Lecture 8:
I: IR evaluation
II: Adaptive Hypertext

Johan Bollen
Old Dominion University
Department of Computer Science

jbollen@cs.odu.edu
http://www.cs.odu.edu/~jbollen
IR evaluation

Why we must evaluate?

1. Myriad of possible IR systems
2. Theoretical and logical appeal
3. User preferences and communities
4. Need for metric and benchmarks to determine relative effectiveness

Golden standard: relevancy

1. IR: query → results → relevancy?
2. Who determines relevancy?

(a) User?
(b) Expert?

3. What is relevancy?

Need for formalization of metrics based on relevancy

1. Able to deal with different queries, collections
2. Objective and quantitative metric
3. Simple and intuitive
Concept of Precision and Recall

Two accepted meanings

1. Your favorite DJ’s

2. Two metrics of IR performance

   Collection v. query

   1. Given we have a query

   2. Given we have a collection

   3. We have a set of documents that are all relevant to query

(a) Relevancy remains to be defined

(b) Let’s assume we have an adequate understanding

4. Our IR system returns its results

5. Two questions

   (a) How many irrelevant results?

   (b) How well do results cover set of all relevant results in collection?
Relevancy formalized
Precision and Recall

Let us denote a user information request as \( I \) and the set of document relevant to that request as \( R \).

For any request \( I \), a set of results is returned: \( A \).

The cardinality of \( A \) and \( R \) is denoted: \(|A|\) and \(|R|\).

We denote the set of relevant results in \( A \) as \( R_a \), so \( R_a \subseteq A \).

\[
\text{Recall} \quad \text{recall} = \frac{|R_a|}{|R|}
\]

or the ratio of relevant results in the answer set over the total number of possibly relevant results for \( I \) in the collection.

\[
\text{Precision} \quad \text{precision} = \frac{|R_a|}{|A|}
\]

or the ratio of relevant results over all results in the answer set.
Precision and Recall
Precision and Recall

Precision
1. Represents how many relevant results returned by IR system
2. Indicates quality of results
3. Corresponds to false positives

Recall
1. Represents how many relevant items were captured in toto by IR system
2. Indicates coverage of results
3. Corresponds to false negatives

Precision v. Recall
1. IR system is a little like fishing
2. Fine net catches a lot of small and large fish: many of them not good catch, but you catch most fish
3. Coarse net catches only large fish: but many fish not caught
4. You can not have both fine and coarse net!
Determining Precision and Recall figures

Assume set of queries: for each query we can determine precision and recall.
Assume a set of queries $Q = \{q_1, q_2, \cdots, q_t\}$. The ideal and complete answer set of all relevant documents for query $q_1$ is:

$$R_{q_1} = \{d_7, d_5, d_6, d_{20}, d_{25}\}$$

We invent an IR model, and apply it to the collection $R$, $R_{q_1}$ subset $R$ using.

It produces a ranking of results:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Document</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>$d_7$</td>
</tr>
<tr>
<td>2</td>
<td>$d_3$</td>
</tr>
<tr>
<td>3*</td>
<td>$d_{20}$</td>
</tr>
<tr>
<td>4</td>
<td>$d_1$</td>
</tr>
<tr>
<td>5</td>
<td>$d_{662}$</td>
</tr>
<tr>
<td>6</td>
<td>$d_{429}$</td>
</tr>
<tr>
<td>7*</td>
<td>$d_{25}$</td>
</tr>
<tr>
<td>8</td>
<td>$d_{99}$</td>
</tr>
<tr>
<td>9</td>
<td>$d_{723}$</td>
</tr>
<tr>
<td>10*</td>
<td>$d_6$</td>
</tr>
</tbody>
</table>

We take into account the ranking produced by the IR system because the user examines results in this order.

