

Advanced Systems Architecture

Progress Report #1 – 23 March 2006

Team: Joao Castro, Nirav Shah, Robb Wirthlin
Faculty supervisor: Chris Magee

ESD.342

Agenda

- Scope of the problem
 - Comparison of materials
- Expected Data & Hypotheses
- Mapping of local neighborhood of an entry
- Summary & Next Steps

Target data sets

- Online information sources
 - Wikipedia
 - Mathworld
 - Encyclopedia Britannica
 - Britannica online
 - Encarta
- Public transportation networks
 - Boston subway system
 - Paris subway system
 - Seattle transit network
 - New York subway system
 - Porto subway system

Scope the problem

	Wikipedia	Mathworld	Encyclopedia Britannica	Online E. B.	Encarta
Date Established	2001	1995	1768	1994	1993
# of Entries	1 million articles, 340 million words	12527 articles (originally based on notes compiled by a single author, Eric W. Weisstein)	31,550 pages in 32 volumes containing 65,000 articles	120,000 articles, 55 million words; CD-ROM (more than 80,000 articles); DVD-ROM (more than 100,000 articles)	41000 articles, standard 68000 articles, premium (originally based on Funk and Wagnalls, Collier's, and New Merit Scholar Encyclopedias)
# of Contributors	1 million (January 2005, 13,000 or more users who made at least five edits that month; 9,000 of these active users worked on its three largest language editions. A more active group of about 3,000 users made more than 100 edits per month, over half of these users having worked in the three largest editions. One-quarter of Wikipedia's traffic comes from users without accounts, who are less likely to be editors.)	1 (Did one person write all this stuff? Yes. With the exception of contributed entries and entries written by MathWorld staff writer Todd Rowland, the entire contents of MathWorld have been written over the last decade by internet encyclopedist Eric Weisstein, with generous assistance from many people in the mathematics and internet communities. Contributions are most welcome.)	4000	4000	Not available
Stability (rate of change)	High	Med	Low (1976 – 30 volumes; 1994 – 32 volumes); Yearly annual update edition	Med	Med
Accessibility (value proposition)	Free	Free	Purchase	Memberships	Memberships
Peer Review	Little; becoming more frequent	Yes	Yes	Yes	Yes (Graduate students at the University of Washington Information School are fact-checking all proposed changes to Encarta. They are trained in research and passionate about corroborating facts.)
Ease of change	Easy	Changes solicited; reviewed; credit given	Hard	Hard	Changes solicited; reviewed; no credit given
Depth of "knowledge"	Shallow	Mixed	Deep (Micropedia)	Mixed	Mixed

