City Research Online

City St George's, University of London

Citation: MacFarlane, A. (2011). Using Multiple Choice Questions to Assist Learning for Information Retrieval. In: Efthimiadis, E. N., FernandezLuna, J. M., Heute, J. F. & MacFarlane, A. (Eds.), Teaching and Learning in Information Retrieval. Information Retrieval Series, 31. (pp. 107-121). Berlin: Springer-Verlag. ISBN 9783642225109 doi: 10.1007/978-3-642-22511-6_8

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/4490/

Link to published version: https://doi.org/10.1007/978-3-642-22511-6_8

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Using Multiple Choice Questions to Assist Learning for Information Retrieval

Andrew MacFarlane

Department of Information Science, City University London, Northampton Square, London EC1V 0HB

Abstract A key issue in teaching and learning in information retrieval – particularly for library and information science students – is the gap in prior knowledge compared with the need for mathematics to conduct and evaluate searches. In this chapter, we examine the use of online Multiple Choice Questions to support these type of students, and narrow this gap between experience and knowledge. We provide some background in terms of related work and the use of MCQ's for assessment. The key areas of search which can be supported by this form of assessment are defined, and these are used to outline a proposed strategy for defining a series of questions to support learning.

1. Introduction

In MacFarlane (2007) a number of different challenges in supporting postgraduate library and information science students was outlined. The challenge has come about for two main reasons; there is evidence of a decline in mathematical skills for student entering tertiary education in the U.K (Croft, 2002) and some key skills such as calculus are no longer at GCSE level (Appleby and Cox, 2002) in the U.K.; many of the students who entered postgraduate LIS courses have a first degree in Arts and Humanities, and have very little exposure to relevant mathematical concepts for search. A number of different solutions were put forward to resolve this problem, including the use of online Multiple Choice Questions to support the students in their leaning. In this chapter we address this issue by examining the parts of information retrieval (search and evaluation) to which MCQ's can most usefully be applied – namely discrete mathematics and numeracy. In section 2 we describe related work. An outline of using MCQ's for assessment is provided in section 3, after which the areas of search which can be supported (and which relate to the mathematics part of the syllabus) are outlined in section 5. Given this background a strategy for implementing MCQ's in search is put forward. A summary and conclusions is provided at the end.

2. Related work

The various pedagogic challenges in teaching and learning search with Information Science students was first outlined in MacFarlane (2007), which focused on three main areas of mathematic which require support in some way to assist in the understanding of search – numeracy, discrete mathematics and probability/statistics. In this light the issue of diagnostic tests, delivery of material and summative assessment was addressed. This lead to further work on using a tutorial style of delivery to support students in MacFarlane (2009), which demonstrated that a pro-active style of teaching has a positive effect on the teaching of mathematics for search. A large scale review of teaching and learning in IR (Fernandez-Luna et al, 2009) indicated that there has been some work on using online tests in assessment e.g. Sacchanand and Jaroenpuntark (2006) which was focused on using a web-based training package in distance learning, however it the research focused on search in general and used other forms of online interactivity such as multimedia to achieve its aims. There is a clear need therefore to investigate the use of MCQ's to support student learning in search, in particular for understanding mathematics.

3. Using MCQ's for assessment

Multiple choice Questions (MCQ's) are one of a number of different types of objective assessments, which can either be used for formative and summative assessment (Higgins and Tatham, 2008). Each question is made up of a stem (the text of the question) with a number of options as the answer to the question; the key – the correct answer, and a number of distracters which are incorrect answers (McKenna and Bull, 1999). Care needs to be taken when designing questions (Higgins and Tatham, 2003) to ensure that no clues are given away as to the answer, encourage surface learning, contain obviously incorrect answers etc.

MCQs are best used when there are clear objective right and wrong answers to a problem. As stated in MacFarlane (2007), much of learning IR cannot be supported this way as important elements contain subjective aspects e.g. information needs, relevance assessment. However the underlying mathematics used to support search (such as Boolean logic, numeracy) do allow objective questions to be posed, and is therefore possible to use them in a supporting role. It is our contention that using MCQ's for summative assessment (MacFarlane, 2007) is not appropriate, and we therefore focus on formative assessment. Using MCQ's to support teaching and learning mathematics for search is useful in the formative context - they can either be used for diagnostic tests in order to ascertain the prior experience of the cohort, or just in a supporting role for those who feel the need for extra tuition.