Result 1: $d_7 \in R_{q_1}$, precision=100%, recall=20%

Result 3: $d_{20} \in R_{q_1}$, precision=66%, recall=40%

Result 5: $d_{25} \in R_{q_1}$, precision=43%, recall=60%

Result 9: $d_6 \in R_{q_1}$, precision=40%, recall=80%

Above 80% recall, precision=0
Precision and Recall curve
Precision at standard recall levels

1. Notion of precision at recall levels
   (a) Given that we accept a certain recall level
   (b) What will the precision of a given IR system be?
   (c) Standard levels: 0, 10, 20, ..., 100 (11 levels)

2. PR calculated over range of queries
   (a) Requirement to extrapolate values
   (b) Aggregate different precision values for level of recall

3. Averaging procedures and aggregate PR numbers
Average precision

Assume $\overline{P(r)}$ indicates average precision at recall level $r$

$P_i(r)$ represents precision at recall level $r$ for query $i$.

then:

$$\overline{P(r)} = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}$$

Need to extrapolate recall levels

For any given query, recall level may not be precisely within standard level.

Use maximum of neighboring recall levels.
Single Value Summaries

1. We often need to know how an IR system performs for single queries
   (a) Reveal exceptions and specific problems
   (b) Compare different IR systems over set of given queries

2. Averaging at relevant results for given query
   (a) Sample precision at relevant results in ranking
   (b) Average values
   (c) Encourages systems that rank relevant results highly

3. R-precision
   (a) Look at total number of relevant documents, $|R_q|$ for query $q$
   (b) Sample precision at ranking in R-th position

4. Precision histograms
   (a) Compare R-precision for different queries across different IR systems
   (b) Subtract R-precision system 1 from R-precision system 2
   (c) Allows comparison of two system for same set for queries
Notes on Precision and Recall

1. excellent DJs
2. validity problems
   (a) Relevance:
      i. Subjective
      ii. Depends on more than term matches
      iii. Different to consistently evaluate across queries
   (b) All or nothing
      i. Set of relevant documents is not so crisp
      ii. Precision and Recall do not take into account degree of relevancy
      iii. Actual objective should be to compare degree of relevancy for query or IR system and for external relevancy estimator
      iv. Novelty and serendipity: butterfly effect on user satisfaction
   (c) Scale
      i. Large collections: difficult to determine all relevant documents
      ii. Difficult to construct IR evaluation framework
Standardized evaluation framework: Reference collections

1. Lack of consistent frameworks for PR evaluation
2. Different systems, collections, queries

CACM and ISI collections

1. CACM
   (a) Communications of the ACM (1979)
   (b) Data includes categories, references, etc
   (c) 52 test information requests
2. ISI
   (a) Sample of Institute of Scientific Information collection
   (b) 25 test information requests
II: Adaptive Hypertext

Johan Bollen
Old Dominion University
Department of Computer Science

jbollen@cs.odu.edu
http://www.cs.odu.edu/~jbollen
Hypertext: a Brief History

1. Hypertext: systems for non-linear text
   (a) Shift from “traditional”, “linear” texts and media
   (b) User determines own sequence, experience
      i. Personalization
      ii. Interactive

2. Central Concept:
   (a) Network of documents
   (b) Document or document subcomponents link to other documents or its subcomponents
   (c) Author refers reader to related material
   (d) Reader can select links and determine own path
Where technology makes a difference.

1. Existing document networks:
   (a) Citation graphs
   (b) Reviews

2. What makes hypertext different?
   (a) Technology: digital storage, allows fast and convenient addition and incorporation of new knowledge
   (b) User experience of seamless transitions
   (c) Mimics associative and automated style of human information
   (d) Integration: links have become integral component of communication

The Memex: mother of all hypertext

1. V. Bush (1945): “As we may think”
2. Augmentation of human intellect: important issue
3. Microfilm based device, personal information assistant
4. Indexing combined with adaptive information links
Later developments

1. Early developments: focus on technology
2. Focus on augmentation of human information processing
3. Strong ties to innovations in Human Computer Interaction

Douglas Engelbart (Bootstrap institute):
1. Probably first implementation of functional hypertext systems (1968): On-Line System (NLS)
2. Collaboration (AUGMENT), Video Conferencing
3. Computer Mouse!
4. Hypertext (Adaptive!), many different formats, change of representation, organization of ideas

ZOG (1972), KMS (1981):
1. Pure hypertext system for organizing organizational data
2. Two screens: user edit text, insert hyperlinks and construct annotations
3. Simplicity of interface

Intermedia System (1986):
1. Education purposes, Apple MacOS
2. Hypertext toolkit for variation of formats
3. Bi-directional links, stored separate from text for maintainability
Later developments, cont.

