

How Do Experienced Information Lens Users Use Rules?

Wendy E. Mackay¹, Thomas W. Malone¹, Kevin Crowston¹,
Ramana Rao², David Rosenblitt¹, Stuart K. Card²

¹*Massachusetts Institute of Technology*

²*Xerox Palo Alto Research Center*

Abstract

The Information Lens provides electronic mail users with the ability to write rules that automatically sort, select, and filter their messages. This paper describes preliminary results from an eighteen-month investigation of the use of this system at a corporate test site. We report the experiences of 13 voluntary users who have each had at least three months experience with the most recent version of the system. We found that:

1. People without significant computer experience are able to create and use rules effectively.
2. Useful rules can be created based on the fields present in all messages (e.g., searching for distribution lists or one's own name in the address fields or for character strings in the subject field), even without any special message templates.
3. People use rules both to prioritize messages before reading them and to sort messages into folders for storage after reading them.
4. People use delete rules primarily to filter out messages from low-priority distribution lists, not to delete personal messages to themselves.

Keywords

Information Lens, rules, filtering, electronic mail.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

Introduction

The Information Lens is a prototype intelligent information sharing system designed to help users select, sort and prioritize their electronic mail (Malone et. al., 1987a, Malone et. al., 1987b). While many of these ideas are now being incorporated in other systems, (Borenstein & Thyberg, 1988, Gorry et. al., 1988) there have as yet been no systematic empirical studies of how people actually use such systems.

This paper describes the preliminary results of an eighteen-month study of the use of the Information Lens at a corporate test site. The full Information Lens provides a variety of features, including automatic rules for processing messages, semi-structured message types, and "anyone servers" for routing public messages to interested receivers. This paper concentrates only on the local rules created by experienced users of the system to process messages on their own workstations. Based on repeated interviews with users and on-line records of the rules users created, we have tried to determine how people with significant experience using Lens used rules.

We were especially interested in the following questions: Can non-programmers use Lens effectively? What kinds of rules prove useful after several months of experience with the system? When do people choose to run their rules? When do people use rules to delete messages automatically? What kinds of strategies do people use to manage their mail, and how do rules fit these strategies?

Creating Lens Rules

Information Lens users can create sets of IF-THEN rules to specify the actions they want performed on their messages and the conditions under which those actions should be taken. Common actions include moving the message into a folder, deleting the message, setting a user-defined characteristic (e.g., "Urgent") on the

message, and printing the message. Users create rules with a Rule Editor (below). In this example, the user has specified that messages addressed to the "Want-Ads" distribution list that also contain the subject "Car" are to be moved to the "Car" folder. Whenever a message meets these criteria, Lens moves the message to the "Car" folder.

Figure One

The screenshot shows a window titled "RULE EDITOR" with a menu bar containing "Done", "Apply", "Cancel", and "Reset". The window is divided into two main sections: "IF:" and "THEN:".

IF:

- Message Type: [icon] [text box]
- From: [text box]
- To: **Want-Ads**
- cc: [text box]
- To or cc: [text box]
- Subject: **Car**
- Text: [text box]
- Marked: [text box]
- [icon] [text box]
- [icon] [text box]

THEN:

- [icon] **MOVE TO: Car**
- [icon] **STOP**
- [icon] [text box]

Users specify when rules should be applied by grouping them into different rulesets. "New mail" rules are applied automatically when new messages are retrieved to the user's workstation, before the user sees the new messages. Some users have several new mail rulesets. For example, one set may be used daily and another may be used only after returning from a trip or a vacation. The remaining rulesets are applied at the user's discretion, on any folder. Most users skim messages in their inboxes before running any of these rulesets. (These "on demand" rulesets can also be run on other folders.)

Method

Our test site is a research laboratory (approximately 60 people) of a major American corporation. It was chosen because it uses the same hardware and software

environment as our prototype system and the members of the laboratory use electronic mail for a wide range of critical and non-critical organizational communication. The participants in our study included a spectrum of job categories, educational backgrounds, and amounts of computer experience.