4. Areas of leaning in IR supportable by MCQ's

We now turn to the specific areas of search in IR, which are mathematically focused and to which MCQ's can be applied to objectively assist teaching and

learning. We identify six different areas of the search process which require knowledge of mathematics to some degree:

- Boolean logic: operational IR systems only use three core Boolean operators: AND (set intersection), OR (set union) and AND NOT (set difference). These are used to connect terms together to create quite complex queries, expanding and narrowing search as necessary.
- Extended Boolean logic: operational IR systems also provide extensions to Boolean logic, such a proximity operators and truncation or wildcards. Proximity operators come in a number of different forms (adjacency, within text block e.g. sentence, paragraph, within a range of words or near) and are use to narrow down the search further – they are a special case of AND (set intersection). Truncation operators are used to specify variations of a useful word e.g. break* would retrieval break, breaks, breakers, breaking etc and are used to expand the search – they are a special case of OR (set union). Truncation operators are generally postfix, but infix or prefix operators can be supported.
- Ranking and ordering documents: Statistics such as term frequency, inverse document frequency and document length are used to building ranking models for ordering documents. The cohort only needs a high level understanding of these concepts, and their effect on search.
- System syntax: operational search systems have their own query syntax and form, which varies from system to system particularly for Extended Boolean operators e.g. within, near proximity operators, *,? Truncation operators. Examples of variations can be found in Dialog (Dialog, 2010) and Factiva (Factiva, 2001).
- Boolean search strategies: there are a number of command line systems which allow quite complex search strategies e.g. successive fractions and building blocks. These require an understanding of set theory and merging sets through both Boolean and Extended Boolean operators (see above).
- Evaluation measures: the core precision and recall measures (Cleverdon, 1967) must be understood, plus important variations such as mean average precision (TREC, n.d.). This assumes binary relevance. Graded relevance assessment measures such as the Discounted Cumulative Gain (DGC) measure can also be tackled (Kekäläinen and Järvelin, 2002).

MacFarlane (2007) applied the 'Mathematical Assessment Task Hierarchy' (MATH) taxonomy defined by Smith et al (1996), to the building blocks for mathematics required for information retrieval (see table 1):

Group A	Group B	Group C
Numeracy. Set theory. Transformation rules e.g. commutativity, associativity.Statistics and probability.	Forming Boolean queries analysed from a users information need.	Search strategies (different uses of Boolean and Adjacency operators and terms). Evaluation of results.

TABLE 1 – BUILDING BLOCKS FOR MATHEMATICS REQUIRED FOR INFORMATION RETRIEVAL (MACFARLANE, 2007)

In this table dependencies flow from left to right e.g. Group B knowledge is dependent on Group A, Group C is dependent on Group B. The logical flow of questions therefore is to address Group A issues followed by Group B then Group C. We apply this framework to the six elements of information retrieval which require mathematical knowledge identified above:

- Group A: Boolean and Extended Boolean Logic. Ranking and Ordering of documents. Precision and Recall measures.
- Group B: Systems search syntax (implemented Boolean and Extended Boolean logic). Applying precision and recall measures.
- Group C: Boolean Search strategies (using systems search syntax). Evaluation of precision and recall measures.

This framework provides us with the strategy we put forward to build a set of MCQ's to support student learning of mathematics for IR.

5. A proposed strategy for implementing MCQ's to support IR learning

In this section we proceed through the taxonomy from group A to group C issues. The example questions provided are based on the authors personal experience of teaching an IR module to LIS students. We provide different views of the mathematics material to be delivered including some which are more mathematical focused, and some which are put in the context of searching using examples of highlight text. Each example given takes the following form:

- Stem: this is the text of the question for the student to answer
- Options: different answers to the question, two types:
    - Key: the correct answer to the question.
    - Distracters: incorrect answers which are plausible and are typical wrong answers for the given question.

The advice given in McKenna and Bull (1999) is used to generate the example questions, readers should refer to this paper on writing pedagogically useful MCQ's. The structure is based directly on the taxonomy derived in MacFarlane (2007) and expanded here. In sections 5.1 to 5.4 we address query based issues.