**Guide (1987):**

1. First commercially available hypertext systems for PC and Apple
2. Integration of text, images and video
3. Hyperlinks: variety of in-text buttons
4. Absence of distinction between authors and readers

**HyperCard (1988):**

1. Closed hypertext system
2. Page: object containing text, GUI
3. Tool to develop interactive applications

**Xanadu (1965):**

1. Term Hypertext: Ted Nelson (1965)
2. Xanadu project: issues of ownership, copyright, version control, link verification, re-use

http://www.sfc.keio.ac.jp/~ted/XUsurvey/xuDation.html
An attempt to generalize

**Three components:**

1. Hypertext network: network of linked documents
2. Interface: visualizes document content and links, implements link actualization
3. User: reads documents, selects links, “navigates” network
Why the WWW operates as a hypertext system.

1. Separation of storage (HTTP server) and retrieval (client)
2. Hyperlinks point to local and non-local resources using URLs
3. Client translates retrieved pages to hypertext
4. Decentralized paradigm offer user illusion of hypertext system
Why the WWW is a BAD hypertext network.

Ted Nelson:

The Web is the minimal concession to hypertext that a sequence-and-hierarchy chauvinist could possibly make...

... The Web is a special effects race, FANFARES ON SPREADSHEETS! JUST WHAT WE NEED! (Instead of dealing with the important structure issues—structure, continuity, persistence of material, side-by-side intercomparison, showing what things are the same.) This is cosmetics instead of medicine.

1. Network can not operate on any level higher than individual HTTP server
2. Singular hyperlinks
3. HTML: special effects race
4. URL: static file identifiers

However, it upended many hypertext developments and apparently stopped most fundamental HT research.
Back to non-Linear text: is it really any better?

Two issues:
1. Use: retrieval and education
2. Design: manual or automated

Evaluations of Use and Design

Use:
1. Juxtaposition of linear, traditional text vs. hypertext is mistaken
2. No data supporting the superiority of hypertext over traditional text
3. Education benefits questioned
4. Retrieval: boolean search engines preferred by users
5. Lost in hyperspace phenomenon: Users loose track of where they are and where they are going

Design:
1. Manual design: uneconomical to the extreme
2. Strong shift to database approaches on WWW: end of hypertext on WWW
3. Automated hypertext linking: text content based IR techniques, database approaches
Adaptive hypertext and Hypermedia

Objectives:
1. Reduce navigation problems
2. Accommodation of different users and user community preferences and goals
3. Support educational outcomes
4. Improve interface

Adaptation process focus:
1. Structure: hyperlink structure
2. Presentation: interface improvements, link annotation, navigation guidance
3. Content: document structure and content
4. User: learning effects

Most systems combine these forms of adaptation User Models:
1. Represents features of user(s)
   (a) Goals: highly sporadic
   (b) Preferences: transient
   (c) Individual Characteristics: stable
   (d) Community Characteristics: very stable
2. Modulates and shapes adaptation process
3. Stated goals and preferences, knowledge models
Examples

Conditional text presentation
1. User model: classification as layman, expert, etc.
2. Conditional text technique:
   (a) Text relevant to concepts is decomposed in segments
   (b) Display condition on classifications
3. Often applied in educational setting

Link hiding based on education goals
1. Definition of hierarchies of tasks
2. Links hidden or displayed depending on dependency rules for learning tasks
3. User selects task and system hides or displays links
4. Often performed in combination with link ontologies

