List Structured Files

© 2001 - 2010 by David S. Burris, Ph.D., CCP, CSP

Computer Science

Sam Houston State University

Huntsville Texas, 77341

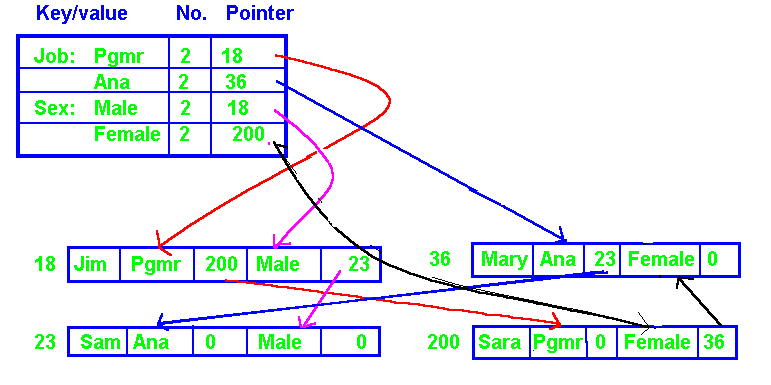
[csc\_dsb@shsu.edu](mailto:csc_dsb@shsu.edu)

Permission is granted to students in my classes use any or all of the materials in this document for their personnel education. All other rights reserved.

***MULTI-LISTS***

|  |  |  |
| --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |
| **Job: Programer**  **Analyst** | **2**  **2** | **18**  **36** |
| **Sex: Male**  **Female** | **2**  **2** | **18**  **200** |

|  |
| --- |
|  |

******

***Logical Record Fromat:***

|  |  |  |
| --- | --- | --- |
| ***Name*** | ***Job*** | ***Sex*** |

***Actual Record Format***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Name*** | ***Job*** | ***Job pointer*** | ***Sex*** | ***Sex pointer*** |

***Circular list are preferred over the implementation shown above (last points to start).***

***MULTI-LISTS***

|  |  |  |
| --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |
| **Job: Pgmr**  **Anal** | **4**  **3** | **18**  **36** |
| **Sex: Male**  **Female** | **5**  **2** | **18**  **36** |
| **Eyes: Blue**  **Green**  **Hazel** | **5**  **1**  **1** | **36**  **18**  **119** |
| **Age: 20**  **21**  **30** | **3**  **3**  **1** | **18**  **119**  **52** |

***Logical Record Format:***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **NAME** | **AGE** | **SEX** | **HAIR** | **JOB** | **EYES** |

***Actual Record Format:***

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Rec** | **Name** | **Age** | ***Pt*** | **Sex** | ***Pt*** | **Hair** | **Job** | ***Pt*** | **Eyes** | ***Pt*** |
|  |  |  |  |  |  |  |  |  |  |  |
| **18** | **Joe** | **20** | ***36*** | **Male** | ***52*** | **Brown** | **Pgmr** | ***119*** | **Green** | ***0*** |
| **36** | **Mary** | **20** | ***416*** | **Female** | ***416*** | **Black** | **Anal** | ***52*** | **Blue** | ***52*** |
| **52** | **Sam** | **30** | ***0*** | **Male** | ***119*** | **Blond** | **Anal** | ***200*** | **Blue** | ***200*** |
| **119** | **Harry** | **21** | ***200*** | **Male** | ***200*** | **Blond** | **Pgmr** | ***228*** | **Hazel** | ***0*** |
| **200** | **Ted** | **21** | ***228*** | **Male** | ***228*** | **Gray** | **Anal** | ***0*** | **Blue** | ***228*** |
| **228** | **Damon** | **21** | ***0*** | **Male** | ***0*** | **Black** | **Pgmr** | ***416*** | **Blue** | ***416*** |
| **416** | **Sara** | **20** | ***0*** | **Female** | ***0*** | **Blond** | **Pgmr** | ***0*** | **Blue** | ***0*** |

The multi-list format allows fast efficient processing of records sharing a common value for a secondary key after a random access.

***MULTI-LISTS***

|  |  |  |
| --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |
| **Job: Pgmr**  **Anal** | **40**  **30** | **18**  **36** |
| **Sex: Male**  **Female** | **50**  **20** | **18**  **36** |
| **Eyes: Blue**  **Green**  **Hazel** | **50**  **10**  **10** | **36**  **18**  **119** |
| **Age: 20**  **21**  **30** | **30**  **30**  **10** | **18**  **119**  **52** |

***Query: AND (logical conjunction)***

**Count and list all employees that are female and have blue eyes and are analysist.**

**#search = MIN(Female, Blue, Analyst)**

**= MIN( 20, 50, 30 )**

**= 20 records.**

**Twenty is an upper bound for the query, the lower bound is zero. *The upper bound is the shortest length list, e.g., 20 records in the example.***