Sections 5.1 and 5.2 focuses on Group A issues in search such as Boolean Logic, Extended Boolean Logic and the ranking and ordering of documents. The information in section 5.3 is then developed further in the context of Group B issues, by addressing the issue of system search syntax. Group C issues are then tackled in section 5.4 by developing search strategies using the search syntax used in section 5.3. Section 5.5 is focused entirely on evaluation issues, and proceeds from group A to C using standard evaluation measures in IR.

5.1 Boolean and extended Boolean operator Querying

The first area to address is Boolean logic, firstly via set theory and then using examples in context. All the simple examples given here focus on AND, but variations would be provided for the OR, AND NOT operators. An understanding of Boolean operators would be tackled first:

- Stem: What is the result set for {1,3,5,6,8,9} AND {1,2,4,6,7,8}
- Key: {1,6,8} – correct! Elements must be in both sets
- Distracters:
    - {1,2,3,4,5,6,7,8,9} - No: this is a Boolean OR (set union), A OR B. Recall that the real world use of AND is different from a formal setting such as this.
    - {3,5,9} – No: this is a Boolean NOT (set difference), A NOT B.
    - {2,4,7} – No: this is a Boolean NOT (set difference), B NOT A.

Figure 1: Example set theoretic question (AND)

Variations would be provided for OR and NOT operators. The stem can utilize Venn diagrams to help the student find the answer (see above). Questions on Boolean transformation rules could also be addressed here e.g. assocativity, communtivity etc. This question set would be followed by MCQ's focused on highlighted text, see figure 2:

- Highlighted Text: "**Information retrieval** is the art of the possible. Most people – no matter what society they live in – are involved in some kind of information seeking. For example the **Bushmen** of the Kala-

hari need information on waterholes in the desert. Although they live in a non-technical society, their thirst brings about an information need, in this case to find a waterhole."

- Stem: Which of the following queries would retrieve this document?
- Key: information AND retrieval – Correct: this document contains both terms in the query.
- Distracters:
    - information AND extraction – No: the text does not contain the right hand term 'extraction'.
    - sahara AND bushmen – No: the text does not contain the left hand term 'sahara'
    - technological AND seekers – No: the text does not contain either query terms.

Figure 2: Example highlighted text question (AND)

Care needs to be taken when mixing AND and OR queries, otherwise two right answers can be provided, complicating the option design e.g. information OR retrieval would also be a correct answer for the example given in figure 2. Proximity operators, a special case of AND can also be tackled (perhaps directly after the questions on the AND operator to help the student distinguish between the two types of query). Proximity operators are tackled next, see figure 3:

- Highlighted Text: "**Information retrieval** is the art of the possible. Most people – no matter what society they live in – are involved in some kind of **information seeking**. For example the **Bushmen** of the **Kalahari** need information on waterholes in the desert. Although they live in a non-technical society, their thirst brings about an information need, in this case to find a waterhole."
- Stem: Assuming a proximity operator named ADJ, which retrieves documents with two terms directly next to each other in the specified order - which of the following queries would retrieve this document?
- Key: information ADJ retrieval – Correct: this document contains both terms in the query and they are next to each other in the specified order.
- Distracters:
    - seeking ADJ information – No: the text does contain both terms and they are directly next to each other, but not in the specified order of required by the query.
    - Bushmen ADJ Kalahari – No: the two terms are in the correct order, but the words 'of' and 'the' are between them.
    - information ADJ seekers – No: the text does not contain term 'seekers'.

Figure 3: Example highlighted text question (ADJ)

Variations for other types of proximity operator could also be provided, e.g. same block of text, with a specified distance etc. Truncation operators, as a special case of OR can be tackled next, see figure 4:

- Highlighted Text: "Information retrieval is the **art** of the possible. Most people – no matter what society they live in – are involved in some kind of information **seeking**. For example the **Bushmen** of the Kalahari need information on waterholes in the desert. Although they live in a non- technical society, their thirst brings about an information need, in this case to find a waterhole"
- Stem: Assuming a postfix only wildcard operator *, which truncated a query term and selects a variety of documents with the [stem]*, which of the following queries would retrieve this document?
- Key: seek* - Correct: the text contains seeking which contains the stem 'seek' with 'ing' picked up by the wildcard operator
- Distracters:
    - Seek*ng – No: this is an infix operation, so although the term 'seeking' is in the document, a postfix only operator would not retrieve this text.
    - *men– No: this is a prefix operation, so although the term 'bushmen' is in the document, a postfix only operator would not retrieve this text.
    - arts* – No: the term 'art' is in the text, but the character 's' after 'art' would mean that the text would not match with this postfix only example.