Domain Knowledge or overlay models
1. Semantic network of domain concepts
2. Representation of user knowledge: value of individual concepts
3. Generation of navigation aids or selective display of content and links based on concept relations

Adaptive Presentation: link and page ordering
1. User vote or rate pages and links
2. Order of link presentation is adapted
3. Individual or collective
Problems

Design:

1. Generated from structural features
2. Text content based IR techniques
   (a) Database queries
   (b) Vector space methods
   (c) User design
3. Use of existing author-defined document networks (citation graphs)
4. Collaborative filtering: user proximities

Use:

1. Assumption is that network is a given
   (a) Navigation aids based on specific navigation problems
   (b) Aid user
2. Symptomatic approach
3. Little evaluation of effectiveness
4. Focus on individual, explicit preferences and goals
5. Little consideration of valid knowledge representation and methods of validation
Basic Idea

1. Replace Design by Use
   (a) Frequency of retrieval sequence indicates associative relation
   (b) Text Independent
   (c) Validity with regard to user preferences

2. Automated link updating
   (a) Registration of document co-retrieval patterns
   (b) Learning Rules change link weights
   (c) Gradual Adaptation to Collective Preferences
Design to Use
Basic Principles

1. Representation
   (a) Weighted Hyperlinks: weights indicate degree of association
   (b) Directed Weighted Graph Representation of Network Structure
      i. Assumption is asymmetric associations
      ii. Corresponds to word association research and WWW hyperlinks

2. Adaptation
   (a) Reinforcements based on

Hyperlink Traversals
(b) Learning Rules Operate on Hyperlink Weights
   i. Small additions to link weights for each traversal
   ii. Aimed at gradual adaptation of structure

3. Interface
   (a) Weighted Ordered Ranking of Hyperlinks
   (b) Dynamic Re-Ordering
   (c) Display “slices” of 10 hyperlinks
Graph Representation

Hypertext Graph Representation (Furner, 1996), (Kleinberg, 1999)

\[ G = (V, L) \]

\[ V = \{v_0, v_1, \cdots, v_{n-1}\} \]

\[ L \subseteq V \times V \]

\((v_i, v_j) \in L\) represents hyperlink \(v_i \in V \rightarrow v_j \in V\)

\(n \times n\) matrix, \(H\), binary values \(h_{ij} \in \{0, 1\}\)

hyperlink: \((v_i, v_j) \in L : h_{ij} = 1\)

absence hyperlink: \((v_i, v_j) \not\in L : h_{ij} = 0\)
Graph Representation Assumptions

1. Hypertext Pages $\sim$ Graph Nodes?
   (a) Homogeneous content
   (b) Singular Concept
2. Hyperlinks $\sim$ Graph Edges
   (a) Singular anchor and target
   (b) Hyperlink Semantics
   (c) Directed Connections

Simplification

1. Feature rather than bug
2. Goal is not to construct “best” hypertext system but study larger class of adaptive HT systems
Learning Rules

1. Frequency:
   (a) Operates on link traversal frequency
   (b) Reinforce existing connections
   (c) Hebbia Learning

2. Transitivity:
   (a) Operates on transitive co-retrievals
   (b) Bridges plausible connections

3. Symmetry:
   (a) Reinforce connection symmetric to those traversed
   (b) Reinforce plausible connections
Learning Rules

\[ G = (V, L) \]

add weight function \( W \),

\[ G = (V, L, W) \]

\( W : V^2 \to \mathbb{R}^+, W(v_i, v_j) \in \mathbb{R}^+ \)

User Request Sequence (RS): \( p = (v_0, v_1, \cdots, v_{n-1}) , n \in \mathbb{N}^+ \)

