This problem is to be done on a group basis following the Trinity University Academic Integrity Policy. Laboratory problems should be submitted electronically (e-mail to cs3194@ariel.cs.trinity.edu) on or before the due date and should contain a problem write-up, source code to any programs and data sets used in solving the problem. The submitted files should be ASCII text files having Unix end-of-line characters (please convert all Windows and Mac text files to Unix format—I have found that Emacs or the Stone Text Tool seem to do a reasonable job of such conversions). If several files need to be submitted, put them in a directory having name your-last-name-problem-set-number and create a tar archive of this file system and attach it to your e-mail problem submission.

Introduction

The third year design project is to design a system for filtering e-mail messages. By filtering, we mean analyze e-mail headers and/or content and determine what classification should be assigned each e-mail message.

Design teams must:

- Research the e-mail classification literature.
- You should study existing designs and implementations for e-mail classification to determine their features and limitations.
- Decide what a better design means in the context of their research effort.
- Complete a prototype implementation of an e-mail classification system which contains the features which make a better system. In this context, the terminology prototype must be a functional prototype. User interface considerations are considered secondary in this problem.

Design Groups

There are 31 third year design students. The following J expressions were used to select 5 design groups.

```
names =: 'Arnold, Katherine M.';'Austin, Ryan E.';'Baker, Chris B.';'Barth, David A.';'Caille, Ryan P.'
names =: names,'Calcaterra, Gina M.';'Choo, Tai-San';'Crowe, Loren D.';'DeFelice, David C.';'Hohimer, Christopher J.'
names =: names,'Houck, David P.';'Leeds, Gregory O.';'Martin, Jess A.';'Martin, Victor'
names =: names,'McKenna, Ryan';'Michalson, Brett P.';'Miller, Travis A.';'Nasr, Carmen T.'
names =: names,'Nevin, Daniel R.';'Papazian, John H.';'Quilici, Bryan D.';'Raisani, Jabbar K.'
names =: names,'Riela, Sandra J.';'Rossmeisl, Joshua S.';'Rubio III, Robert';'Salter, Eric L.'
names =: names,'Scholz, Ryan B.';'Schwartz, Scott L.';'Schwarz, Paul D.';'Smith, Christopher A.';'Stoeckl, Kieth R.'
names =: names,'Wack, Rosellen K.';'Wilbraham, Ryan M.';'Wilmes, Theodore R.'
groups =: 'Group 1';'Group 2';'Group 3';'Group 4';'Group 5';'Group 6'
(/: /: (28$1 0 0 0 0 0),12$1 0 0 0 0 0) { ((34?34){>names),>groups
```

Group 1
Barth, David A.
Smith, Christopher A. (Group Leader)
Scholz, Ryan B.
Schwarz, Paul D.
Papazian, John H.
Raisani, Jabbar K.
Group 2
Rubio III, Robert
Calcaterra, Gina M.
McKenna, Ryan
Miller, Travis A.
Martin, Victor
Caille, Ryan P.
Group 3
Austin, Ryan E.
Crowe, Loren D.
Michalson, Brett P.
Choo, Tai-San
Stoeck, Kieth R.
Martin, Jess A.
Group 4
Baker, Chris B.
Quilici, Bryan D.
Nevin, Daniel R. (Group Leader)
Wack, Rosellen K.
Wilbraham, Ryan M.
Leeds, Gregory O.
Group 5
Schwartz, Scott L. (Group Leader)
Arnold, Katherine M.
Houck, David P.
Wilmes, Theodore R.
Nasr, Carmen T.
Group 6
Hohimer, Christopher J. (Group Leader)
Riela, Sandra J.
DeFelice, David C.
Salter, Eric L.
Rossmeisl, Joshua S.

Each group must meet and select a group leader before the beginning of the second class period (January 22, 2003) and e-mail the instructor cs3194@ariel.cs.trinity.edu the results of this selection.

Design groups must submit preliminary design documents to the instructor by e-mail cs3194@ariel.cs.trinity.edu before the beginning of the fourth class period (February 12, 2003).

E-Mail Classification Designs

The idea of a better e-mail classification system must be independently formulated by each design group. It is expected that groups will engage in a competition with each other in much the same fashion as companies might if they were designing and marketing competing e-mail classification software products.

Design Project Solution [ HTML ] [ PS ] [ PDF ]