Linking Specialized Online Medical Discussions to Online Medical Literature

Sam Stewart[1], Syed Sibte Raza Abidi[1], Allen Finley[2]

[1]NICHE Research Group, Faculty of Computer Science, Dalhousie University, Halifax, Canada

[2]IWK Health Centre/Dalhousie University, Halifax, Canada

April 26, 2010

Acknowledgement

This work is carried out with the aid of a grant from the International Development Research Centre (IDRC), Ottawa, Canada

Outline

- Introduction
- Project Framework
  - Literature Search Strategy
- Preliminary Results

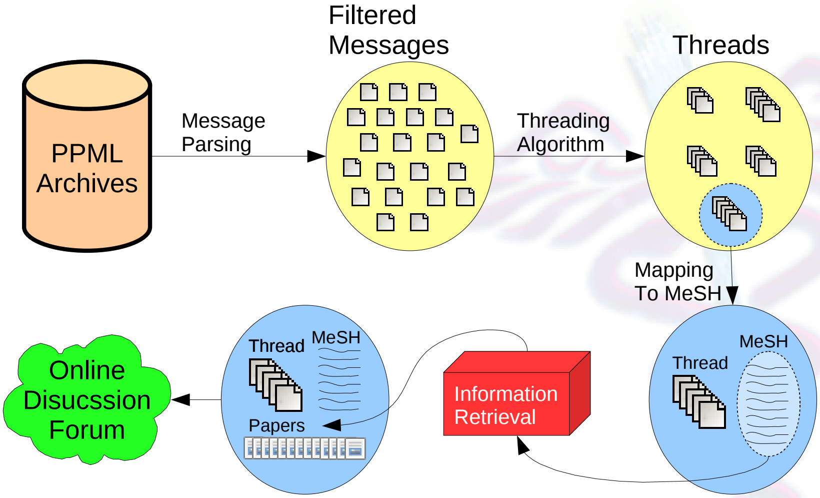
Introduction

- Pediatric pain management is a complex subject, as children lack the cognitive ability to properly express their pain, which can lead to incorrect interventions [1].
- The problem is exacerbated due to lack of specialized knowledge or training on the subject [2].
- Because of the temporal and physical restrictions that clinicians face, traditional educational systems are not a plausible solution
- Web 2.0 technologies provide alternate knowledge dissemination mediums for clinicians to converge and share their knowledge on the subject of pediatric pain

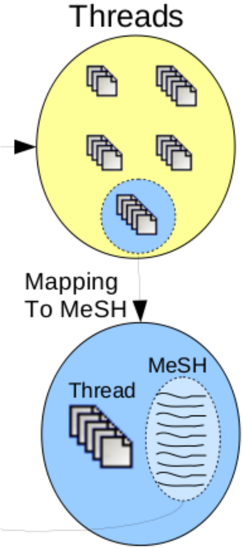
PPML

- One such initiative is the Pediatric Pain Mailing List (PPML), which brings together over 700 pediatric pain practitioners from around the world to share their clinical experiences and seek advice
- The knowledge shared on the PPML is practice-based rather than evidence based
- To generate comprehensive healthcare knowledge around a specific topic, it is important to augment the practice-based (tacit) knowledge on the PPML with explicit knowledge
- The goal of this project is to establish *knowledge linkages* between discussions on the PPML and published literature on Pubmed.

Project Outline



Mapping to MeSH



- Pubmed is indexed by the MeSH lexicon
- The threads are parsed and connected to formal MeSH terms, which will be used to query Pubmed.
- This project used the Metamap program [3] to process the text and connect it to MeSH.
- A set of MeSH terms is created for each thread, along with a score indicating the strength of the connection

Metamap

- Metamap uses its own NLP parser called SPECIALIST [3] to identify all the nouns and noun-phrases in the thread
- It maps each term to one or more terms from the MeSH vocabulary
- Each mapping is assigned a score that is a measure of the strength of the mapping.
- The Metamap scoring systems provides a baseline measure of how well the mapped MeSH term represents the original term in the thread

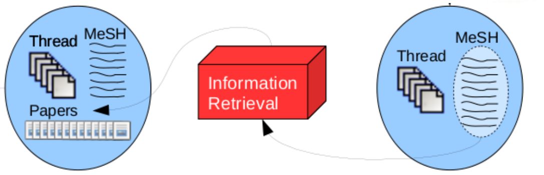
Example

Sample Statement

''The report stated that when **music therapy** is used, **the babies** required **less pain medication**. Does anyone know of any **published reports** of **empirical research** demonstrating the effect?''

Source	**MeSH Term**	**Score**
music therapy	Music Therapy	1000
the babies	Infant	966
less pain medication	Pain	660
less pain medication	Pharmaceutical Preparations	827
published reports	Publishing	694
empirical research	Empirical Research	1000

Literature Search Strategy



- The objective of the search strategy is to *passively* link the threads to published medical literature
- The naive approach would be to search Pubmed and retrieve all the papers that contain every MeSH terms returned. If no papers exist the algorithm would drop the lowest scoring terms and reiterate

Naive Approach

- The naive approach has several problems
  - It doesn't provide any kind of ordering on the resulting papers
  - It doesn't fully utilize the MeSH scores
  - It doesn't take into account the possibility of incorrect mappings
- One of the challenges of mapping free text with Metamap is its inaccuracy [4, 5, 6].
- The presence of a false MeSH term with a high MeSH score will prevent the retrieval of useful papers

Extended Boolean Information Retrieval (eBIR) I

- The eBIR system incorporates query weights into the traditional BIR model
- Let the set of query terms be $A = \{(A_1, s_1), \ldots, (A_n, s_n)\}$, where A_i is the i^{th} query term, and s_i is the associated score
- Let d_{A_i} be an indicator of whether the document D contains the MeSH query d_{A_i}.
- Let the OR and AND queries be

$$Q_{OR(p)} = \{(A_1, s_1) \text{ OR }^p \ldots \text{ OR }^p (A_n, s_n)\}$$

$$Q_{AND(p)} = \{(A_1, s_1) \text{ AND }^p \ldots \text{ AND }^p (A_n, s_n)\}$$

Extended Boolean Information Retrieval (eBIR) II

- The p-norm scores for each of the searches is given in equations (1) and (2).

$$sim(D, Q_{OR(p)}) = \left[\frac{s_1^p d_{A_1}^p + s_2^p d_{A_2}^p + \ldots + s_n^p d_{A_n}^p}{s_1^p + s_2^p + \ldots + s_n^p} \right]^{1/p} \quad (1)$$

$$sim(D, Q_{AND(p)}) = 1 - \left[\frac{s_1^p (1 - d_{A_1})^p + \ldots + s_n^p (1 - d_{A_n})^p}{s_1^p + \ldots + s_n^p} \right]^{1/p} \quad (2)$$

- The selection of p effects the relative strengths of the returned scores.

Modified IR Algorithm

- The problem with applying the eBIR algorithm to this project is that it doesn't address the issue of specialized domains
- MeSH keywords such as *Pediatrics* or *Pain* could be implicitly representative of all conversations on the list
- To solve the problem of specialized domains it was decided that a *specialized filter* would be added, adding an AND operator to the query

Modified IR Algorithm

- The new query would modify the search query by adding *Infant*, *Child* and *Adolescent* to the set of MeSH terms, as demonstrated in equation (3)
  - Let (M_i, m_i) be MeSH term i and the associated Metamap score.

$$\begin{aligned} Q &= [\mathit{Infant}\ \mathrm{OR}^p\, \mathit{Child}\ \mathrm{OR}^P\, \mathit{Adolescent}]\ \mathrm{AND}^P \\ &\quad [(M_1, m_1)\ \mathrm{OR}^P (M_2, m_2)\ \mathrm{OR}^P \ldots (M_n, m_n)] \end{aligned} \tag{3}$$

- Using the eBIR algorithm the next step would be to apply query weights to the terms in the specialized filter and then find a suitable value for p.

Modified IR Algorithm

- This project decided instead to modify the eBIR equation slightly, making the AND filter a strict filter, and leaving the query weights on the OR operator

$$sim(D,Q) = [1-(1-d_I)(1-d_C)(1-d_A)](m_1 d_{M_1} + \ldots + m_n d_{M_n}) \quad (4)$$

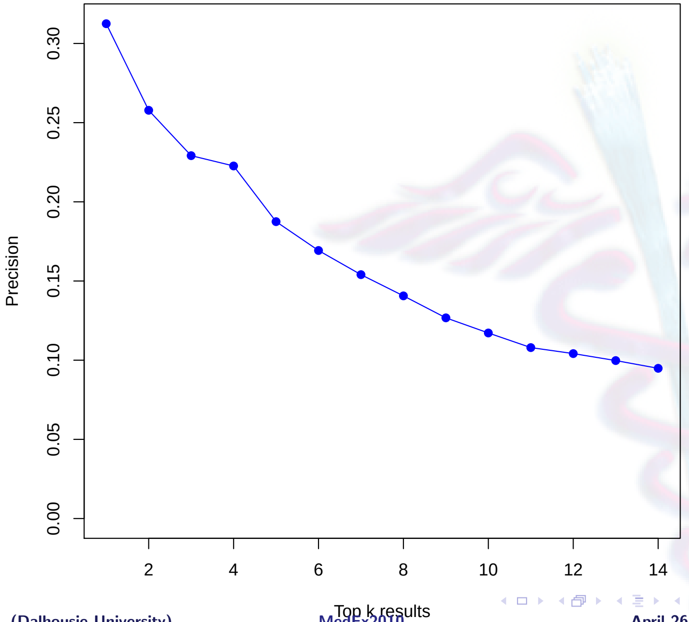
Preliminary Results

- A pilot study was conducted on all messages from 2007 and 2008
- The threads were reviewed to determine
  1. the accuracy of the message parsing
  2. the accuracy of the thread assignment
  3. The accuracy of the papers returned
- The message parsing was successful on 76% of the messages
- The threading was successful on 90% of the messages