RS Sub-path Triplet: \( p_s = (v_i, v_j, v_k) , j, k \in \mathbb{N}^+ \)

\[ j = i + 1, k = i + 2, i < n - 2 \]
Learning Rules

Frequency (F) or $F(p_s, W)$:

$$W(v_i, v_j)_t = W(v_i, v_j)_{t-1} + r_f$$

$$W(v_j, v_k)_t = W(v_j, v_k)_{t-1} + r_f$$

Symmetry (S), or $S(p_s, W)$:

$$W(v_j, v_i)_t = W(v_j, v_i)_{t-1} + r_s$$

$$W(v_k, v_j)_t = W(v_k, v_j)_{t-1} + r_s$$

Transitivity (T), or $T(p_s, W)$:

$$W(v_i, v_k)_t = W(v_j, v_k)_{t-1} + r_t$$
Example
Interface

Please choose from the following list the word most related to "act" (of "act"):

- file
- effect
- change
- statement
- exercise
- action
- education
- writing
- More words...

If you are tired, click here to exit the experiment.

Please choose from the following list the word most related to "action" (of "act"):

- movement
- change
- effort
- development
- use
- influence
- history
- action
- take
- course
- More words...

If you are tired, click here to exit the experiment.
Interface

Displaying links

1. Display shows groups of 10 links
2. Rank ordered: first slice displays 10 strongest links
3. User can shift from one group to the other
Adaptive Hypertext System Construction
Why?

1. What Could we learn from such system?
   (a) Network Development: Prototype for General Class of Adaptive Information Systems
   (b) Network Structure: Representation of Community Hyperlink Preferences
   (c) General Properties of Adaptive Information Systems

2. Why is This Important?
   (a) Study of Adaptive Information Systems Efficiency
   (b) General Study of Concept Associations
   (c) Relation Designer vs. User Preferences
   (d) Empirical Study of Intelligent Networks
Experiments

1. Open Participation
   (a) No Registration of Subjects
   (b) No Data Gathered on Characteristics\Motives of Subjects

2. Instructions to Browse Experimental Hypertext System
   (a) Associative Browsing: Position based Hyperlink Selection
   (b) Request not to Apply High-level Strategies

3. Specific Choice of Node Labels
   (a) English nouns
   (b) 150 most frequent

4. Two Experiments
   (a) Frequency + Transitivity ($r_f = 1$, $r_t = 0.5$)
   (b) Frequency + Transitivity + Symmetry ($r_s = 0.3$)
Results

1. Participation
   (a) ±6000 requests
   (b) 600 participants \ experiment

2. Resulting Networks:
   (a) Structure
   (b) Network Development
   (c) Interaction of Learning Rules
   (d) Reliability and Validity
### Table 1: Sample of matrix representing final structure of adaptive hypertext network trained in experiment 1

<table>
<thead>
<tr>
<th>number</th>
<th>label</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>act</td>
<td>0</td>
<td>8.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>(2)</td>
<td>action</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(3)</td>
<td>age</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>(4)</td>
<td>amount</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>(5)</td>
<td>area</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(6)</td>
<td>art</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>(7)</td>
<td>attention</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(8)</td>
<td>authority</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(9)</td>
<td>bed</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(10)</td>
<td>blood</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>number</td>
<td>label</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>(1)</td>
<td>act</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(2)</td>
<td>action</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(3)</td>
<td>age</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(4)</td>
<td>amount</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(5)</td>
<td>area</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(6)</td>
<td>art</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
</tr>
<tr>
<td>(7)</td>
<td>attention</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
</tr>
<tr>
<td>(8)</td>
<td>authority</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(9)</td>
<td>bed</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(10)</td>
<td>blood</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: Sample of matrix representing final structure of adaptive hypertext network trained in experiment 2
General Characteristics

Matrix Experiment 1

density \( \frac{2709}{22350} = 0.121 \)

weight value \( \bar{x} = 3.07, \max = 80 \)
**Centrality**

Un-normalized Centrality node $i$:

$$
 c_i = \frac{\sum_{j=0}^{n} w_{ij} + \sum_{j=0}^{n} w_{ji}}{2}
$$

(1)