This site has used electronic mail extensively for over a decade and a half. Essentially all employees rely on electronic mail for both formal and informal communication. Users receive an average of 30 to 35 electronic messages per day. Many of these messages come from a variety of distribution lists on both work and non-work topics. The existing mail system in the lab operates in a networked workstation environment that permits the use of separate windows for composing, reading and browsing messages or folders. Messages can be marked with system-defined characteristics, such as moved or deleted, or with characteristics specified by the user. Users can create and name as many folders as they like for storing messages.

The members of this laboratory represent lead users of information technology. Studying them should provide insights into the typical office environment five to ten years hence.

Software

The Information Lens provides additional features to the existing electronic mail system. The Lens software is written InterLisp-D and LOOPS and runs on Xerox 1100 series workstations. Version 1 was introduced to a small group of users in January of 1987. Based on feedback from these users, version 2 was created and distributed to a larger group of users in September, 1987. Version 3 was developed for a new version of the basic mail system and underlying LISP environment. It incorporates suggestions from version 2, but omits other features, including message types and the anyone server (see Malone et. al., 1987). Version 3 became generally available in January 1988, but participants could only use it after their workstations were upgraded to the new system software. Thus, start dates for version 3 range over a period of several months. Although we have not explicitly encouraged users to use version 3, it continues to be actively used and enhanced at the test site.

Participants

Members of the laboratory were told about the Information Lens and asked if they would be interested in using it. Because we were interested in active users of electronic mail, we solicited participation via electronic mail.

Over 25 people used one or more versions of the Information Lens. However, for this paper we are primarily interested in how experienced users create and use rules. Therefore, we present results from the 13 users who have used version 3 of Lens for at least three months. Some have used version 3 for as long as seven months and over half have extensive experience with earlier versions of Lens. The group includes a manager, a manager/secretary team, four computer programmers, one research scientist with formal training in computer science and six other research scientists (from a variety of physical and social sciences) without formal training in computer science. In order to help disguise the identities of our subjects in the discussion below, we randomly vary the sex of the subject in our descriptions.

Those omitted from this analysis have either stopped using Lens or have not yet accumulated three months of experience with version 3. The designer of version 3 was omitted because many of his rules were created to test the system and cannot be distinguished from his regular mail processing rules. The primary reason for discontinuing use of Lens was conversion to a workstation that does not support the Lens software.

Two people (one programmer and one manager) reported that they did not find Lens sufficiently useful to continue using it. The programmer reported being "overwhelmed" by mail in the first interview and had subscribed to a large number of distribution lists. By the third interview, she had removed herself from most distribution lists and reported few problems keeping up with her mail. She noted that she would like to use the Lens anyone server, when it became available. The manager was given the earliest version of Lens, which was slow and had numerous bugs. He used it for several days before giving up, and never decided to try the newer versions of Lens. (The rest of his group did, however, try the newer versions of Lens.)

Data Collection

Interviews were conducted prior to the study and at three month intervals during the course of the study. Each interview was scheduled for an hour in the participant's office. Participants were asked to estimate the daily numbers of messages sent and received, the number of mail folders, the size of the inbox and the number of distribution list subscriptions. These were checked against the actual numbers for the day and participants were asked if the day was typical.

Participants were also asked open-ended questions about major problems and successes with electronic mail. They used this an opportunity to describe their current communication patterns, successful mail management strategies, and problems that needed to be addressed. They also described their use of Lens, including a description of their rules and how they relate to general mail handling strategies. (Refer to Mackay (1988) for a detailed description of the interview format and the actual questions.)

On-line data was also automatically collected, including (approximately) weekly snapshots of the rulesets, the hierarchy of mail folders and the distribution lists to which each user belonged.

The quantitative data described below are derived from a snapshot of the rules and folders of the 13 most experienced Lens users. For each user, we chose the first day after the user had had three month's experience with version 3 to create a representative, but not exhaustive, sample of rules.

Results

Can non-programmers write rules?

