Processing Address Lists in Icon

Ralph E. Griswold

Department of Computer Science, The University of Arizona

Introduction

Version 8.1 of the Icon program library contains a collection of programs for processing address lists. These programs check the correctness of address lists, filter them for designated entries, sort them, and format mailing labels.

The format of addresses lists processed by these programs is loosely structured. This allows such lists to be created and maintained using any text editor and allows them to be used for a variety of purposes (not just for addresses, although that term is used here for simplicity). The lack of structure, on the other hand, allows ambiguities and the possibility of incorrectly organized data. These programs are no substitute for a database system or an application specifically dedicated to the handling of mailing lists.

Address List Format

An address list, in the sense the term is used here, is a sequence of *entries*. Each entry begins with a *header* line, in which the first character is a #. Subsequent lines contain the address information in a natural format with a few constraints that are necessary if some of the programs described in the next section are to be used. For example, an address list might look like this:

```
#
Mr. Fred Burfle
1010 Wayside Lane
Scottsdale, AZ       85254
#
Prof. M. Elwood Mork
5235 Courtland Blvd., Apt. 23
Minneapolis, MN   55432
      .
      .
      .
```

Since a # at the beginning of a line constitutes a header, a # cannot appear as the first character of a line in an entry. One work-around for this problem is to put a blank in front of a # that otherwise would appear at the beginning of a line in an entry.

Within an entry, a line whose first character is a * is considered to be a comment and is not treated as significant text. For example, such comment lines are ignored when formatting mailing labels. Comment lines can be used for information like telephone numbers.

The # that starts a header line can be followed by one or more *designator* characters. Several of the programs can select only those entries with specific designators.

The choice of designator characters is up to the user. For example, #a might be used to designate active accounts, while #b might be used to designate bad addresses.

Organization of Entry Information

Some of the programs that process address lists expect the entries to be in a specific form. For example, the first line of an entry (after the header) is expected to be a name if the entry is an actual address.

Similarly, for addresses in the United States, the last line of an entry is expected to be the city, followed by a comma, followed by the postal-code abbreviation for the state, followed by one or more blanks, followed by the ZIP code. See the examples above.

For an address outside the United States, the last line is expected to consist only of the country name, in all uppercase letters.

Programs

The following programs are available for processing address lists:

adlcheck	Checks lists for bad data. Options include checking the state and ZIP code (U.S. only), country name, and for fitting in the confines of a standard one-up mailing label.
adlcount	Counts the number of labels in a list with optional restriction to entries with specified designators.
adlfiltr	Filters a list, outputting only those entries with specified designators.
adllist	Lists "fields" of address list entries, including addressee name, city, state, ZIP code, and country.
adlsort	Sorts address list entries by addressee name, ZIP code, or country.
labels	Produces one-up mailing labels for designated entries.

See the programs themselves for detailed documentation.