10 most central nodes

<table>
<thead>
<tr>
<th>label</th>
<th>centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>knowledge</td>
<td>416.75</td>
</tr>
<tr>
<td>experience</td>
<td>324.75</td>
</tr>
<tr>
<td>development</td>
<td>261</td>
</tr>
<tr>
<td>mind</td>
<td>218.25</td>
</tr>
<tr>
<td>education</td>
<td>210.25</td>
</tr>
<tr>
<td>method</td>
<td>204.5</td>
</tr>
<tr>
<td>thought</td>
<td>191.25</td>
</tr>
<tr>
<td>system</td>
<td>181</td>
</tr>
<tr>
<td>body</td>
<td>178.25</td>
</tr>
<tr>
<td>person</td>
<td>168</td>
</tr>
</tbody>
</table>
# Cluster Analysis

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Age, time, century, day, evening, moment, period, week, year</td>
</tr>
<tr>
<td>Space</td>
<td>Place, area, point, stage</td>
</tr>
<tr>
<td>Movement</td>
<td>Action, change, movement, road, car</td>
</tr>
<tr>
<td>Control</td>
<td>Authority, control, power, influence</td>
</tr>
<tr>
<td>Cognition</td>
<td>Knowledge, fact, idea, thought, interest, book, course, development,</td>
</tr>
<tr>
<td></td>
<td>doubt, education, example, experience, language, mind, name, word,</td>
</tr>
<tr>
<td></td>
<td>problem, question, reason, research, result, school, side, situation,</td>
</tr>
<tr>
<td></td>
<td>story, theory, training, use, voice</td>
</tr>
<tr>
<td>Intimacy</td>
<td>Love, family, house, peace, father, friend, girl, hand, body, face, head,</td>
</tr>
<tr>
<td></td>
<td>figure, heart, church, kind, mother, woman, music, bed, wife</td>
</tr>
<tr>
<td>Vitality</td>
<td>Boy, man, wife, health</td>
</tr>
<tr>
<td>Society</td>
<td>Society, state, town, commonwealth</td>
</tr>
<tr>
<td>Office</td>
<td>Building, office, work, room</td>
</tr>
</tbody>
</table>
## Hyperlink Development

<table>
<thead>
<tr>
<th>Rank</th>
<th>Country</th>
<th>Authority</th>
<th>Authority</th>
<th>Authority</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Thing</td>
<td>Society</td>
<td>Rate</td>
<td>Power</td>
<td>Authority</td>
</tr>
<tr>
<td>3</td>
<td>Government</td>
<td>Rate</td>
<td>Society</td>
<td>Effect</td>
<td>Effect</td>
</tr>
<tr>
<td>4</td>
<td>Policy</td>
<td>Theory</td>
<td>Theory</td>
<td>Knowledge</td>
<td>Influence</td>
</tr>
<tr>
<td>5</td>
<td>Film</td>
<td>Problem</td>
<td>Problem</td>
<td>Government</td>
<td>Government</td>
</tr>
<tr>
<td>6</td>
<td>Heart</td>
<td>Job</td>
<td>Government</td>
<td>Society</td>
<td>Law</td>
</tr>
<tr>
<td>7</td>
<td>Group</td>
<td>Service</td>
<td>Job</td>
<td>Money</td>
<td>Knowledge</td>
</tr>
<tr>
<td>8</td>
<td>Room</td>
<td>Money</td>
<td>Money</td>
<td>Father</td>
<td>Society</td>
</tr>
<tr>
<td>9</td>
<td>Paper</td>
<td>Knowledge</td>
<td>Effect</td>
<td>Law</td>
<td>Order</td>
</tr>
<tr>
<td>10</td>
<td>Order</td>
<td>Day</td>
<td>Service</td>
<td>Rate</td>
<td>Money</td>
</tr>
</tbody>
</table>

| Requests: | 0 | 600 | 1200 | 2400 | 4800 |
Speed of Network Development

1. Network States saved at fixed intervals

2. Development Speed = \( \text{CORR}(s_k, s_{final}) - \text{CORR}(s_{k+1}, s_{final}) \)
Expectations Interaction of Learning Rules

