The Current State of MetaMap and MMTx

UMLS Webcast

Alan (Lan) R. Aronson
Lister Hill Center/NLM/NIH
alan@nlm.nih.gov

August 20, 2009

(updated December 10, 2009)

Outline

- Historical background
- Distribution modes
- MetaMap and MMTx* similarities
- MetaMap and MMTx differences
- Recent MetaMap development

*MMTx – MetaMap Transfer

Historical Background

- Programs that map biomedical text to a thesaurus
  - CLARIT (*Evans et al., 1991*)
  - SAPHIRE (*Hersh et al., 1990*)
  - **MetaMap *(Aronson et al., 1994)***
  - Metaphrase (*Tuttle et al., 1998*)
  - **MMTx (*2001*)**
  - KnowledgeMap (*Denny et al., 2003*)
  - Mgrep (*2009*)
- Characteristics of MetaMap/MMTx
  - Linguistic rigor
  - Flexible partial matching
  - Emphasis on thoroughness rather than speed

MetaMap/MMTx Example

- PMID – 19529903

- TI –Bile duct stricture due to caused by portal

  Stricture of bile duct | Causing | Hepatic

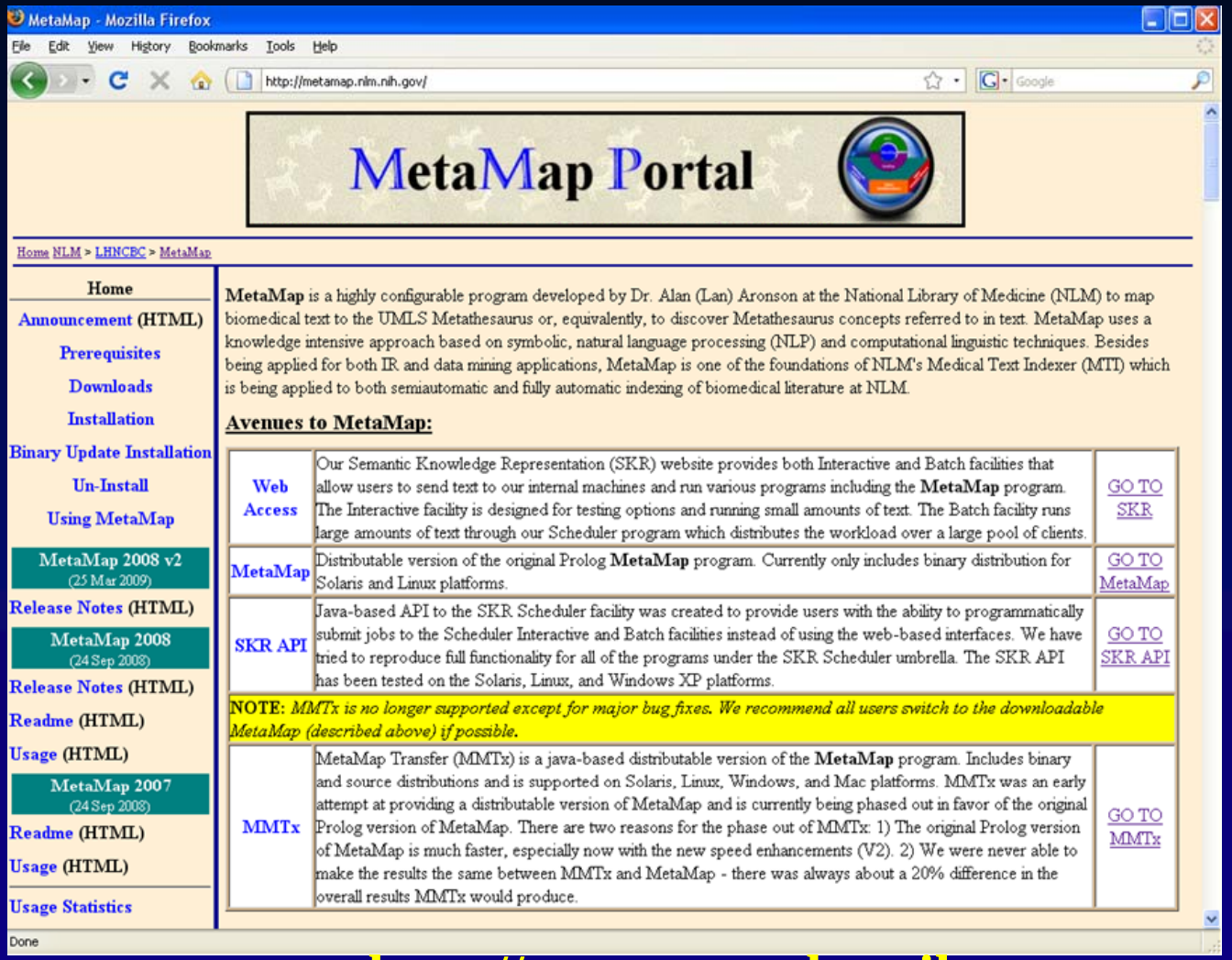
  biliopathy: Treatment with one-stage

  Administration procedure | One | Phase

  portal-systemic shunt and biliary bypass.