Figure 4: Example highlighted text question (Wildcard)

All operators presented so far can then be used in a series of questions with much more complex Boolean expressions in them, see figure 5:

- Highlighted Text: "**Information retrieval** is the art of the possible. Most people – no matter what society they live in – are involved in some kind of **information seeking**. For example the **Bushmen** of the **Kalahari** need information on waterholes in the desert. Although they live in a non- technical society, their thirst brings about an information need, in this case to find a waterhole"

- Stem: Which of the following Boolean expressions would retrieve the above text?
- Key: (information ADJ retriev*) AND (information ADJ seek*) - Correct: all proximity operators in combination with wildcards and the connecting AND operator will retrieve this text.

- Distracters:
    - (Bush* ADJ Kalahari) AND (information ADJ seek*) – No: all the search terms are valid, but 'of the' is between the terms 'Bushmen' and 'Kalahari', so the left hand adjacency operators would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.
    - (information ADJ retriev*) AND (information ADJ science) – No: the text does not contain the term 'science' therefore the right hand adjacency operator would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.
    - (information ADJ retrieval) AND (information ADJ seeker*) – No: the right hand expression contains a wildcard operator 'seekers*' which would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.

Figure 5: Example highlighted text question (Boolean expression)

Common errors and misunderstanding on Boolean expressions must be addressed. The reader experienced in Boolean logic may regard the question design to be at somewhat a low level, but should not the remedial nature of this material at the Group A stage, required for this cohort before they tackle problems at the Group B level. Questions on Boolean transformation rules can also be provided using Boolean expressions to augment the set theoretic questions.

5.2 Ranking techniques

The first element to tackle when looking at ranking techniques, is to get the student to understand the concept of ranking, before tackling the concept of term weighting which drives the ranking process. This is because term weighting is somewhat involved and requires knowledge of difference statistical elements, while ranking (when weighting is complete) is a straightforward process. An example of a ranking question is given in figure 6.

	Doc 1	Doc 2	Doc 3	Doc 4
Engine	1	18	15	10
Repair	19	0	15	11
Services	4	10	1	33
Total				

- Stem: Given the table above showing the search term and document weight pairs, what is the order of ranking for the documents
- Key: 1st Doc 4, 2nd Doc 3,3rd Doc 2, 4th Doc 1 – Correct: the accumulated rates Doc 1(54), Doc 3 (31), Doc 2 (28), Doc (24)
- Distracters:
    - 1st Engine (44), 2nd Repair (45), 3rd Services (48) – No: this is wrong on two counts, we are concerned with the documents not term ordering, and the ordering is incorrect.
    - 1st Services (48), 2nd Repair (45), 3rd Engine (44) – No: Although the ordering is correct, , we are concerned with the documents not term ordering.
    - 1st Doc 1, 2nd Doc 2, 3rd Doc 4, 4th Doc4 – No: the ordering is inverse to what it should be i.e. Doc 1(54), Doc 3 (31), Doc 2 (28), Doc (24)

Figure 6 – Example ranking question

Several variations of this could be provided. The student is then in a position to tackle statistic information associated with each weight. We tackle Inverse Document Frequency (figure 7), Term Frequency (figure 8) and Document Length (figure 9). An appropriate form for this type of material is the presentation of true/false statements for the students to examine, selecting the correct number of true answers for a given set.

- Stem: Recall the definition of Inverse Document frequency (IDF) in the notes. How many of the following statements are true?
    1. IDF of 26 is more indicative of search usefulness than a IDF of 303.
    2. IDF of 567 is more indicative of search usefulness than a IDF of 303.
    3. IDF of 303 is less indicative of search usefulness than a IDF of 250.
    4. IDF of 304 is less indicative of search usefulness than a IDF of 303.
- Key: Two are correct, statements 1 and 3 – the smallest figure in IDF is regarded as more likely to be useful in a search context.
- Distracters:
    - Three – No: Two statements are incorrect, 2 and 4 where higher IDF's are said to be more indicative of search usefulness, the opposite is the case – the smallest figure in IDF is regarded as more likely to be useful in a search context.
    - One – No: Two are correct, statements 1 and 3 – the smallest figure in IDF is regarded as more likely to be useful in a search context.

