Words can’t describe it: streamlining PRECIS just for laughs!

Christine Jacobs and Clément Arsenault

When researchers at the Musée...pour rire, the international museum of humour in Montreal (Quebec, Canada), started to research film and video extracts for incorporation into the museum exhibits, it became obvious that a database with very specific subject access was necessary. A modified version of PRECIS, dubbed PRECIS-MO, was adopted. The adaptation and implementation are discussed.

The Musée... pour rire Documentation Centre

In April 1993, Montreal (Quebec, Canada) welcomed a new institution to its cultural horizon—Un Musée...pour rire, the international museum of humour, opened its doors to the public. Its mission is to celebrate, preserve, honour and spread the world’s laughter.

With 3000 square metres of exhibition space, the museum mounts bilingual (French- and English-language) shows of major proportions; however, the exhibits do not consist solely of the traditional presentations of artifacts. Humour is intangible, more of a virtual subject than an array of identifiable objects. Therefore, most of the exhibits incorporate video clips (excerpts of film, video and television productions), a more suitable medium for expressing the intangible.

When the research for clips for the Museum’s opening exhibits started, no structure was in place to ensure that the material gathered by the researchers would be preserved and made accessible for further use. They screened and selected as they went. If the researchers needed to find sequences again, they had to rely on their memories to do so. With no database in place, this resulted in a great deal of confusion, frustration and countless errors, as well as an enormous waste of time. It soon became obvious that a documentation centre was needed in order to store and organize all the material and information.

One of the Centre’s first activities was the creation of a video clip database. The researchers screen and catalogue television comedy series, specials and movies. As they are screening, they pick out sequences and themes that they feel are particularly well done, or pertinent in terms of subject, point of view, or method of presentation.

Because the Museum presents thematic exhibits, it became clear that good subject access to the video clips was a priority. Subject heading and keyword systems were considered. However, none gave the precision necessary for the level of subject access the researchers needed.

PRECIS, which is used by the National Film Board of Canada (NFB) for indexing films and videos, came to the Centre’s attention. The NFB has been using the system successfully since 1978, to provide access that is specific in terms of subjects, aspects and forms.

PRECIS

PRECIS is a complex string indexing system in which the indexer uses approximately 40 codes and operators to create strings of terms. In print products the strings are shunted (permuted) so that the various terms become access points, always within the context of the rest of the string. It is a very expressive system that permits quite sophisticated formulations, but it also demands a high level of expertise. We will not attempt to describe the system here, but refer readers interested in a full explanation to Austin, Dykstra and Richmond.¹

The drawbacks of PRECIS—complexity and cost—were stumbling blocks. The researchers are specialists in the field of humour, not documentation, cataloguing or indexing. They were the obvious ones to do the...
STREAMLINING PRECIS

indexing, since they could identify the 'funny bits' and other aspects of value to them in their work, but they had neither the time nor the inclination to learn the complete PRECIS system.

The cost of using PRECIS through a central utility was prohibitive. An aspect of PRECIS that is discussed rarely in the literature is that its consistent application is a time-consuming (and therefore money-consuming) process. The PC version of PRECIS used at the NFB was not yet on the market, and although it would have solved some of the cost problems in terms of straight dollars and in terms of time, it would not have reduced the complexity, since it is a fairly complete implementation of Derek Austin's PRECIS.

At the NFB some consideration had been given to using a modified version of PRECIS that would have been suitable for an online system (no shunting of terms, and no typographical concerns). With an ongoing commitment to print products, however, this idea was shelved. When Christine Jacobs was approached by the Centre as an indexing specialist with PRECIS experience, she immediately thought of the modifications considered at that time, and proceeded to simplify them even further.

PRECIS-MO Implementation

The aim was to produce a bare-bones version of PRECIS that would still provide enough grammatical framework to prompt the indexers in their analysis and to promote consistency.

The decision was made to store the PRECIS strings as a field in the catalogue record in a FoxPro database. The coding would be used only at the manual stage to construct the strings: the 'top down' entry would then be stored in the PRECIS field (searchable with Boolean operators). This liberated the indexers from the manipulation coding.

In the end we were left with six of the primary operators (0, 1, 2, 3, 4, 6), and two of the secondary operators (p, q); dates were enclosed in parentheses and connectives were enclosed in square brackets. See Figure 1 for the list of operators and their meanings.

The downward substitute convention was retained, to improve readability of entries where necessary. This involves the substitution of a natural-language phrase for terms that link awkwardly, or form subsets. An effort is made to retain the terms 'as is,' but it may be necessary to change the form of one or more terms. To compensate for this, it is important to use truncation techniques when searching. The code indicating a downward substitute, a down arrow (↓), is followed by a number indicating the number of lines for which it is substituting. See Example 1 in Figure 2.

This is a very slim version of PRECIS, but reasonably faithful to the original grammar! In keeping with the spirit of PRECIS, natural, relatively informal terms are preferred. Since so much was modified, we felt that we should provide a name for it that would make this clear. PRECIS-MO works well in French and English (PRECIS-modifié/modified), is easy to say, and hints at the Montreal origin.