***Query: OR (logical disjunction)***

**Count and list all employees that are female or have blue eyes or are analysist.**

**#search = #female + #blue + #analysist**

**= 20 + 50 + 30**

**= 100 records.**

**Note that 100 is the upper bound and the lower bound for the query is 50 records. *For "OR", the upper bound is the sum of the list lengths and the lower bound is the maximum length list. Some records may be processed more than once.***

***MULTI-LISTS***

|  |  |  |
| --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |
| **Job: Pgmr**  **Anal** | **40**  **30** | **18**  **36** |
| **Sex: Male**  **Female** | **50**  **20** | **18**  **36** |
| **Eyes: Blue**  **Green**  **Hazel** | **50**  **10**  **10** | **36**  **18**  **119** |
| **Age: 20**  **21**  **30** | **30**  **30**  **10** | **18**  **119**  **52** |

|  |
| --- |
|  |

**QUERY:**

**How many employees are Male and have blond hair?**

Pre-search statistic is only an upper bound.

**How many employees have blond hair?**

Pre-search statistic is only an upper bound.

**How many employees are male, and have blue eyes, and blond hair?**

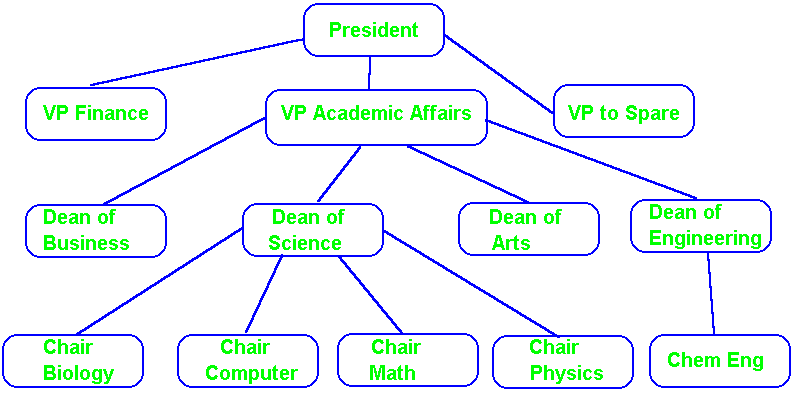
Pre-search statistic upper bound is length of shortest length key list with the lower bound of zero.

**How many employees are male, or have blue eyes, or age 30?**

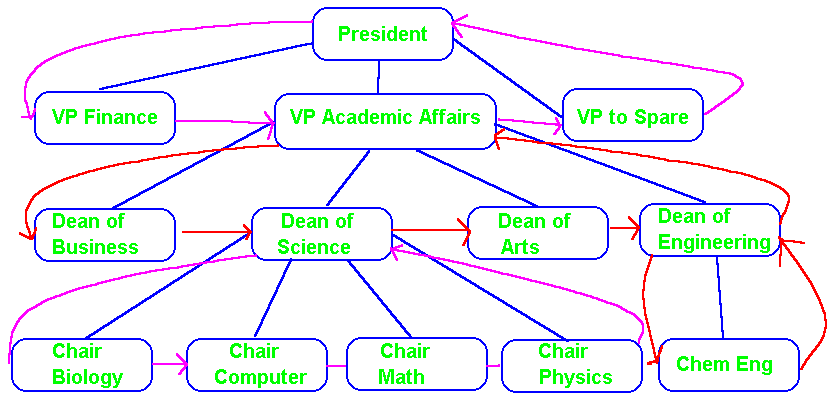
Pre-search statistic upper bound is sum of key list lengths with the lower bound being the longest length list.

**How many employees are over the average age?**

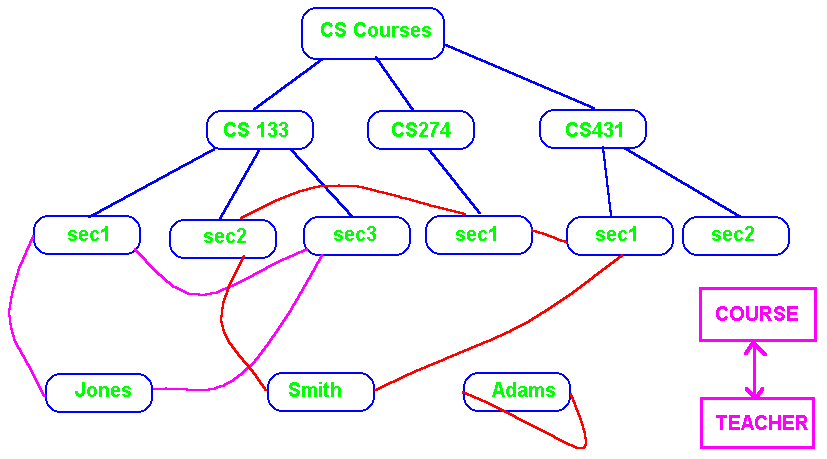
**How many employees do not have blue eyes?**