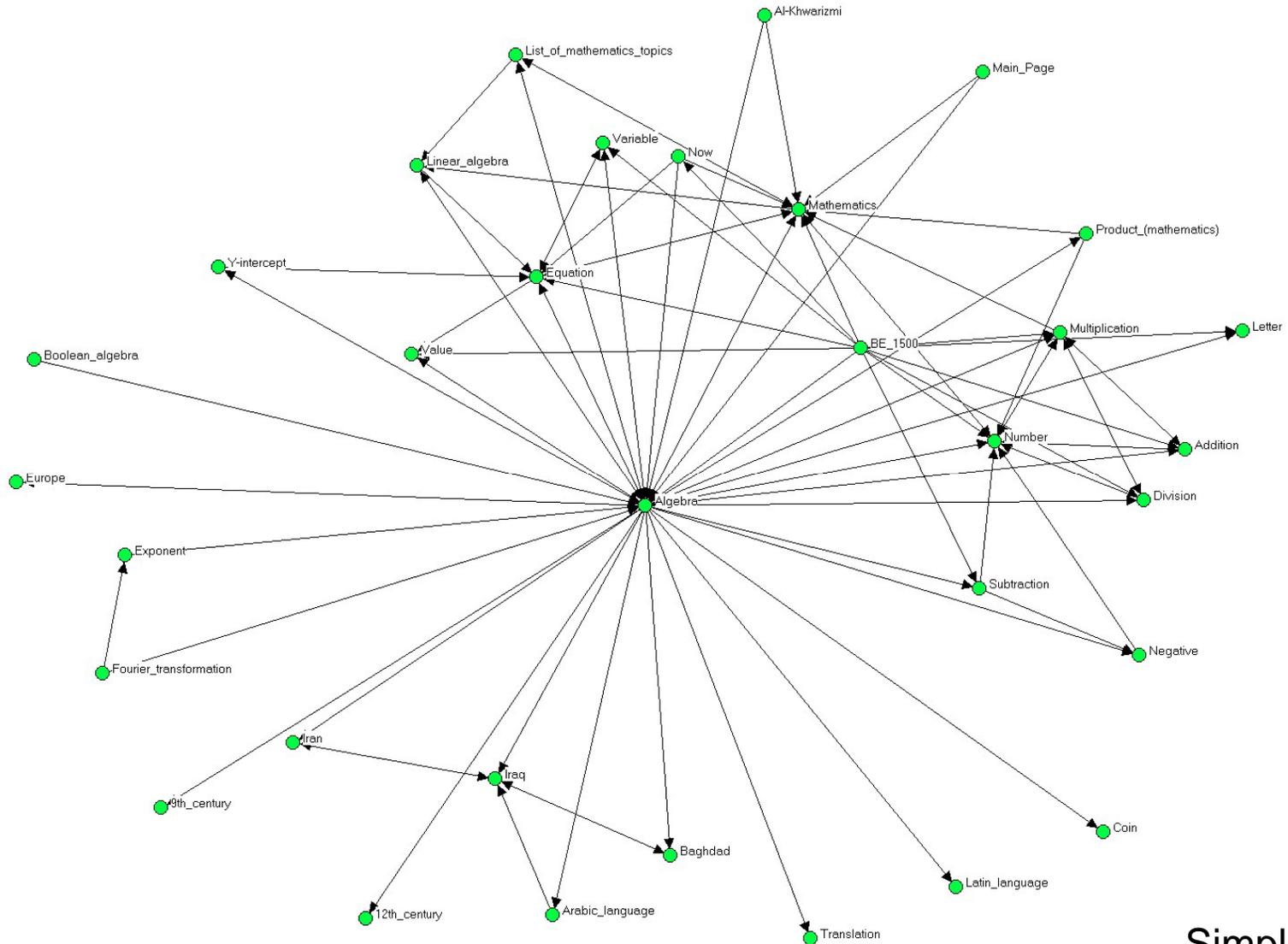
Scope of problem

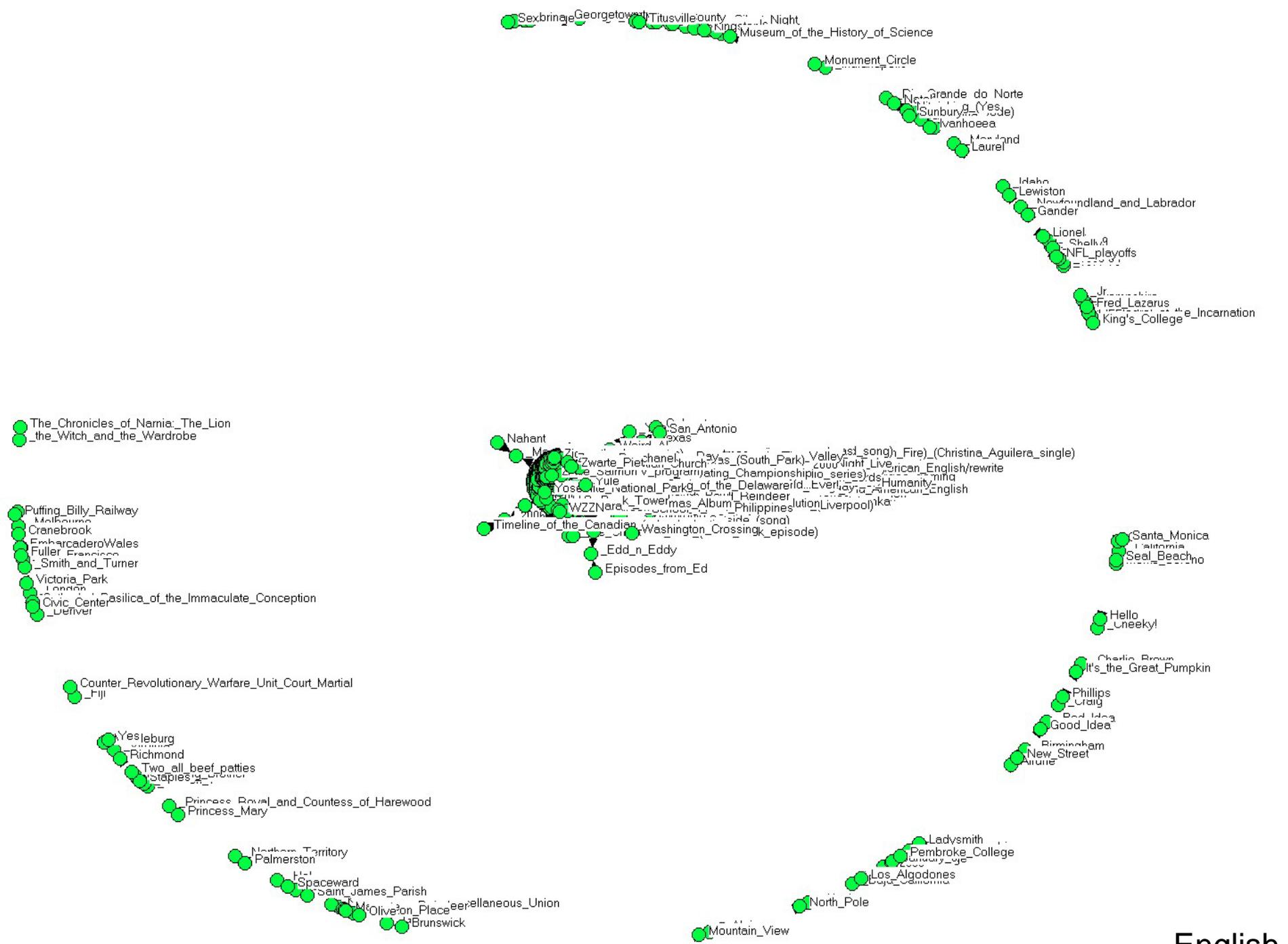
	Seattle transit network	Boston Subway system	Paris subway system	Porto subway system	New York subway system
Date Established		1897	1900	2003	IRT 1904
# of subway stops		122 stops; 6 connections	380 stops; 81 connections	>70	490 stops; lots of connections
# of responsible agencies, etc.		1	1	1	1
Stability (rate of change)				Built in one go.	
Accessibility (value proposition)	Price per km	Price per km, \$1.25 flat rate except for outlying stations	Price per km	Price per km	Price per km
Oversight (regulation)	Public/private board	Public company (the "T")		Publicly owned company	Public company
Ease of change	Years	Years	Years	Years	Years
Distribution of Average Path Length		Long			Short

Expected data/hypothesis

- Completeness, accuracy, neutrality
 - metrics used by Wikipedia for a “Featured Article”
- Network properties
- System properties

Example of neighborhood





English

Summary & Next Steps

- Understand the organization structure and society rules (Britanica
- Observe differences across
 - topics
 - publications
- Evolution of article relationships through time