CSCI 3901 FINAL PROJECT – BREAKDOWN ANALYSIS

Submission by: Nadipineni Hemanth Kumar [B00899473]

Date: 2021-11-02

Overview of the project:

With the abundance of storage space in the new era of technology, people tend not to delete the pictures, thus creating difficulty in finding images and videos. [The study of families, family history, and the tracing of their lineages is known as **Genealogy**. Oral interviews, historical records, genetic analysis, and other records are used by genealogists to gather information about a family and to prove kinship and pedigrees of its members[i].]

The main objective of this project is to develop a system that connects family tree information (Focusing only on the biological relations) with a photo repository and the metadata associated with the photos.

- The system consists of a family tree database and a media archive.
- Store the information of individuals and their relations in the family tree database.
- Store the information of photos in media archive.
- Information can be incomplete in database except individual_name in the family tree database and filename in the media archive.
- With all this information, we need to give output for the queries given by genealogists (such as list of ancestors of person X for Z generations).

Abbreviations:

- A – a person named 'A'
- B – a person name 'B'
- nA - the number of generations A is back from the common ancestor
- nB - the number of generations B is back from the common ancestor

Formulae used:

- Degree of cousinship: min(nA, nB)-1
- Degree of removal: |nA-nB|

Relationships defined:

- Ancestor/Descendent Relation: when either of nA or nB is 0.
- Sibling: when min{nA, nB} is 1.
- X cousins Y removed; where X = Degree of cousinship, Y = Degree of removal.
- Sibling Y removed; when min{nA, nB} is one and degree of removal>0;
  Y = Degree of removal

Classes defined:

- PersonIdentity: Identifier for a person in the tree
- FileIdentifier: Identifier for the file
- BiologicalRelation: Class to specify the cousinship and level of removal.

Critical Analysis:

1) **What comes into the program?**

   Family Tree Database:
   - Name of the person (personName) [String] [Mandatory; NOT NULL]: from addPerson()
   - Date of Birth (dateOfBirth) [String]: from recordAttributes()
   - Location of Birth (locationOfBirth) [String]: from recordAttributes()
   - Date of Death (dateOfDeath) [String]: from recordAttributes()
   - Location of Death (locationOfDeath) [String]: from recordAttributes()
   - Gender (gender) [String]: from recordAttributes()
   - Occupation (occupation) [String]: from recordAttributes()
   - References to source material (reference) [String]: from recordReference()
   - Notes on the person (notesOfPerson) [String]: from recordNote()
   - Parent or Child relations (relationOfPerson) [String]: from recordChild()
   - Ceremony Relations (ceremonyRelationOfPerson) [String: List of partners]: from recordPartnering()
   - Partnering Dissolutions (dissolutionOfPerson) [String: List of dissolutions]: from recordDissolution()

   Media Archive:
   - The location of a file (fileLocation) [String] [Mandatory; NOT NULL]: from addMediaFile()
   - Date of the file taken (dateFile) [String]: from recordMediaAttributes()
   - Location of the file (locationFile) [String]: from recordMediaAttributes()
   - The name of city the file is taken at (cityFile) [String]: from recordMediaAttributes()
   - The province at which the file is taken at (provinceFile) [String]: from recordMediaAttributes()
   - The country at which the file is taken at (countryFile) [String]: from recordMediaAttributes()
   - Tags for the file (tagesFile) [String]: from tagMedia ()
   - Persons in the file (presonsFile) [String: List]: from peopleInMedia()

2) **What transformations do I need to make to the data?**
   - Find the nearest common ancestor for the persons A and B.
   - Find nA for the person A.
   - Find nB for the person B.

- Calculate the degree of cousinship using nA and nB
- Calculate the degree of removal using nA and nB

3) **What part of the data is processed right away?**
   - Set up a relation as soon as the person is added with other relevant data

4) **What part of the data do I need to keep longer?**
   - **a.** The names of the people
   - **b.** The identifier
   - **c.** The relations among the people
   - **d.** The notes for each person
   - **e.** The references for each person
   - **f.** The name of the file
   - **g.** The identifier of the media file
   - **h.** The tags of the media file
   - **i.** The location at which the media file is taken
   - **j.** The dates at which the media files were taken

5) **What goes out of the program?**
   - The name of a person when used findPerson() to locate the individual in the tree.
   - The name of the person when used findMediaFile() to locate the individual
   - The name of the person when used findName() to identify a person using ID.
   - The name of the file when used findMediaFile() with a file identifier.
   - The relation between two persons when used findRelation() between person A and person B.
   - The descendants of a person within X generations, where X is passed to descendents().
   - The ancestors of a person within X generations, where X is passed to ancestors().
   - The notes and references of a person when used notesAndReferences().
   - The names of files linked to a tag within a date range, when used notesAndReferences(). [Note: Null indicates no restriction on dates]
   - The names of files taken at a location within a date range, when used findMediaByLocation(). [Note: Null indicates no restriction on dates]
   - The names of files, taken within a date range, that contains any person in a list of people, when used findIndividualsMedia(). [Note: Null indicates no restriction on dates]
   - The names of files that contains immediate child/children of a person in ascending order of dates taken (Sort by recent) when used findBiologicalFamilyMedia().

Assumptions made:

All the assumptions in the problem statement are to be considered.

Odd Cases:

- The case when all the persons have just the name and no other information.
- The case when all the files have just the filename and no other information.
- The case when all the persons and the files have just the names.

Known Limitations:

- Relationships are limited to only biological.
- Filetypes are limited to only photos and videos.

What is important for the solution to do:

- Creating a tree with family relations.
- Calculation of the nearest common ancestor.
- Calculating degree of cousinship and degree of removal.

Notes:

The information present in this document is taken from the problem statement.
I consent that I neither shared nor showed this document to anyone.

[i] En.wikipedia.org. 2021. *Genealogy - Wikipedia*. [online] Available at: <https://en.wikipedia.org/wiki/Genealogy> [Accessed 4 November 2021].