****

**Hierarchical Structure 1:M Relationship**

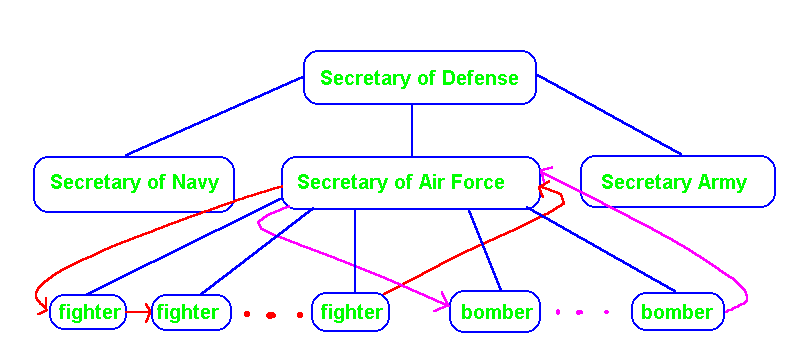
****

**Implemented using Circular Linked Lists**

****

**Network M:N Relationship**

**Note the use of the “bridge” node / record “section) to implement the many-to-many relationship between course and teacher.**

****

**A node can be the owner of multiple sets (relationships). The air force owns fighters and bombers yet they are different.**

**Multilist File and Directory Creation**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | **1** |  |  |  | **2** |  |  |  |  |
|  | **Input**  **File** |  | **Assign**  **Addresses**  **(AD) & strip keys** |  | **File**  **Sorted**  **by AN, AD**  **with keys** |  | **Generate**  **Key/AD/AN**  **Triples** |  | **Key/AD/AN**  **File sorted by AD/AN** |  |  |
|  | **Sorted by accession number(AN), keys, & non-key data** |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **3** |  | **3** |  | **4** |  |  |  |  |  |  |
|  | **Sort by Key**  **and AD within Key** |  | **File sorted by key, AN with in key: has AN** |  | **Construct Key Directory with head of list AD and List Length. Generate all link addresses (LA).** | | |  | **File sorted by Key, AD within Key containing RN & LA** |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | **5** |  |  |  | **6** |  |  |  |  |
|  | **File containing Key Directory** |  | **Sort by RN** |  | **File,**  **Key/AD/RN**  **Sorted by RN** |  | **Insert LA and create Master File** |  | **List Structured File in RN order** |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | **Contains AD, Keys, LA, & Non-Key data.** |  |  |
|  | **RN is record number** | | |  |  |  |  |  |  |  |  |
|  | **AD is DASD address** | | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

**Assume a relative file. Assign addresses as 1, 2, •••, N.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RN** | **Name** | **Eye** | **Sex** | **Other** |
|  |  |  |  |  |

**Input:**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **101** | **Joe** | **Blue** | **M** | **•••** |  |  |  |  |  |  |  |
| **102** | **Mary** | **Green** | **F** | **•••** |  |  |  |  |  |  |  |
| **103** | **Sam** | **Blue** | **M** | **•••** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  | Directory |  |  |  |
| **Block1** |  |  |  |  |  |  | **Block 4** | **Key** | **Point** | **LL** |  |
| **1** | **101** | **Joe** | **Blue** | **M** | **•••** |  |  | **Blue** | **1** | **2** |  |
| **2** | **102** | **Mary** | **Green** | **F** | **•••** |  |  | **F** | **2** | **1** |  |
| **3** | **103** | **Sam** | **Blue** | **M** | **•••** |  |  | **Green** | **2** | **1** |  |
|  |  |  |  |  |  |  |  | **Joe** | **1** | **1** |  |
| Block 2 |  |  |  |  |  |  |  | **M** | **1** | **2** |  |
| **1** | **101** | **Joe** |  |  |  |  |  | **Mary** | **2** | **1** |  |
| **1** | **101** | **Blue** |  |  |  |  |  | **Sam** | **3** | **1** |  |
| **1** | **101** | **M** |  |  |  |  |  |  |  |  |  |
| **2** | **102** | **Mary** |  |  |  |  |  | **Blue** | **1** | **101** | **3** |
| **2** | **102** | **Green** |  |  |  |  |  | **Blue** | **3** | **103** | **0** |
| **2** | **102** | **F** |  |  |  |  |  | **F** | **2** | **102** | **0** |
| **3** | **103** | **Sam** |  |  |  |  |  | **Green** | **2** | **102** | **0** |
| **3** | **103** | **Blue** |  |  |  |  |  | **Joe** | **1** | **101** | **0** |
| **3** | **103** | **M** |  |  |  |  |  | **M** | **1** | **101** | **3** |
|  |  |  |  |  |  |  |  | **M** | **3** | **103** | **0** |
| Block 3 |  |  |  |  |  |  |  | **Mary** | **2** | **102** | **0** |
| **1** | **101** | **Blue** |  |  |  |  |  | **Sam** | **3** | **103** | **0** |
| **3** | **103** | **Blue** |  |  |  |  |  |  |  |  |  |
| **2** | **102** | **F** |  |  |  |  | Block 5 | **Blue** | **1** | **101** | **3** |
| **2** | **102** | **Green** |  |  |  |  |  | **Joe** | **1** | **101** | **0** |
| **1** | **101** | **Joe** |  |  |  |  |  | **M** | **1** | **101** | **3** |
| **1** | **101** | **M** |  |  |  |  |  | **F** | **2** | **102** | **0** |
| **3** | **103** | **M** |  |  |  |  |  | **Green** | **2** | **102** | **0** |
| **2** | **102** | **Mary** |  |  |  |  |  | **Mary** | **2** | **102** | **0** |
| **3** | **103** | **Sam** |  |  |  |  |  | **Blue** | **3** | **103** | **0** |
|  |  |  |  |  |  |  |  | **M** | **3** | **103** | **0** |
|  |  |  |  |  |  |  |  | **Sam** | **3** | **103** | **0** |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Block 6** | **File** | **Format** |  |  |  |  |  |  |  |  |  |
| **Address** | **Name** | **NL** | **Eye** | **EL** | **Sex** | **SL** | **Other** |  |  |  |  |
| **1** | **Joe** | **0** | **Blue** | **3** | **M** | **3** | **•••** |  |  |  |  |
| **2** | **Mary** | **0** | **Green** | **0** | **F** | **0** | **•••** |  |  |  |  |
| **3** | **Sam** | **0** | **Blue** | **0** | **M** | **0q** | **•••** |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