- Four – No: Two statements are incorrect, 2 and 4 where higher IDF's are said to be more indicative of search usefulness, the opposite is the case – the smallest figure in IDF is regarded as more likely to be useful in a search context

Figure 7 – Example Inverse Document Frequency question

- Stem: Recall the definition of Term frequency (TF) in the notes. How many of the following statements are true?
    1. TF of 26 is more indicative of search usefulness than a TF of 3.
    2. TF of 56 is more indicative of search usefulness than a TF of 3000.
    3. TF of 38 is less indicative of search usefulness than a TF of 25.
    4. TF of 38 is less indicative of search usefulness than a TF of 45.
- Key: Two – Correct: statements 1 and 4 are correct. The largest figure in TF is regarded as more likely to be useful in a search context.
- Distracters:
    - One – No: Statements 1 and 4 are correct. The largest figure in TF is regarded as more likely to be useful in a search context.
    - Three – No: Two statements are incorrect, 2 and 3 where lower TF's are said to be more indicative of search usefulness, the opposite is the case – the larger figure in TF is regarded as more likely to be useful in a search context.
    - Four – No: Two statements are incorrect, 2 and 3 where lower TF's are said to be more indicative of search usefulness, the opposite is the case – the larger figure in TF is regarded as more likely to be useful in a search context.

Figure 8 – Example Term Frequency question

The concepts of TF is used to understand Document Length (DL), therefore this can be tackled in the question set next.

- Stem: Recall the definition of Document Length (DL) in the notes. Assuming a term occurs 10 times in a document, how many of the following statements are true:
    1. DL of 1000 is more indicative of search usefulness than a DL of 100.
    2. DL of 349 is more indicative of search usefulness than a DL of 1011.

3. DL of 460 is less indicative of search usefulness than a DL of 278.
4. DL of 318 is less indicative of search usefulness than a DL of 996.

- Key: Two are correct, statements 2 and 3 – the smallest figure in DL (or the shortest document) is regarded as more likely to be useful in a search context.
- Distracters:
  - One – No: Two are correct, statements 2 and 3 – the smallest figure in DL (or the shortest document) is regarded as more likely to be useful in a search context.
  - Two – No: Two statements are incorrect, 1 and 4 where higher DL's are said to be more indicative of search usefulness, the opposite is the case – the smallest figure in DL (or the shortest document) is regarded as more likely to be useful in a search context.
  - Three – No: Two statements are incorrect, 1 and 4 where higher DL's are said to be more indicative of search usefulness, the opposite is the case – the smallest figure in DL (or the shortest document) is regarded as more likely to be useful in a search context.

Figure 9 – Example Document Length question

Variations of each of these questions could be provided for sets which have 0 to 4 correct answers in them. A set of questions on how weighting relates to ranking could also be asked.

5.3 Query Syntax

In our examples we assume Dialog syntax (Dialog, 2010). We present examples of correct and incorrect syntax and/or use, and then take the examples from section 5.1, asking the same kind of questions using the appropriate Dialog form. Figure 10 show an example of correct and incorrect syntax – this lends itself to the choice of true/false from a set of given statements:

- Stem: Assuming the Dialog search syntax, how many of the following search statements are correct?
  1. SEARCH information
  2. S retrieval
  3. C information AND retrieval
  4. SELECT extraction
- Key: Two: Correct - statements 2 and 4 are valid Dialog search statements, statement 1 is wrong as dialog does not support the command

'search', statement 3 is wrong as you can only apply combine to sets already generated by select statements.

- Distracters:
    - – No: Incorrect - statements 2 and 4 are valid Dialog search statements, statement 1 is wrong as dialog does not support the command 'search', statement 3 is wrong as you can only apply combine to sets already generated by select statements.
    - – No: Incorrect - statements 2 and 4 are valid Dialog search statements, statement 1 is wrong as dialog does not support the command 'search', statement 3 is wrong as you can only apply combine to sets already generated by select statements
    - – No: Incorrect - statements 2 and 4 are valid Dialog search statements, statement 1 is wrong as dialog does not support the command 'search', statement 3 is wrong as you can only apply combine to sets already generated by select statements

Figure 10 – Example Dialog Syntax question

Further questions on sessions in Dialog could also be asked to address the confusion students have in building a search statement with the service. Variations could include testing any confusion between systems when using different online systems e.g. the truncation operator ? as against *. Building on this we then test the students with valid Dialog forms, but testing Boolean logic again in this context – see figure 11 (modified version of figure 5).

- Highlighted Text: "**Information retrieval** is the art of the possible. Most people – no matter what society they live in – are involved in some kind of **information seeking**. For example the **Bushmen** of the **Kalahari** need information on waterholes in the desert. Although they live in a non- technical society, their thirst brings about an information need, in this case to find a waterhole"

- Stem: Which of the following Boolean expressions would retrieve the above text?
- Key: SELECT (information(w)retriev?) AND (information(w)seek?) - Correct: all proximity operators in combination with wildcards and the connecting AND operator will retrieve this text.
- Distracters:
    - SELECT (Bush?(w)Kalahari) AND (information(w)seek?) – No: all the search terms are valid, but 'of the' is between the terms 'Bushmen' and 'Kalahari', so the left hand adjacency operators would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.

- SELECT (information(w)retriev?) AND (information (w)science) – No: the text does not contain the term 'science' therefore the right hand adjacency operator would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.
- SELECT (information(w)retrieval) AND (information (w)seeker?) – No: the right hand expression contains a wildcard operator 'seekers?' which would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.

Figure 11 – Example Dialog Boolean question

A variation would be to augment these Boolean expression question sets with examples of correct and incorrect Dialog syntax. The next stage is to build on questions presented on Boolean expressions and Dialog syntax in terms of search strategies.

5.4 Search strategies

At this stage we can start looking at higher level search concepts (group C), such as search strategies. We concentrate on the types of strategies available on command line interfaces such as Dialog i.e. quicksearch, building blocks and successive fractions. We assume some confusion between strategies particularly building blocks and successive fractions, which require knowledge of intermediate sets and building a final search from these sets.

- Stem: Given the search terms 'human', 'computer', 'interaction' which of the following is a valid form of the 'building blocks' search strategy?
- Key: Correct – sets are built up one by one, and the final result set generated at the end of the strategy.
  Set1 = human
  Set2 = computer
  Set3 = interaction
  Set4 = 1 AND 2 AND 3
- Distracters:
  - Set1 = human AND computer AND interaction
    – No: this is an example of a quicksearch – in building blocks sets are built up one by one, and the final result set generated at the end of the strategy.
  - Set1 = human
    Set2 = computer AND Set1
    Set3 = interaction AND Set2
    – No: this is an example of successive fractions, which builds up intermediate sets incrementally – in building blocks sets

are built up one by one, and the final result set generated at the end of the strategy.

- Set1 = human
  Set2 = computer
  Set3 = interaction
  Set4 = 1 OR 2 OR 3
  – No: while the first three stages are correct for a Building Blocks strategy, AND is always applied to sets, as each intermediate set represents a facet, therefore the application of OR at the last stage is invalid in this kind of search strategy.

Figure 12 – Example search strategy question

The issue of confusing applying AND and OR between and within facets can also be tested further. Other strategies such as 'Citation Pearl Growing' can also be investigated. Building on this we can then present questions on search strategies using valid Dialog search syntax (see figure 13).

- Highlighted Text: "**Information retrieval** is the art of the possible. Most people – no matter what society they live in – are involved in some kind of **information seeking**. For example the **Bushmen** of the **Kalahari** need information on waterholes in the desert. Although they live in a non- technical society, their thirst brings about an information need, in this case to find a waterhole"

- Stem: Which of the following Boolean expressions would retrieve the above text?
- Key: Set1 = SELECT (information(w)retriev?)
  Set2 = SELECT (information(w)seek?)
  COMBINE 1 AND 2
  - Correct: all proximity operators in combination with wildcards and the connecting AND operator will retrieve this text.
- Distracters:
  - Set1 = SELECT (Bush?(w)Kalahari)
    Set2 = SELECT (information(w)seek?)
    COMBINE 1 AND 2
    - No: all the search terms are valid, but 'of the' is between the terms 'Bushmen' and 'Kalahari', so the left hand adjacency operators would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.
  - Set1 = SELECT (information(w)retriev?)
    Set2 = SELECT (information (w)science)
    COMBINE 1 AND 2

- No: the text does not contain the term ‘science’ therefore the right hand adjacency operator would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.

- Set1 = SELECT (information(w)retrieval)
  Set2 = SELECT (information (w)seeker?)
  COMBINE 1 and 2
  - No: the right hand expression contains a wildcard operator ‘seekers?’ which would not match the text. Both right hand and left had expressions either side of the AND operator must be satisfied for the text to be retrieved.