Search Strategy Results

- Since the search strategy returns a ranked list of papers, precision-recall curves would be helpful
- Unfortunately, it is not possible to calculate recall, since it requires knowledge of the number of pertinent papers in Pubmed.
- Therefore we plotted precision against the top k results returned, to see what effect returning more papers has on the outcome.

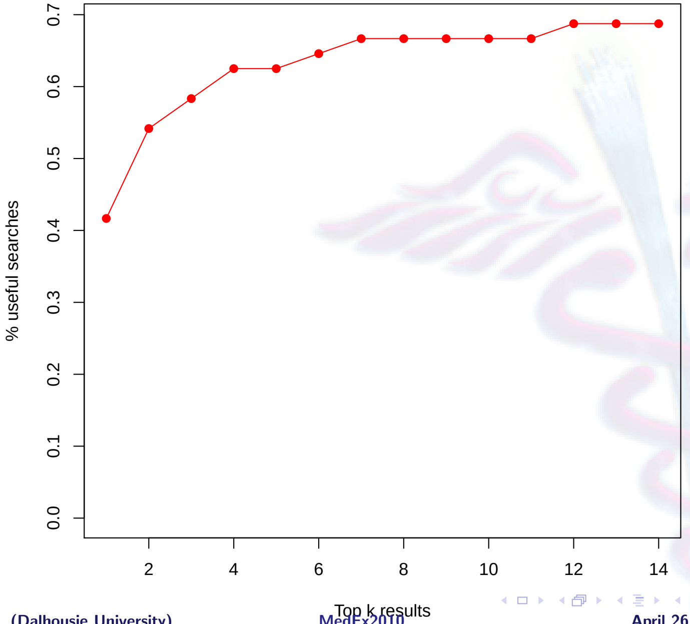
Precision vs. k



Search Strategy Results

- The results are promising, with a high of .3125
- The decreasing nature of the line implies that as we consider papers that are less pertinent, the precision of the search decreases.
- It does not give any advice about a definitive value for k
- The next figure shows, for different values of k, the percentage of searches that return any pertinent results

Utility vs. *k*


% useful searches
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2
4
6
8
10
12
14
Top k results

Utility

- From the figure it seems that there is a dropoff after four papers, though the function is still increasing.
- This means that presenting only four papers would cost about a 10% drop in utility, with a corresponding precision of 0.23
- This result will be useful when the final discussion forum is created.

Conclusion

- The project successfully demonstrates the potential for a knowledge linkage system
- Using the keywords produced from processing free text to retrieve published literature presents a unique IR problem
- Dealing with the Metamap terms and scores requires more than tradtional BIR
- The eBIR system provides this functionality, and slight changes to the system result in a successful linkage strategy

Future Work

- The next step is to implement the system and get feedback from real users
- The search strategy should be compared with other options
- More time should be spent looking into the variables within the eBIR algorithm
- Tweaking the Metamap scores, either within the Metamap system or through post-processing, should also be explored

T. Atherton.

Children's experiences of pain in an accident and emergency department.

Accident and Emergency Nursing, 10:79–82, 1991.

Suzanne Caty, Jocelyne Tourigny, and Irene Koren.

Assessment and management of children's pain in community hospitals.

Journal of Advanced Nursing, 22(4):638–645, 1995.

Alan R. Aronson.

Metamap: Mapping text to the umls metathesaurus.

http://skr.nlm.nih.gov/papers/references/metamap06.pdf, July 2006.

Wendy W Chapman, Marcelo Fiszman, John N Dowling, Brian E Chapman, and Thomas C Rindflesch.

Identifying respiratory findings in emergency department reports for biosurveillance using metamap.

MEDINFO, 2004.

H S Chase, D R Kaufman, S B Johnson, and E A Mendonca.

Voice capture of medical residents' clinical information needs during an inpatient rotation.

Journal of the American Medical Informatics Association, 16:387–394, 2009.

C E Jr Kahn and D L Rubin.

Automated semantic indexing of figure captions to improve radiology image retrieval.

Journal of the American Medical Informatics Association, 16:280–286, 2009.

Thank You

Questions?