**XL implies a link field for key X.**

**Complete the following Mulitlist**

1. **Using the direct address.**
2. **Using the Name field as the key.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Key** | **LL** | **Pt** |  |  |  |  |  |  |  |  |  |  |  |
| **Job**  **PGMR**  **ANAL**  **Sex**  **Male**  **Female** |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Jill** | **PGMR** |  | **26** | **F** |  |  |  |  |
|  |  |  |  |  | **10** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Sam** | **PGMR** |  | **26** | **M** |  |  |  |  |
|  |  |  |  |  | **18** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Sue** | **ANAL** |  | **22** | **F** |  |  |  |  |
| **Key** | **Pt** |  |  |  | **12** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Mary** | **PGMR** |  | **22** | **F** |  |  |  |  |
|  |  |  |  |  | **16** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Joe** | **ANAL** |  | **20** | **M** |  |  |  |  |
|  |  |  |  |  | **28** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | **Name** | **Job** | **JPT** | **Age** | **Sex** | **Spt** |  |  |  |  |  |
|  |  |  | **Record Format** | | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Multilist Using Direct Address**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Key** | **LL** | **Pt** |  |  |  |  |  |  |  |  |  |  |  |
| **Job**  **PGMR**  **ANAL**  **Sex**  **Male**  **Female** | **3**  **2**  **2**  **3** | **10**  **28**  **18**  **10** |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Jill** | **PGMR** | **18** | **26** | **F** | **12** |  |  |  |
|  |  |  |  |  | **10** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Sam** | **PGMR** | **16** | **26** | **M** | **28** |  |  |  |
|  |  |  |  |  | **18** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Sue** | **ANAL** | **0** | **22** | **F** | **16** |  |  |  |
| **Key** | **Pt** |  |  |  | **12** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Mary** | **PGMR** | **0** | **22** | **F** | **0** |  |  |  |
|  |  |  |  |  | **16** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Joe** | **ANAL** | **12** | **20** | **M** | **0** |  |  |  |
|  |  |  |  |  | **28** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | Name | **Job** | **JPT** | **Age** | **Sex** | **Spt** |  |  |  |  |  |
|  |  |  | **Record Format** | | |  |  |  |  |  |  |  |  |

**“0” indicates a null link.**

**Logical pointers have an advantage over physical pointers in that they are not tied to the hardware. This makes the file more portable. On the other hand, physical pointers do not have to translated at runtime hence are faster. Preference should probably be given to the logical pointers due to the device independence and advantages gained with respect to archiving (backup).**