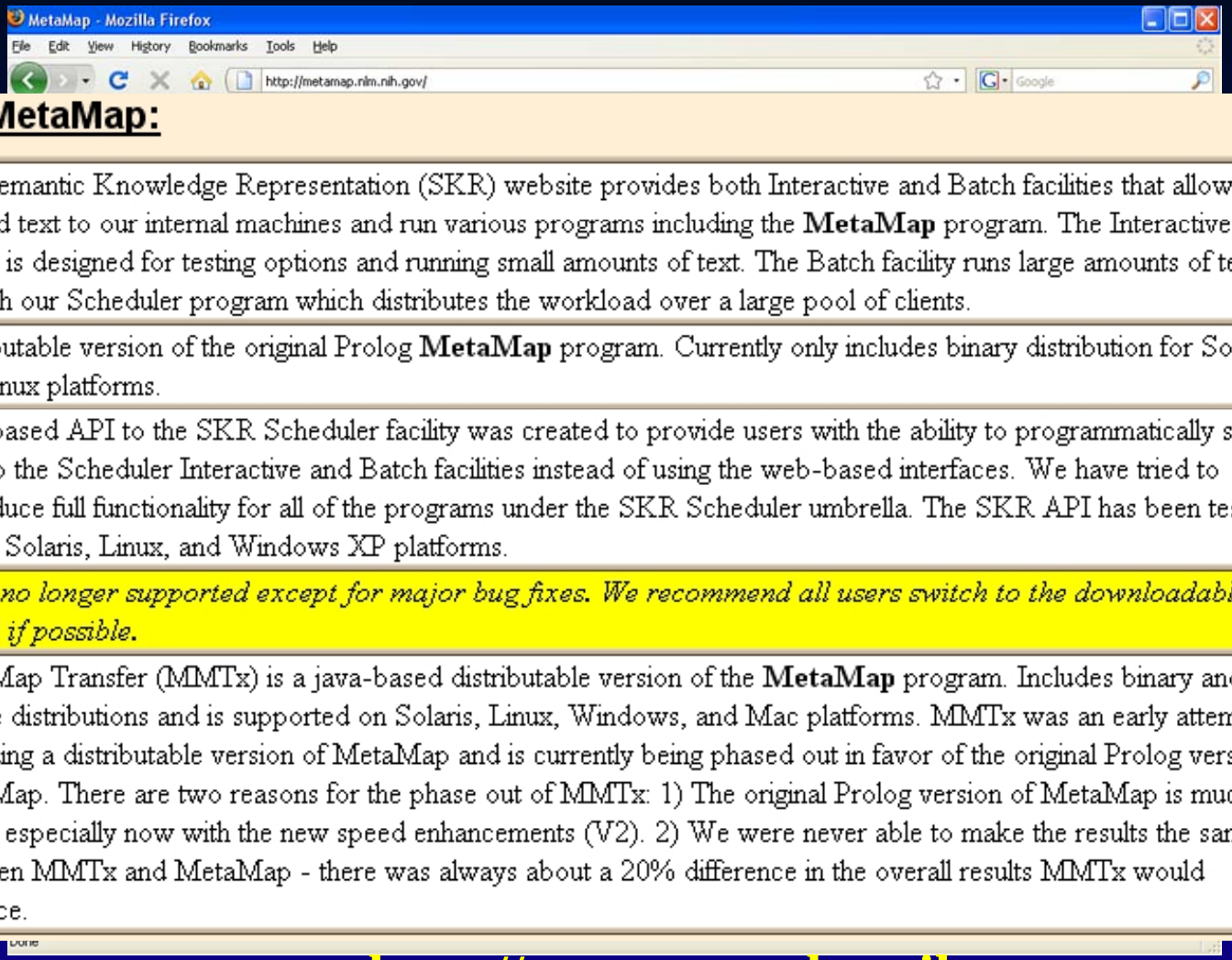
  Portasystemic shunt | Biliary | Bypass

