p2}$