**Mulitlist Using Direct Name Field**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Key** | **LL** | **Pt** |  |  |  |  |  |  |  |  |  |  |  |
| **Job**  **PGMR**  **ANAL**  **Sex**  **Male**  **Female** | **3**  **2**  **2**  **3** | **Jill**  **Joe** Sam **Jill** |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Jill** | **PGMR** | Sam | **26** | **F** | Sue |  |  |  |
|  |  |  |  |  | **10** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Sam** | **PGMR** | Mary | **26** | **M** | Joe |  |  |  |
|  |  |  |  |  | **18** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **Sue** | **ANAL** | **0** | **22** | **F** | Mary |  |  |  |
| **Key** | **Pt** |  |  |  | **12** |  |  |  |  |  |  |  |  |
| **Jill** | **10** |  |  |  |  |  |  |  |  |  |  |  |  |
| **Joe** | **28** |  |  |  |  |  |  |  |  |  |  |  |  |
| **Mary** | **16** |  |  |  | **Mary** | **PGMR** | **0** | **22** | **F** | **0** |  |  |  |
| **Sam** | **18** |  |  |  | **16** |  |  |  |  |  |  |  |  |
| **Sue** | **12** |  |  |  |  |  |  |  |  |  |  |  |  |
| **Indexed** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Directory** |  |  |  |  | **Joe** | **ANAL** | Sue | **20** | **M** | **0** |  |  |  |
|  |  |  |  |  | **28** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | Name | **Job** | **JPT** | **Age** | **Sex** | **Spt** |  |  |  |  |  |
|  |  |  | **Record Format** | | |  |  |  |  |  |  |  |  |

**“0” indicates a null link.**

***INVERTED LISTS***

***Logical “Person” record format using “NAME” as the key.***

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **address** | **NAME** | **AGE** | **SEX** | **HAIR** | **JOB** | **EYES** | **DEPT** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **18** | **Joe** | **20** | **Male** | **Brown** | **Pgmr** | **Green** | **Acc** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **36** | **Mary** | **20** | **Female** | **Black** | **Anal** | **Blue** | **DP** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **52** | **Sam** | **30** | **Male** | **Blond** | **Anal** | **Blue** | **DP** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **119** | **Harry** | **21** | **Male** | **Blond** | **Pgmr** | **Hazel** | **Inv** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **200** | **Ted** | **21** | **Male** | **Gray** | **Anal** | **Blue** | **Sales** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **228** | **Damon** | **21** | **Male** | **Black** | **Pgmr** | **Blue** | **Prod** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **416** | **Sara** | **20** | **Female** | **Blond** | **Pgmr** | **Blue** | **Sales** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Directory** | **Key** | **Point** | **Directory actually maintained as B-Tree** |
|  | **Damon** | **228** | **or related technology to allow efficient** |
|  | **Harry** | **119** | **insertion, deletion, and modification of** |
|  | **Joe** | **18** | **records while allowing random lookup** |
|  | **Mary** | **36** | **and sequential processing of records** |
|  | **Sam** | **52** | **in ascending key order.** |
|  | **Sara** | **416** |  |
|  | **Ted** | **200** |  |

**Typical user questions:**

1. **List all employees having a specific job, e.g., programmer.**
2. **List all employees based on sex.**
3. **List all employees with a specific eye color.**
4. **List all employees sharing a specific age.**

***INVERTED LISTS***

***Logical “Person” record format using “NAME” as the key.***

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **address** | **NAME** | **AGE** | **SEX** | **HAIR** | **JOB** | **EYES** | **DEPT** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **18** | **Joe** | **20** | **Male** | **Brown** | **Pgmr** | **Green** | **Acc** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **36** | **Mary** | **20** | **Female** | **Black** | **Anal** | **Blue** | **DP** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **52** | **Sam** | **30** | **Male** | **Blond** | **Anal** | **Blue** | **DP** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **119** | **Harry** | **21** | **Male** | **Blond** | **Pgmr** | **Hazel** | **Inv** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **200** | **Ted** | **21** | **Male** | **Gray** | **Anal** | **Blue** | **Sales** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **228** | **Damon** | **21** | **Male** | **Black** | **Pgmr** | **Blue** | **Prod** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **416** | **Sara** | **20** | **Female** | **Blond** | **Pgmr** | **Blue** | **Sales** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Directory** | **Key** | **Point** | **Directory actually maintained as B-Tree** |
|  | **Damon** | **228** | **or related technology to allow efficient** |
|  | **Harry** | **119** | **insertion, deletion, and modification of** |
|  | **Joe** | **18** | **records while allowing random lookup** |
|  | **Mary** | **36** | **and sequential processing of records** |
|  | **Sam** | **52** | **in ascending key order.** |
|  | **Sara** | **416** |  |
|  | **Ted** | **200** |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |  |
| **Job: Pgmr**  **Anal** | **4**  **3** | **18, 119, 228, 416**  **36, 52, 200** | **INVERTED** |
| **Sex: Male**  **Female** | **5**  **2** | **18 ,52, 119, 200, 228**  **36, 416** | **LIST** |
| **Eyes: Blue**  **Green**  **Hazel** | **5**  **1**  **1** | **36, 52, 200, 228, 416**  **18**  **119** | **DIRECTORY** |
| **Age: 20**  **21**  **30** | **3**  **3**  **1** | **18, 36, 416**  **119, 200, 228**  **52** | **(Partially versus fully inverted)** |

**Index allows 1:M relationship, e.g., secondary key to records with key. Indexes are used to implement foreign key concept in relational database.**