Figure 13 – Example Dialog Boolean Search Strategy question

In figure 13, the building blocks strategy is used, but variations of different strategies, also tackling important misconceptions.

5.5 Measures in IR evaluation

We now turn to evaluation measures, proceeding from group A questions on calculations for precision and recall, group B questions on interpreting and evaluating those figures and group C question on high level evaluation using query examples from earlier query sets. Figure 14 shows an example for mean average precision, addressing typical misconceptions of students on the measure.

- Stem: If relevant documents occur at positions 1, 3 and 7 of the results list, assuming that 10 documents are relevant, what is the correct value for average precision on this search.
- Key: 0.21 – Correct: Average precision is calculated by adding the calculations of every relevant document instance at N retrieved, and dividing by the total number of relevant documents known, therefore the result for this search would be (1.0+0.667+0.428)/10 = 0.21.
- Distracters:
  - 21% – No: All precision figures are report between 0-1. Average precision is calculated by adding the calculations of every relevant document instance at N retrieved, and dividing by the total number of relevant documents known, therefore the result for this search would be (1.0+0.667+0.428)/10 = 0.21
  - 2.09 – No: The final stage of dividing but the number of relevant documents has not been applied. Average precision is calculated by adding the calculations of every relevant document instance at N retrieved, and dividing by the total number of relevant documents known, therefore the result for this search would be (1.0+0.667+0.428)/10 = 0.21

- 0.20 – No: this is a rounding error. Average precision is calculated by adding the calculations of every relevant document instance at N retrieved, and dividing by the total number of relevant documents known, therefore the result for this search would be (1.0+0.667+0.428)/10 = 0.21

Figure 14 – Example group A question on precision calculations

A further addition to the question set would be to ensure that the student understood that either negative (-) precision or results greater than 1.0 are invalid. The issue of interpreting precision figures can then be tackled (see figure 15).

- Stem: With regard to precision measures, how many of the following statements are true?
    1. A precision of 0.21 is superior to 0.15
    2. A precision of 0.615 is superior to 0.616
    3. A precision of 0.987 is inferior to 0.876
    4. A precision of 0.543 is inferior to 0.742
- Key: Two – Correct: statements 1 and 4 are correct. The largest value of precision indicates that a search has provided more relevant documents in the retrieved set and/or higher up the ranking.
- Distracters:
    - One – No: statements 1 and 4 are correct. The largest value of precision indicates that a search has provided more relevant documents in the retrieved set and/or higher up the ranking.
    - Three – No: statements 2 and 3 are incorrect. The largest value of precision indicates that a search has provided more relevant documents in the retrieved set and/or higher up the ranking.
    - Four – No: statements 2 and 3 are incorrect. The largest value of precision indicates that a search has provided more relevant documents in the retrieved set and/or higher up the ranking.

Figure 15 – Example group B question on interpreting precision

Variations of this question could address the issue of Recall, and perhaps other measures such as BPREF or DCG. We can the complete the full circle by looking a search strategies using Dialog syntax and examine the effect on precision and recall by narrowing and expanding searches. Figure 16 gives an example on the effect of narrowing down searching and its effect on precision.

- Stem: Which of the following Dialog search strategies would narrow down a search, thereby increasing precision at the expense of recall.

- Key: Set1 = SELECT (information(w)retriev?)
  Set2 = SELECT (information(w)seek?)
  COMBINE 1 AND 2
  - Correct: of all the queries presented, this would yield the narrowest set of results (a small set of documents), and would increase precision.
- Distracters:
  - Set1 = SELECT (information AND retriev?)
    Set2 = SELECT (information(w)seek?)
    COMBINE 1 AND 2
    - No: the AND operator used to generate Set1 is more expansive than the adjacency operator (w), and would therefore increase recall at the expense of precision. It is therefore a more expansive search.
  - Set1 = SELECT (information OR retriev?)
    Set2 = SELECT (information(w)seek?)
    COMBINE 1 AND 2
    - No: the OR operator used to generate Set1 is more expansive than the adjacency operator (w), and would therefore increase recall at the expense of precision. It is therefore a more expansive search
  - Set1 = SELECT (informat?(w)retriev?)
    Set2 = SELECT (informat?(w)seek?)
    COMBINE 1 and 2
    - No: the truncation operator on the term 'informat?' used to generate both Set1 and Set is more expansive than 'information', and would therefore increase recall at the expense of precision. It is therefore a more expansive search

Figure 16 – Example group C question queries and effect on precision

Variations of this would include recall examples, different forms of search strategy, and different combinations of both Boolean and extended Boolean operators.