Figure 1. Operators and codes used in PRECIS-MO

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Place / environment</td>
</tr>
<tr>
<td>1</td>
<td>Key system; object of an action</td>
</tr>
<tr>
<td>2</td>
<td>Action</td>
</tr>
<tr>
<td>3</td>
<td>Agent; doer of an action</td>
</tr>
<tr>
<td>4</td>
<td>Point-of-view</td>
</tr>
<tr>
<td>6</td>
<td>Form</td>
</tr>
<tr>
<td>p</td>
<td>whole/part relationship with preceding term</td>
</tr>
<tr>
<td>q</td>
<td>quasi-generic relationship with preceding term</td>
</tr>
<tr>
<td>(</td>
<td>encloses dates</td>
</tr>
<tr>
<td>]</td>
<td>encloses connectives such as and, by</td>
</tr>
<tr>
<td>↓</td>
<td>downward substitute</td>
</tr>
</tbody>
</table>

All terms except those enclosed in parentheses and square brackets should be controlled in the thesaurus. Forms of terms in a substitute phrase may vary from the form of the controlled terms.

What wasn't included

Operator 5 (for indicating selected instances, such as study regions) was not included because we felt there would be little call for it. On the odd occasion when it might be useful, the analysis could be handled with multiple strings. However, it would always be possible to include it at a later date, if the requirements arose.

Operators s, t, and u, which indicate types of actions and cause/effect relationships, were eliminated. It was felt that distinguishing between types of actions, necessary if the strings are being shunted, was too sophisticated for our needs. Although interesting to do, it would have added to the complexity of the researchers' task without affecting the form of the entries. Operator 2 is used for all types of actions.

Rather than appearing on separate lines linked by secondary operator f or g, coordinate concepts are linked by [and] on the same line.

None of the various codes necessary for print products was included. The downward connective code Sv was replaced with brackets that enclose the connecting term. The Sd date code was replaced with parentheses.

What about the thesaurus?

It would have been nice to have a thesaurus that interfaced smoothly with the dictionary of entries. However, reality being what it is, we have had to settle for the tried-and-true method of index cards. When money and software permit, this will be automated.

In fact, the thesaurus problem has been somewhat of a challenge. English-language material is indexed in
STREAMLINING PRECIS

Examples:
1 bosses
2 contract negotiations [with]
142 television talk show sidekicks
1 television talk shows
6 situation comedies

Entry: Bosses. Contract negotiations with television
talk show* sidekicks—Situation comedies.

(∗Note that search terms must be truncated to allow
for terms in the downward substitutes)

0 États-Unis
1 artistes [et] chanteurs
2 fabrication
4 parodie

Fabrication—Parodie.

0 Canada
1 elections (1993)
2 role [of]
3 women
6 skits

Entry: Canada. Elections (1993). Role of women:
lawyers—Skits

Figure 2. Encoded PRECIS-MO terms and the resulting entries

From Log Driver's Waltz, directed by John Weldon. Reproduced with the permission of The National
Film Board of Canada. Indexed under the terms:
Log Driver's Waltz:
Logging industry. Log drivers. Dancing
abilities—Songs—Animated films

English and French-language material in French. So
we are working with two base thesauri. Since the work
involved in producing bilingual thesauri is immense,
these are maintained separately; no effort is made to
coordinate the terms or the hierarchies between the
general French and English word pools.

However, it was felt that one of the long-term goals
of the Centre was to produce a bilingual Humour
Thesaurus. To accommodate this requirement we have
divided each thesaurus into two. We are, effectively,
working with four thesauri. The Humour Thesauri
essentially comprise the terms used with operators 4
and 6—the point-of-view and form terms. These sub-
sets are being developed bilingually, and more effort is
put into the vocabulary. Unfortunately, humour terms
are very fluid and varied, so it may be a long time
before there is agreement about what means what!

The researchers were given basic training in the-
saurus theory, and are responsible for using the the-
sauri and creating new terms. However, the thesauri
are managed by the Centre's chief librarian and his
assistant.

Problems in implementation

There have, of course, been some difficulties in the
implementation. Even streamlined, PRECIS is not a
simple system to teach, learn or use.

Each researcher has a personal style of indexing and
of deciding what is funny. Although they have been
given training in indexing, they are researchers first
and indexers/cataloguers second, so it falls to the librarian to iron out the wrinkles and to balance disparities where possible.

The relationships expressed in comedy are often complex, and even with the PRECIS-MO framework, the researchers find it difficult to analyze events into one string. To get around this, two or more strings are often used to express concepts that theoretically should be expressed in one.

Staff turnover among the researchers is a concern since training is time-consuming. PRECIS is a system that requires a good grasp of the grammatical and theoretical hows and whys, if it is going to work well.

Although it was expected that thesaurus maintenance would be labour intensive, it is even more so than anticipated. As a result, the control of the general word pools is fairly light, and the effort is put into maintaining the Humour Thesauri. Proper maintenance would require the hiring of a permanent part-time person, which is not feasible at the present time.

Are we happy with it?

The PRECIS structure has proved useful in teaching subject analysis. The researchers often knew when something was funny, and what made it funny, but not necessarily how to describe what they saw. The basic questions—did anything happen? If yes, to whom or what did it happen? Who, or what did it? Where did it happen?—coupled with the prompts inherent in the operators, provide them with a framework through which to view their material, focusing their attention on the relationships and events.

Although still in its early days, the system is working well. It definitely provides the level of specificity needed by the researchers. The PRECIS-MO framework ensures good coverage of all aspects of the subject(s), and the subject analysis is as precise as is necessary.

The researchers are also happy with the flexibility of PRECIS-MO. Although the terms are controlled in the thesauri, they are free to add to them, using as formal or as informal words as are necessary to describe the (sometimes bizarre) material. Whether a variety comedy special on the problems of adolescence (serious subjects like suicide and AIDS treated in a series of skits and stand-up routines), an episode of Murphy Brown, or an interview with a comedian, the Indexer Vol. 19 No. 2 October 1994

From The Awful Fate of Melpomenus Jones, directed by Gerald Potterton. Reproduced with the permission of the National Film Board of Canada. Indexed under the terms:
The Awful Fate of Melpomenus Jones:
1. Young clergymen. Politeness and shyness—Animated films
2. Goodbyes and leave—takings—Animated films
3. Leacock, Stephen. Short stories—Animated films

The Indexer Vol. 19 No. 2 October 1994
STREAMLINING PRECIS

material has been amenable to the type of analysis PRECIS requires.

Having one centralized system rather than a variety of systems run by the individual researchers has been a major benefit, justifying the efforts put into the creation of the system.

References
2. Thank you to John Leide (McGill University) for his suggestion.
3. Thank you to my PRECIS colleagues at the NFB, Sheila Kunst and Christiane Talbot, for their encouragement and advice—C.J.

Art form recognized

We reproduce below the introductory note for the index to Sky watcher’s handbook: the expert reference source for the amateur astronomer, edited by James Muirden, published by W.H. Freeman. The note is written by James Muirden, and quoted here by permission.

While indexing this book I have discovered that although indexing is supposed to be a science, it is really an art. Like all art forms, the perfect index exists only in the mind. The indexer does not know the predilections, needs, or expectations of the reader. In addition, the best index is only an approximation to the content of the book it is supposed to unfold.

Realising this has made me feel confident enough to state my aims, thereby also revealing my shortcomings. They have been:
1. To try to short-cut the reader’s journey by supplying ample cross-references.
2. To give the reader a comprehensive idea of the range of topics covered in each chapter.
3. To avoid irritating strings of blank page numbers after a single entry.
4. To give particular guidance to entries that may not be exactly where the reader expects to find them.

Expanding on (4), as an example, the separate features in a section dealing exclusively with Martian features have not been separately indexed, but features mentioned elsewhere in the chapter have (or should have been) indexed.

I have also tended to include the subjects of illustrations (not, on the whole, tables) in the index.

(Note the use of ‘should’, ‘tended’, and ‘on the whole’, and remember that I am talking about an art form.)

All these precepts, and others, have surely been followed before, and more consistently. But in one respect, this index breaks new ground: it contains the answer to a puzzle. Therefore, to avoid spoiling the challenge, the reader is strongly advised to read the book from cover to cover and decide what the puzzle is, before attempting to use the index.

Telling authors how

Much advice has been offered worldwide lately to authors about indexes. We reproduced in our last issue ‘The author and the index’, the relevant sections from Nancy Mulvany’s volume, Indexing books, newly published by the University of Chicago Press.

In Australia, in the Campus Review of Macquarie University, Sydney, for 5–11 May 1994, Robin Derricourt advises academic authors on obtaining or providing indexes to their books: ‘Index the readers’ key’. Describing the index as ‘the menu, the store directory, the telephone directory for scholarly or educational purposes’, he states that a good index should start from the diverse needs of the readers: ‘the marginal for one reader is the core issue for another’. Soundly, he advises ‘A computer may help to sort, list and alphabetize entries but for a scholarly book it does not create an index of the quality and form that readers need’.

In the UK, in The Author, Autumn 1994, ‘The tail and the dog: a book indexer speaks’, Douglas Matthews wrote for authors of how he works as an indexer, still holding ‘that the author is the best person to compile an index (assuming him or her to be technically competent)’. He attributes a further role to the indexer, that of test-driver, pointing out mistakes, ambiguities, inconsistencies, even stylistic infelicities. Precepts from the northern hemisphere accord with those from the the south: ‘Computers and word processors help, but there can be no shortcutting that basic, dogged, analytical reading of the whole work’.

Christine Jacobs is a freelance librarian specializing in PRECIS; Clément Arsenault is former Librarian of the Documentation Centre for the Musée . . . pour rire.