**Ex: List all programmers or list everyone with blue eyes, .i.e, specific key value to all records containing the value.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **18** | **Joe** | **20** | **Male** | **Brown** | **Pgmr** | **Green** | **Acc** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **36** | **Mary** | **20** | **Female** | **Black** | **Anal** | **Blue** | **DP** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **52** | **Sam** | **30** | **Male** | **Blond** | **Anal** | **Blue** | **DP** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **119** | **Harry** | **21** | **Male** | **Blond** | **Pgmr** | **Hazel** | **Inv** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **200** | **Ted** | **21** | **Male** | **Gray** | **Anal** | **Blue** | **Sales** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **228** | **Damon** | **21** | **Male** | **Black** | **Pgmr** | **Blue** | **Prod** |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **416** | **Sara** | **20** | **Female** | **Blond** | **Pgmr** | **Blue** | **Sales** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Directory** | **Key** | **Point** | **Directory actually maintained as B-Tree** |
|  | **Damon** | **228** | **or related technology to allow efficient** |
|  | **Harry** | **119** | **insertion, deletion, and modification of** |
|  | **Joe** | **18** | **records while allowing random lookup** |
|  | **Mary** | **36** | **and sequential processing of records** |
|  | **Sam** | **52** | **in ascending key order.** |
|  | **Sara** | **416** |  |
|  | **Ted** | **200** |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |  |
| **Job: Pgmr**  **Anal** | **4**  **3** | **18, 119, 228, 416**  **36, 52, 200** | **INVERTED** |
| **Sex: Male**  **Female** | **5**  **2** | **18 ,52, 119, 200, 228**  **36, 416** | **LIST** |
| **Eyes: Blue**  **Green**  **Hazel** | **5**  **1**  **1** | **36, 52, 200, 228, 416**  **18**  **119** | **DIRECTORY** |
| **Age: 20**  **21**  **30** | **3**  **3**  **1** | **18, 36, 416**  **119, 200, 228**  **52** | **(Partially versus fully inverted)** |

**Typical user questions:**

1. **List all employees having a specific job, e.g., programmer.**
2. **List all employees based on sex.**
3. **List all employees with a specific eye color.**
4. **List all employees the sharing a specific age.**

Inverted List

**Directory Structure Questions (relations)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **1:1** |  |  |  |  |  |  |
|  | Key | Pointer |  | Attribute | **1:M** | Record |  |
|  | **Damon** | **228** |  | **Job** |  | **Person** |  |
|  | **Harry** | **119** |  |  |  |  |  |
|  | **Joe** | **18** |  | **Sex** |  | **Person** |  |
|  | **Mary** | **36** |  |  |  |  |  |
|  | **Sara** | **416** |  | **Eye Color** |  | **Person** |  |
|  | **Sam** | **52** |  |  |  |  |  |
|  | **Ted** | **200** |  | **Age** |  | **Person** |  |

**EX: List all persons with specific attribute (eye color is hazel). If all fields in the record are treated as keys, the list is termed full inverted versus partially inverted (not all fields are keys).**

**Directory Structure**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Key | Pointer |  |  |  |  |  |  |  |  |  |  |  |
|  | **Damon** | **228** |  |  | 18 | **Joe** | **20** | **Male** | **Brown** | **Pgmr** | **Green** | **Acc** |  |
|  | **Harry** | **119** |  |  |  |  |  |  |  |  |  |  |  |
|  | **Joe** | **18** |  |  | 36 | **Mary** | **20** | **Female** | **Black** | **Anal** | **Blue** | **DP** |  |
|  | **Mary** | **36** |  |  |  |  |  |  |  |  |  |  |  |
|  | **Sara** | **416** |  |  | 52 | **Sam** | **30** | **Male** | **Blond** | **Anal** | **Blue** | **DP** |  |
|  | **Sam** | **52** |  |  |  |  |  |  |  |  |  |  |  |
|  | **Ted** | **200** |  |  | 119 | **Harry** | **21** | **Male** | **Blond** | **Pgmr** | **Hazel** | **Inv** |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | 200 | **Ted** | **21** | **Male** | **Gray** | **Anal** | **Blue** | **Sales** |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | 228 | **Damon** | **21** | **Male** | **Black** | **Pgmr** | **Blue** | **Prod** |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | 416 | **Sara** | **20** | **Female** | **Blond** | **Pgmr** | **Blue** | **Sales** |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Key / Value** | **Num** | **Pointers** |  | ***Considerations:*** |
|  | **Job:** Pgmr  Anal | 4  3 | 18, 119, 228, 416  36, 52, 200 |  | **1) Whole Record Insertion**  **2) Whole Record Deletion** |
|  | **Sex:** Male  Female | 5  2 | 18,52, 119, 200, 228  36, 416 |  | **3) Qualifier Modification (non key)**  **4) Primary Key Modification** |
|  | **Eyes:** Blue  Green  Hazel | 5  1  1 | 36, 52, 200, 228, 416  18  119 |  | **5) Secondary Key Modification** |
|  | **Age:** 20  21  30 | 3  3  1 | 18, 36, 416  119, 200, 228  52 |  | **B-Trees are used to implement the logically sequential directory.** |