MetaMap/MMTx Distribution Modes



http://metamap.nlm.nih.gov

MetaMap/MMTx Distribution Modes



Avenues to MetaMap:

Web Access	Our Semantic Knowledge Representation (SKR) website provides both Interactive and Batch facilities that allow users to send text to our internal machines and run various programs including the **MetaMap** program. The Interactive facility is designed for testing options and running small amounts of text. The Batch facility runs large amounts of text through our Scheduler program which distributes the workload over a large pool of clients.	GO TO SKR
MetaMap	Distributable version of the original Prolog **MetaMap** program. Currently only includes binary distribution for Solaris and Linux platforms.	GO TO MetaMap
SKR API	Java-based API to the SKR Scheduler facility was created to provide users with the ability to programmatically submit jobs to the Scheduler Interactive and Batch facilities instead of using the web-based interfaces. We have tried to reproduce full functionality for all of the programs under the SKR Scheduler umbrella. The SKR API has been tested on the Solaris, Linux, and Windows XP platforms.	GO TO SKR API
NOTE: *MMTx is no longer supported except for major bug fixes. We recommend all users switch to the downloadable MetaMap (described above) if possible.*		
MMTx	MetaMap Transfer (MMTx) is a java-based distributable version of the **MetaMap** program. Includes binary and source distributions and is supported on Solaris, Linux, Windows, and Mac platforms. MMTx was an early attempt at providing a distributable version of MetaMap and is currently being phased out in favor of the original Prolog version of MetaMap. There are two reasons for the phase out of MMTx: 1) The original Prolog version of MetaMap is much faster, especially now with the new speed enhancements (V2). 2) We were never able to make the results the same between MMTx and MetaMap - there was always about a 20% difference in the overall results MMTx would produce.	GO TO MMTx

http://metamap.nlm.nih.gov

MetaMap and MMTx Similarities

- Same purpose: mapping biomedical text to concepts in the UMLS Metathesaurus
- Same basic algorithm
  - Tokenization and parsing into phrases
  - Variant generation
  - Candidate retrieval
  - Candidate evaluation
  - Final mapping construction

MetaMap and MMTx Differences (1/2)

- Algorithmic details
  - Overall organization of the algorithm
  - Tokenization
- Results
  - Occasional differences, MetaMap's generally preferred
- Programming language
  - Prolog/C (MetaMap)
  - Java (MMTx)

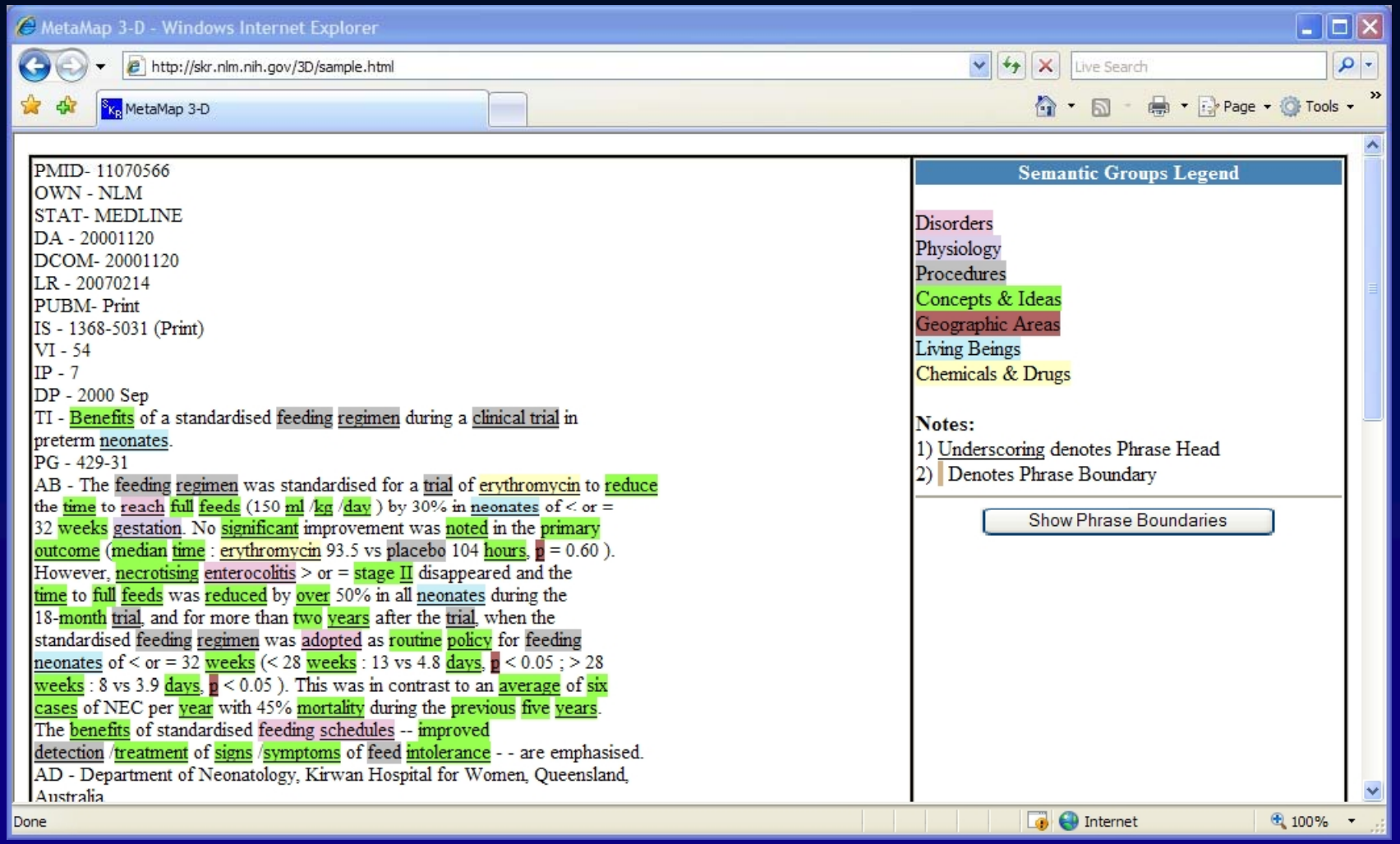
MetaMap and MMTx Differences (2/2)