All participants in the study, regardless of computer experience, were able to effectively write their own rules. All users created rules for themselves and none mentioned any difficulty in writing rules. One user, with no computer training, described his first experience with the rule editor as follows: "It's obvious. You just go into the boxes and type whatever you want." Another non-programmer commented: "I like the fact that one can begin very simply, e.g. just sorting according to one's own name in the to: or cc: fields and then, after seeing how that works, progressively elaborate." The authors of complex rules, defined as rules with more than one field or containing boolean logic within a field, were usually programmers. However, one of the secretaries also wrote a number of complex rules.

Users have the option of writing arbitrarily-complex rules in a LISP-like language, but our interviews indicate most choose to use the graphical rule editor provided with Lens. This result is consistent with the finding by Jeffries & Rosenberg (1987) that both programmers and non-programmers were able to specify mail filtering rules faster when using a form like those we used than when using a procedural language.

What kinds of rules do people write?

The 13 participants in the study created a total of 190 rules, ranging from 2 to 35 rules per person (approximately 15 each). We classified the rules along two dimensions: the field tested in the IF part of the rule and the action used in the THEN part (see Appendix 1). Within each field, we have also categorized the most common types of tests in that field (e.g., for a distribution list, for the user him or herself, or for some other person).

Over half (51%) of the rules involve the processing of distribution lists. All but two of the participants (85%) use at least one distribution list rule. Lens allows some people to stay on distribution lists from which they would otherwise remove themselves. It also allows them to ignore messages from distribution lists they are required to be on, but do not find useful.

The next most common kind of rule identifies messages with some arbitrary string in the RECIPIENTS field (33%). Note that string matches within a field, while not equivalent to natural language parsing, are still more effective than general matches within the body of the message. Knowing that a string refers to a person or a distribution list limits the range of possible matching errors.

Strings can identify sources of messages in different ways. For example, the string ".BITNET" identifies messages from outside the company, often professional colleagues and friends. Strings can also identify a collection of related distribution lists. For example, the string "AI" selects messages addressed to AI, AI-digest, and AI-interest (as well as, unfortunately, Sailing, etc).

Since the subject is less constrained than other fields, we expected it to be less useful in rules than other fields, which usually contain the names of known users or distribution lists. A surprisingly high percentage (29%) of the rules select items in the subject field. (Appendix 1 lists the categories of subject fields.)

Another important type of rule is that which contains a reference to the user. While they comprise only 13% of the total number of rules, 85% of the participants had at least one and many had two or three. (Note that for every reference to oneself, there are an arbitrarily large number of possible references to distribution lists, which explains the low percentage of this type of rule in the general pool of rules.)

It is tempting, but not accurate, to believe that users with

more rules find Lens more useful. In fact, some of the users who reported the most satisfaction with Lens had only two to four rules. Also, some users face mail management problems for which Lens is useful; others do not.

When do people run rules?

Lens was originally designed to only allow people to run rules prior to reading messages. However, Mackay (1988) found that some users preferred to sort messages after reading them. She called the first kind of users "prioritizers" and the second kind "archivers" to characterize the differences.

Lens version 3 allows users to either or both. We found that experienced users often have strong preferences for one or the other. Five people use "new mail" rulesets exclusively, three use "on demand" rulesets exclusively, and five use both kinds of rulesets. Somewhat more than half of the total rules are in "new mail" rulesets.

When do people use delete rules?

One of the most common questions about Lens is what rules people use to delete messages. Approximately one fifth (21%) of the total number of rules delete messages (Appendix 1). Of these, 74% (20/27) delete messages from low-priority distribution lists.

Only five delete rules (13%) refer to a person's name. In each case, the rules use multiple fields to further qualify the type of message. In two cases, the messages deleted are announcements of events sent by an outsider or part-time employee to the entire lab. The senders of these messages are misusing the system, but appear to be outside the usual social pressures that would normally prevent this behavior. Even though it is unlikely that the Lens user will want any particular messages from either sender, the rules are still qualified to filter out only messages about the unwanted subject. Personal messages from either person would not be deleted.

In general, deletion rules appear to be more complex than rules that move messages to folders. If we define complex rules as those with more than one field (counting "TO: or CC:" as a single field), then 60% of delete rules are complex, whereas only 32% of move rules are complex. One explanation for this is that deleting is more "dangerous" than moving because an error has greater consequences, and users are more likely to qualify them.

How do people manage their inboxes?

Rules reflect the strategies users have for managing their

inboxes. We found three primary strategies for handling mail in inboxes: *Keep it all, move unimportant messages, or move important messages.* The most interesting Lens rules occurred with the second and third strategies.

Move the unimportant messages. This strategy moves inessential messages out of the inbox and uses the inbox as a repository for unprocessed mail and things to do. Rules can help identify low priority messages. Several individuals reported that they moved messages from low priority distribution lists into folders only to discover that they never looked at them. Some decided this provides them with useful information about their priorities and helps them decide which distribution lists to delete. Others let the low priority messages pile up and avoid the clutter in their inboxes.

Another strategy is to use newmail rules to automatically delete unimportant messages. The deleted messages appear on the screen with a line drawn through them, so the user has a chance to see them before they are actually removed. ("I don't remove myself from mailing lists, I just automatically delete the message. I occasionally spot something in the deleted messages and take a look.") This person also pointed out that it is easier to write a rule than get off a distribution list. He changes his rules from delete to move, depending on his current work load and whether the topic is currently relevant. "It makes the cost of staying on a distribution list very low." One person uses a similar strategy for identifying potential rules for handling messages in his inbox. "If a rule doesn't fire on a set of messages in the inbox, I let them accumulate and use [those messages] to determine whether a special rule is needed."

The author of Lens version 3 took advantage of a convention in the mail system, in which RE: is automatically appended to the subject line when a user replies to a particular message. If a "boring" message arrives, e.g. **Subject: Tyrolia ski binding recall**, for which subsequent discussion is likely to be uninteresting, she can tell Lens to create a rule that automatically deletes messages with the subject: **Subject: RE: Tyrolia ski binding recall**. (This is an example of "rule-creation-by-example.") Only four participants use this feature, and all subscribed to an above-average number of distribution lists. ("This boring business is a real win!") Another user, who rarely reads distribution lists, tried it for a while and did not find it particularly useful.

Move the important messages. In this strategy, all important messages are moved to a high priority folder and the inbox stores unimportant items. Rules can help identify and move important messages. Five participants use this strategy and usually identify personally-addressed messages as important. All are willing to discipline themselves to read priority folders first.

One senior researcher stays off all voluntary distribution lists and relies on colleagues to provide him with technical information. Mail helps him maintain an active correspondence with colleagues around the world; these make up two-thirds of his messages. His rules correspond closely to his folders, identifying messages addressed to him personally and messages sent via "BITNET". He has very few rules, but they make a major difference in his use of electronic mail. "When I've been gone, Lens is fantastic. I was away for a week and it was a pleasure to have my mail sorted correctly."

Another user converted to this strategy after using Lens for many months. He was initially distrustful of Lens, and used it only to sort previously-skimmed messages in his inbox. He refused to use new mail rules. ("Nothing happens automatically - I look, then sort.") This user allowed his inbox to reach an unmanageable state with over 1100 messages. Note that, at a rate of handling 1 message per minute, this would take over 18 hours to process.) Rather than taking the time to process them, he collects "old" messages from the beginning of his inbox, moves them into a dated folder, and hopes there are no important messages within.

This person tried a new strategy, almost by accident. He decided to create two rules to handle mail while he was away. The first identifies all personally-addressed messages and places them into a "priority" folder. The second identifies all messages related to a conference he was running and sends them to the conference administrator. When he returned, he found that this strategy was "very, very useful" and continued using these two rules. He has stopped trying to keep on top of his inbox and instead, reads only the "priority" folder on a regular basis. He reads whatever is left in the inbox whenever he gets a chance and either files or deletes the messages manually. He said these two rules "changed my life" and no longer describes his mail as being out of control.

Conclusions

Electronic mail has created a set of new problems for its users. It drastically lowers the cost of communication, especially the cost of sending multiple copies of the same message, thus increasing the amount of mail each user receives. Yet the time available to process the increased number of messages is limited and users must decide how to trade off mail processing with other activities. Missing messages, such as meeting announcements or budget deadlines, may have dire consequences. If it takes an average of one minute to process a single message, then a manager who has allocated 30 minutes a day to read mail will only be able to process about 30 messages. Yet he or she may receive twice that many.

Information Lens rules provide one method for managing this flood of communication. This study explores what happens when this ability is given to real users actively trying to manage their mail.

We found that different mail processing strategies are reflected in the use of rules. Rules designed to fire before mail is read ("New mail" rulesets) help users identify important messages and act on them first. Some users find this the most useful aspect of Lens and rely on it to identify messages (usually personally addressed) that should be read immediately.

Rules designed to be fired after messages have been read

are also used to manage mail; both to help store messages in such a way that they can be easily retrieved and to remove messages that are no longer useful. Users who create such rules are more often interested in managing their ability to retrieve information.

References

Borenstein, N.S. & Thyberg, C.A. (September 1988). Cooperative Work in the Andrew Message System. *Proceedings on the Conference for Computer-Supported Cooperative Work*. Portland, Oregon.

Gorry, G.A., Burger, A.M., Chaney, R.J., Long, K.B., Tausk, C.M. . (September 1988). Computer Support for Biomedical Work Groups. *Proceedings on the Conference for Computer-Supported Cooperative Work*. Portland, Oregon.

Jeffries, R and Rosenberg, J. . (April 1987). Form-based and a Language-Based User Interface for Instructing a Mail Program. *CHI + GI Conference Proceedings*. Toronto, Canada.

Mackay, W.E. (October 1989). Diversity in The Use Of Electronic Mail: A Preliminary Inquiry. *ACM Transactions on Office Information Systems*, Vol. 6(4).

Malone, T.W., Grant, K.R., Lai, K.Y., Rao, R. & Rosenblitt, D.R. (1987). Semistructured Messages Are Surprisingly Useful for Computer-Supported Coordination. *ACM Transactions on Office Information Systems*, 5(2), 115-131.

Malone, T.W., Grant, K.R., Turbak, R.A., Brobst, S.A., & Cohen, M.D. (1987). Intelligent Information-Sharing Systems. *Communications of the ACM*, 30, 484-497.

Appendix 1

	ACTION			RULESET		TOTAL
	Move	Delete	Other	New Mail	On Demand	
Recipients	109 57%	27 14%	4 2%	79 42%	58 31%	137 72%
Distribution list	63 33%	20 11%	3 2%	45 24%	38 20%	83 44%
Other string	28 15%	4 2%	0 0%	18 9%	14 7%	32 17%
Self	12 6%	2 1%	1 1%	12 6%	3 2%	15 8%
Other person	6 3%	1 1%	0 0%	4 2%	3 2%	7 4%
From	29 15%	10 5%	0 0%	27 14%	12 6%	39 21%
Other string	10 5%	6 3%	0 0%	13 7%	3 2%	16 8%
Other person	10 5%	2 1%	0 0%	8 4%	4 2%	12 6%
Self	9 5%	2 1%	0 0%	6 3%	5 3%	11 6%
Subject	41 22%	14 7%	0 0%	38 20%	18 9%	56 29%
Project/system	11 6%	1 1%	0 0%	4 2%	8 4%	12 6%
Other string	13 7%	11 6%	0 0%	19 10%	5 3%	24 13%
Distribution list	12 6%	1 1%	0 0%	10 5%	3 2%	13 7%
Meeting/seminar	4 2%	1 1%	0 0%	4 2%	2 1%	6 3%
Phone message	1 1%	0 0%	0 0%	1 1%	0 0%	1 1%
Text or Body	2 1%	4 2%	1 1%	7 4%	0 0%	7 4%
Characteristic	10 5%	0 0%	4 2%	9 5%	3 2%	12 6%
Total = 190 Rules						