Foreign Keys in Relational Database (logically in records on many side of relation):

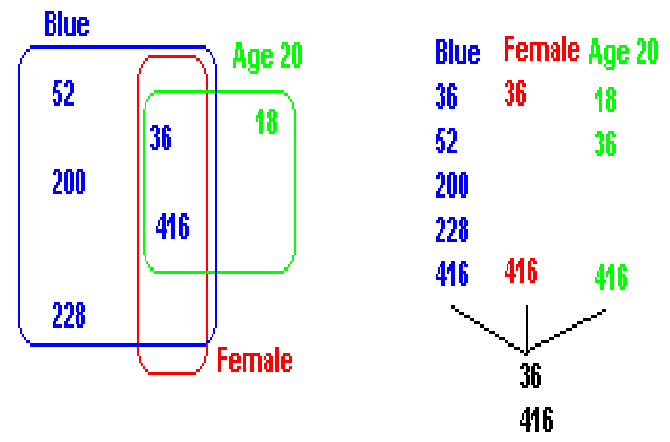
Employer -----> Employee (each employee has key of employer):

***INVERTED LISTS***

|  |  |  |
| --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |
| **Job: Pgmr**  **Anal** | **4**  **3** | **18, 119, 228, 416**  **36, 52, 200** |
| **Sex: Male**  **Female** | **5**  **2** | **18 ,52, 119, 200, 228**  **36, 416** |
| **Eyes: Blue**  **Green**  **Hazel** | **5**  **1**  **1** | **36, 52, 200, 228, 416**  **18**  **119** |
| **Age: 20**  **21**  **30** | **3**  **3**  **1** | **18, 36, 416**  **119, 200, 228**  **52** |

***Query: AND (logical conjunction) using “index:” Exact Pre-Search Statistic!***

**Count and list all employees that are Female and have Blue eyes and are AGE 20.**

****

**Select \*Name**

**From Person**

**Where Sex = ‘Female’ and Eyes = ‘Blue’ and Age = 20**

***Query: OR (logical disjunction) using” index:” Exact Pre-Search Statistic!***

**Count and list all employees that are Female or have Blue eyes or are age 20.**

**Select \* from Employee where Sex = ‘Female’ or Eyes = ‘Blue’ or Age = 20**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Blue** | **36, 52, 200, 228, 416** |  |  |  |  |  |
|  | **Female** | **36, 416** |  | **18, 36, 52, 200, 228, 416** | | |  |
|  | **Age** | **18, 36, 416** |  |  |  |  |  |

***INVERTED LISTS***

|  |  |  |
| --- | --- | --- |
| ***KEY/VALUE*** | ***NO.*** | ***POINTERS*** |
| **Job: Pgmr**  **Anal** | **4**  **3** | **18, 119, 228, 416**  **36, 52, 200** |
| **Sex: Male**  **Female** | **5**  **2** | **18 ,52, 119, 200, 228**  **36, 416** |
| **Eyes: Blue**  **Green**  **Hazel** | **5**  **1**  **1** | **36, 52, 200, 228, 416**  **18**  **119** |
| **Age: 20**  **21**  **30** | **3**  **3**  **1** | **18, 36, 416**  **119, 200, 228**  **52** |

**QUERY:**

**How many employees are Male and have blond hair?**

Pre-search statistic is only an upper bound.

**How many employees have blond hair?**

Sequential search of all records is required.

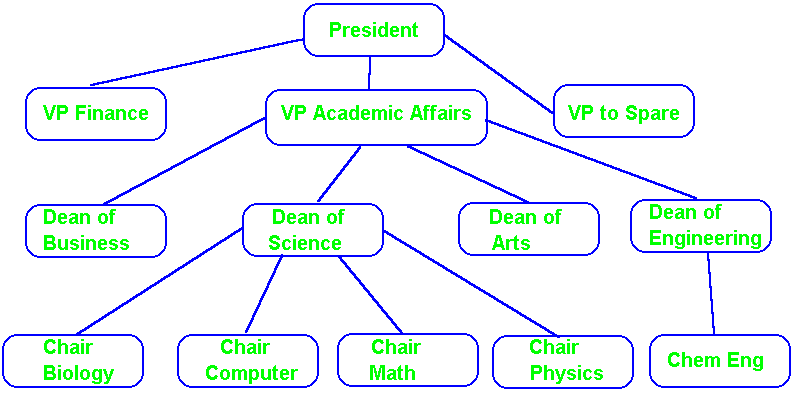
**How many employees are male, have blue eyes, and have blond hair?**