6. Summary and conclusion

We propose a strategy for building a set of MCQ's in order to support Library and Information Science students – inexperienced with the mathematics required for search – and assist their learning. We propose a strategy which would build up knowledge from simple calculations and operations in evaluation and search, to higher level knowledge in two steps. Issues in group A of the taxonomy would be addressed first e.g. an understanding of Boolean operators. This would then allow the testing of knowledge at group B e.g. the formulation of Boolean expressions and translating them to a relevant online system search syntax. Group C knowledge would then build on the knowledge gained

by testing knowledge of search strategies. Our strategy would eventually tackle the issue of the effect of searching on evaluation, providing a holistic picture of the application of mathematics to information retrieval problems.

The examples given in the chapter are just that, and are by no means exhaustive. The next stage is to implement this strategy and build a full set of questions for the test. We envisage a set of around 20/30 questions each for three question sets, at group A, B and C levels of the taxonomy we put forward above. Iterative refinement to the question sets will be essential, in order to build up knowledge about the cohort and the problems they have in understanding the mathematics required for search. Any problems we outline here are by no means exhaustive.

References

Appleby J, Cox W, (2002) "The transition to higher education". In: Kahn P, Kyle J, (eds) *Effective learning and teaching in mathematics and its applications*, Kogan Page 3-19.

Cleverdon CW. (1967). The Cranfield test on index language devices, ASLIB Proceeding, 19. In: Spark Jones, K and Willet, P. (eds), Readings in Information Retrieval, Morgan Kaufmann, San Francisco, 47- 59.

Croft, T. (2002). "Mathematics: the teaching , learning and support of non-specialists". In: Kahn, P and Kyle, J., (eds) *Effective learning and teaching in mathematics and its applications*, Kogan Page, 144-157.

Dialog. (2010). Dialog Pocket Guide. http://support.dialog.com/searchaids/dialog/pocketguide/. Accessed 19th February 2010.

Factiva. (2001). Inside Out – the complete reference for Factiva.com. http://factiva.com/learning/F-646InsideOutguide.pdf. Accessed 19th February 2010.

Fernandez-Luna JM, Huete JF, MacFarlane A, Efthimiadis EN, (2007) Teaching and learning information retrieval, Inf Retrieval 12:201-226.

Higgins E, Tatham L, (2003) Exploring the potential of Multiple-Choice Questions in assessment, Learning and Teaching in Action 2:1, winter. http://www.celt.mmu.ac.uk/ltia/issue4/higginstatham.shtml. Accessed 16 February 2010.

Higgins E, Tatham L, (2008) Assessing by multiple choice question (MCQ) test, The Higher Education Academy.

http://www.ukcle.ac.uk/resources/trns/mcqs/index.html. Accessed 16 February 2010.

Kekäläinen J, Järvelin K. (2002). Using Graded Relevance assessments in IR evaluation. Journal of the American Society for Information Science and Technology, 53:13, 1120-1129.

McKenna C, Bull J, (1999) Designing effective objective test questions: an introductory workshop, CAA Centre, Loughborough University, 17 June. http://www.caacentre.ac.uk/dldocs/otghdout.pdf. Accessed 16 February 2010.

MacFarlane A, (2007) Pedagogic challenges in Information Retrieval - teaching mathematics to Postgraduate Information Science students. In: Huete J, Fernandez Luna JM, MacFarlane A, Ounis I, First International Workshop on Teaching and Learning of Information Retrieval (TLIR 2007). http://www.bcs.org/server.php?show=nav.8704. Accessed 12 February 2010.

MacFarlane A, (2009) Teaching mathematics for search using a tutorial style of delivery, Inf Retrieval 12:162-178.

Smith G, Wood L, Crawford K, Coupland M, Ball, G, Stephenson B, (1996). Constructing mathematical examinations to assess a range of knowledge and skills, International Journal of Mathematical Education in Science and Technology, 30, 47-63.

Text Retrieval Conference (TREC) web site. (n.d.). http://trec.nist.gov/. Accessed 19th February 2010.