- Platform availability
  - MMTx: Solaris, Linux, Windows, OS X
  - MetaMap: Solaris, Linux, Windows (soon), OS X (soon)
- Performance
  - MetaMap is 2-5 times faster than MMTx (as of 2008)

Recent/Current MetaMap Development

- Technical algorithm enhancements resulting in at least 3x speedup in MetaMap execution
  - MetaMap is now 3-10 times faster than MMTx (2009)
- Further technical development
  - Migration from Sun/Solaris to Linux environment
  - Update to current Berkeley DB to prepare for
  - Migration from Quintus to SICStus Prolog
- MetaMap now detects negation (via NegEx)
- MetaMap 3D (colorized MetaMap output)

MetaMap 3D


MetaMap 3-D - Windows Internet Explorer
http://skr.nlm.nih.gov/3D/sample.html
Live Search
MetaMap 3-D
Page
Tools
PMID- 11070566
OWN - NLM
STAT- MEDLINE
DA - 20001120
DCOM- 20001120
LR - 20070214
PUBM- Print
IS - 1368-5031 (Print)
VI - 54
IP - 7
DP - 2000 Sep
TI - Benefits of a standardised feeding regimen during a clinical trial in preterm neonates.
PG - 429-31
AB - The feeding regimen was standardised for a trial of erythromycin to reduce the time to reach full feeds (150 ml /kg /day) by 30% in neonates of < or = 32 weeks gestation. No significant improvement was noted in the primary outcome (median time : erythromycin 93.5 vs placebo 104 hours, p = 0.60). However, necrotising enterocolitis > or = stage II disappeared and the time to full feeds was reduced by over 50% in all neonates during the 18-month trial, and for more than two years after the trial, when the standardised feeding regimen was adopted as routine policy for feeding neonates of < or = 32 weeks (< 28 weeks : 13 vs 4.8 days, p < 0.05 ; > 28 weeks : 8 vs 3.9 days, p < 0.05). This was in contrast to an average of six cases of NEC per year with 45% mortality during the previous five years. The benefits of standardised feeding schedules -- improved detection /treatment of signs /symptoms of feed intolerance - - are emphasised.
AD - Department of Neonatology, Kirwan Hospital for Women, Queensland, Australia
Semantic Groups Legend
Disorders
Physiology
Procedures
Concepts & Ideas
Geographic Areas
Living Beings
Chemicals & Drugs
Notes:
1) Underscoring denotes Phrase Head
2) | Denotes Phrase Boundary
Show Phrase Boundaries
Done
Internet
100%

Pointers: Website and Contributors

http://metamap.nlm.nih.gov

Alan (Lan) R. Aronson (alan@nlm.nih.gov)
James G. Mork
Willie J. Rogers
François M. Lang