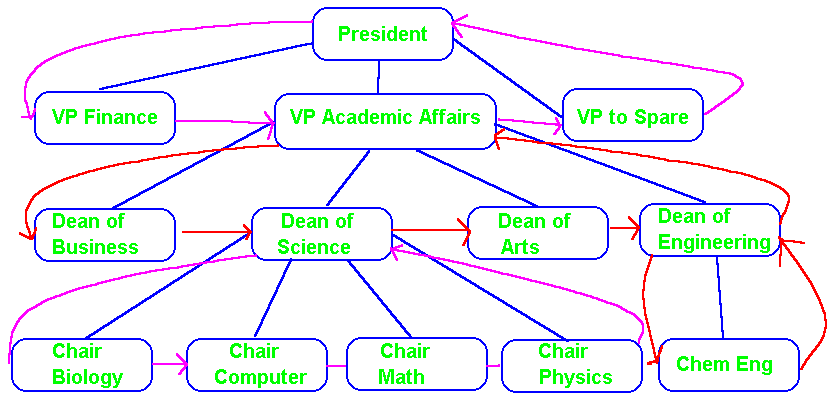
Pre-search statistic by merging addresses of male and blue eyes is an upper bound. The lower bound is zero and may only be determined by retrieving the records.

**How many employees are over the average age?**

**How many employees do not have blue eyes?**

**Each inverted list directly implements a 1:M relationship.**

****

****

**Implemented using Circular Linked Lists**

**Use of circular list are common for network architecture with the pointers buried in the nodes. Relational database prefer keeping all pointers in the directory. Logically the circular list does however exist.**

**Inverted List Creation**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **List Structured File** |  | **AD, sorted by AN, keys (optional), not-key data** |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | **1** |  |  |  | **2** |  |  |  |  |
|  | **Input**  **File** |  | **Assign**  **Addresses**  **(AD)** |  | **File**  **Sorted**  **by AD** |  | **Generate**  **Key/AD**  **Pairs** |  | **Key/AD**  **File sorted by AD** |  |  |
|  | **Sorted by accession number(AN), contains keys, & non-key data** |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | **3** |  |  |  | **4** |  |  |  |  |
|  |  |  | **Sort by Key**  **& AD within Key** |  | **File sorted by key, AD within key** |  | **Construct Inverted List Key Directory** |  | **File w / Inverted List Directory** |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | **Key 1**  **AD**  **AD**  **Key 2**  **AD**  **AD**  **•••** |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

**Assume a relative file. Assign addresses as 1, 2, •••, N.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **AN** | **Name** | **Eye** | **Sex** | **Other** |
|  |  |  |  |  |

**Input:**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **101** | **Joe** | **Blue** | **M** | **•••** |  |  |  |  |  |  |  |
| **102** | **Mary** | **Green** | **F** | **•••** |  |  |  |  |  |  |  |
| **103** | **Sam** | **Blue** | **M** | **•••** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Block1** |  |  |  |  |  |  |  |  |  |  |  |
| **AD** | **AN** | **Name** | **Eye** | **Sex** | **Other** |  | **Block 4** | **Key** | **LL** | **Point** |  |
| **1** | **101** | **Joe** | **Blue** | **M** | **•••** |  |  | **Blue** | **2** | **1, 3** |  |
| **2** | **102** | **Mary** | **Green** | **F** | **•••** |  |  | **F** | **1** | **2** |  |
| **3** | **103** | **Sam** | **Blue** | **M** | **•••** |  |  | **Green** | **1** | **2** |  |
|  |  |  |  |  |  |  |  | **Joe** | **1** | **1** |  |
| **Block 2** |  |  |  |  |  |  |  | **M** | **2** | **1, 3** |  |
| **1** | **Joe** |  |  |  |  |  |  | **Mary** | **1** | **2** |  |
| **1** | **Blue** |  |  |  |  |  |  | **Sam** | **1** | **3** |  |
| **1** | **M** |  |  |  |  |  |  |  |  |  |  |
| **2** | **Mary** |  |  |  |  |  |  |  |  |  |  |
| **2** | **Green** |  |  |  |  |  |  |  |  |  |  |
| **2** | **F** |  |  |  |  |  |  |  |  |  |  |
| **3** | **Sam** |  |  |  |  |  |  |  |  |  |  |
| **3** | **Blue** |  |  |  |  |  |  |  |  |  |  |
| **3** | **M** |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Block 3** |  |  |  |  |  |  |  |  |  |  |  |
| **Blue** | **1** |  |  |  |  |  |  |  |  |  |  |
| **Blue** | **3** |  |  |  |  |  |  |  |  |  |  |
| **F** | **2** |  |  |  |  |  |  |  |  |  |  |
| **Green** | **2** |  |  |  |  |  |  |  |  |  |  |
| **Joe** | **1** |  |  |  |  |  |  |  |  |  |  |
| **M** | **1** |  |  |  |  |  |  |  |  |  |  |
| **M** | **3** |  |  |  |  |  |  |  |  |  |  |
| **Mary** | **2** |  |  |  |  |  |  |  |  |  |  |
| **Sam** | **3** |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |