CS 1555/2055: Database Management Systems
Department of Computer Science, University of Pittsburgh

Relational Algebra

---

1. [18 points: 6 points each] Consider running a webmail service. Assume that the data for all users of this service is represented in four relations with the following schemas:

   - **Mail** (mailID, senderID, timeSent, replyTo, subject, msgText, urgent)
   - **Recipients** (mailID, recipientID)
   - **Users** (userID, fname, lname, cell, city, country)
   - **Labels** (mailID, label)

   Assume that senderIDs and recipientIDs are valid userIDs in the Users relation, and each user has a unique ID. Further, replyTo values are valid mailIDs in the Mail relation (or NULL) and each mail has a unique ID. timeSent represents a full timestamp (date and time) of when a mail was sent, and urgent is a boolean flag set by the sender as to whether or not the mail should be considered to be urgent. Finally, assume that a single mail can have multiple labels applied to it, and all values of the label attribute are simple strings.

   Assume that these relations have the following cardinalities: **Mail**: 2000, **Recipients**: 3500, and **Users**: 100, and **Labels**: 2500 (the service is still being beta tested...). Find the arity and cardinality of the following relations (where exact values can not be determined, give the min and max values). State any assumptions that you make. Assume that $\bar{\bowtie}$ is the left outer join operator.

   $a)$ $\pi_{country}(Users)$

   $b)$ $Mail \bar{\bowtie}_{Mail.mailID=Recipients.mailID} Recipients$

   $c)$ $Mail * Labels$

2. [82 points total] Consider the relational schema mentioned in the description, write the **relational algebra** expression(s) for each of the following queries. For the time of each mail use the standard format **'YYYY-MM-DD HH:mm:ss'**:

   $a)$ [15 points] List the sender's first and last name, his/her user ID, the subject, and the time sent of all urgent mail.

   $b)$ [15 points] List all user IDs of users who have sent at least one urgent mail and live outside the USA.

   $c)$ [15 points] Count the number of mails sent during the year of 2015 that have the label "important" but were not flagged as urgent.

d) [17 points] List the first and last names of all the users who have sent mails to the everyone Ada Lovelace has sent a mail to.

e) [20 points] List all mail IDs of all mails that are not replies and have either a sender or recipient that lives in the USA.

What to submit

In a **single** typed file (PDF), please submit your answers to all questions. Name the file in following way: `hw#-YOUR_USERNAME.pdf` (where `#` is the number of the assignment). **No hand-written assignments are accepted. Any hand-written and scanned assignment will not be graded.** If you have trouble generating any of the relational algebra symbols, for example R $\bowtie_{A=B}$ S, use a descriptive word, e.g., R Join(A=B) S.

Submit your assignment by the due date. **There is no late submission.**