1. Introduction of New Plausible Links by Transitivity and Symmetry
   - Reinforce connections not traversed
   - Introduce required variation into system

2. Selection by Frequency Reinforcement
   (a) If new connections are worthwhile, they will be traversed
   (b) Traversal leads to Frequency reinforcements

Evolution of network structure is shaped by interaction of introduction of links by Transitivity and Symmetry, and selection by Frequency.
Cumulative reinforcements from learning rules for knowledge-research hyperlink

- freq
- trans
- symm

requests x 200

cumulative reinforcement value
Cumulative reinforcements from learning rules for nature-life hyperlink

- freq
- trans
- symm

Cumulative reinforcement value vs requests x 200
Adaptation Problems

**Divergence**  Final Structure may diverge strongly

1. Random initialization
2. Positive feedback in hyperlink selection
   (a) Reinforced links move up in ranking
   (b) Higher probability of selection
3. Amplification of random initialization differences

**Invalidity**  What does final structure represent?

1. Does structure correspond to actual user preferences?
   (a) How can we know?
   (b) What is system really adapting too?
2. Is structure mere artifact of adaptation system?

**Re-Test Reliability**  Do repeated measurements of the same quantity results in similar measurements?

**Validity**  Do measurements actually correspond to the measured quantity?

Or:

1. Do Adaptive Hypertext Networks Converge to Similar Structures When Being Trained by Similar Groups of Subjects

2. Does the Final Structure Correspond to the Users’ Actual Hyperlink Preferences?

**Requires:**

1. Repeated Measurements

2. Cross-Validation
Validity and Reliability of HT Development

1. Reliability: Degree to Which Separately Generated Networks Converge to Central Location (dispersion)

2. Validity: Degree to Which Separately Generated Networks Converge to User Preferences that Determine Browsing or Navigation
Simulation Approach

1. Simulation of Browsing Behavior
   (a) Based on Model of Document Relations
   (b) Selection of Strongest Associations

2. Generation of Multiple Networks
Simulation Principles

1. Selection of Hyperlinks based on two principles:
   (a) Highest Level of Association
   i. Element of decision user is expect to make based on own knowledge of association value ($\neq$) system link weight
   ii. Higher level of association = higher probability of selection
   (b) Rank Order of Hyperlinks in Interface
   i. Links are displayed in groups
   ii. Rank ordered according to system link weights

2. Generation of Multiple Networks
   (a) Simulation of repeated user link selections
   (b) Multiple network comparison: reliability and validity
Selection Probabilities

Hyperlink $i$, selection probability $r(i) = f(r''(i), r'(i), w(i))$

group rank $r'(i)$, within group rank $r''(i)$

$$r'(i) = \text{int}\left(\frac{r(i)}{10}\right)$$

$$r''(i) = r(i) - 10r'(i)$$

$0 < r(i) \leq 150$ nodes

$0 < r'(i) \leq 15$ (15 sets of 10 hyperlinks for 150 nodes)

and

$0 < r''(i) \leq 10$. 

$$s(i) = 0.96496r''(i) \cdot 0.25r'(i).$$

Hyperlink selection probability estimator

$$e(i) = s(i) \cdot w(i)$$
Results

1. 50 networks generated
2. Central network defined as one whose hyperlinks weights are mean value of hyperlink weights over 50 network
3. Correlation of generated networks to “central” network: reliability
4. Correlation of generated networks to simulation’s model of document relations: validity
   - validity: 0.5712**
   - reliability: 0.7971**
   - \( C(H_1, H_2) = 0.63** \)

Conclusion: relatively high validity, high reliability.
Cross-Validation to Word Association Norms

1. University of South Florida Free Closed Word Association Norms

2. FSG looked up (Forward Association Strength (Cue → target) for 150 network nodes: matrix $U$

3. Correlation calculated: $C(H_1, U) = 0.168^{**}$, $C(H_2, U) = 0.192^{**}$

Why such low correlations?

1. Word Association Norms $\sim$ recall
2. Low density of all matrices involved
Readings for next class: April 10th, 2003
