

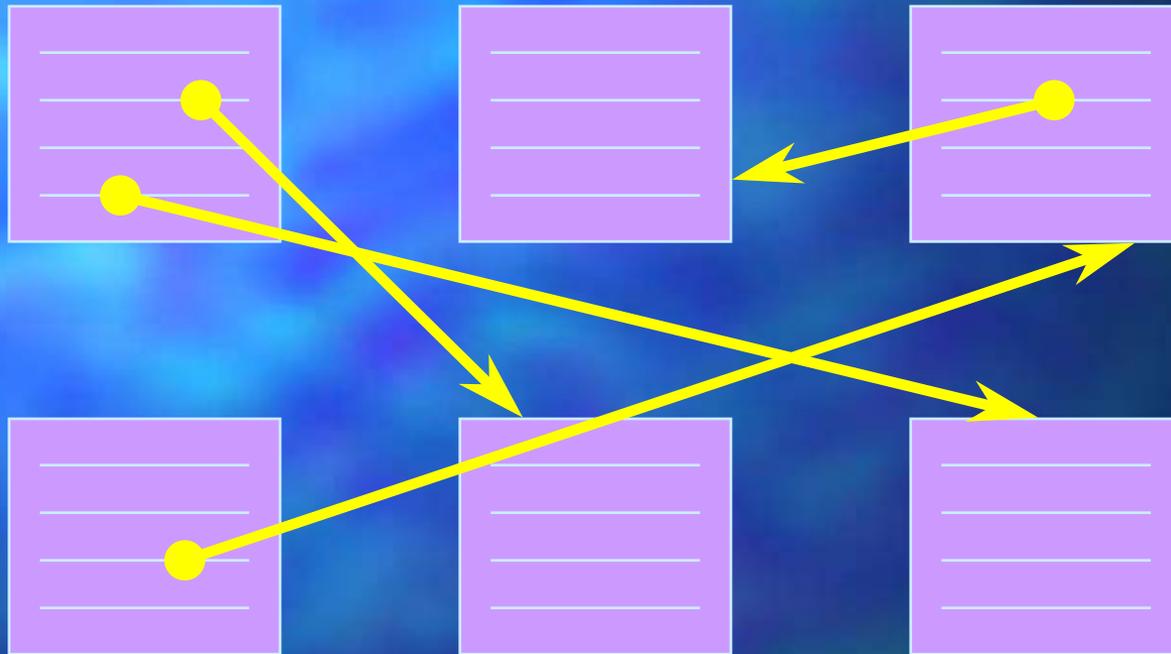
# **Hypertext and Hypermedia**

---

# Definition

- “A database that has active cross-references and allows the reader to ‘jump’ to other parts of the database as desired”
  - Schneiderman, 1989
- Parts of the database called nodes
- Cross-references are called links
- Links tied to a specific point in document, called an anchor

# Definition



# Definition

---

- A link connects two nodes and is normally directed
  - Source node
  - Destination node
  - Normally associated with specific part of source node
    - Anchor
  - Sometimes destination is part of a node
    - Source anchor
    - Destination anchor

# Definition

---

- Most hypertext facilities have a backtrack facility
- Loops are possible
- Some hypertext systems give an indication that a link leads to an already visited node

# Definition

---

- Nodes + Links = Hyperdocument
  - Information content
- Hypertext system
  - Software which lets one read and write hyperdocument
- Hypertext
  - A hypertext system containing a hyperdocument

# Other Definitions

---

- First

- “Hypertext, or non-sequential writing with free user movement along links, is a simple and obvious idea. It is merely the electronification of literary connections as we already know them”

# Other Definitions

---

- Second

- “We can define hypertext as the use of the computer to transcend the linear, bounded and fixed qualities of the traditional written text”

# Other Definitions

---

- Third

- “Mechanisms are being devised which allow direct machine-supported references from one textual chunk to another; new interfaces provide the user with the ability to interact directly with these chunks and to establish new relationships between them. These extensions of the traditional text fall under the general category of hypertext.”

# Other Definitions

## ■ Fourth

- “Hypertext, at its most basic level, is a DBMS that lets you connect screens of information using associative links. At its most sophisticated level, hypertext is a software environment for collaborative work, communication, and knowledge acquisition. Hypertext products mimic the brain’s ability to store and retrieve information by referential links for quick and intuitive access.”

# Other Definitions

## ■ Fifth

- “Hypermedia is Theodore Nelson’s term for computer-mediated storage and retrieval of information in a nonsequential fashion. An extension of Nelson’s earlier coinage, “hypertext” (for non-sequential writing), hypermedia implies linking and navigation through material stored in many media: text, graphics, sound, music, video, etc. But the ability to move through textual information and images is only half the system: a true hypermedia environment also includes tools that enable readers to rearrange the material.”

# Other Definitions

---

## ■ First

- Ted Nelson, *All or One and One for All*, in *Hypertext '87 Papers*, University of North Carolina, Chapel Hill, North Carolina, pp. v-vii

## ■ Second

- G.P. Landow and P. Delany, *Hypertext, Hypermedia and Literary Studies: The State of the Art* in P. Delany and G.P. Landow (Eds.) *Hypermedia and Literary Studies*, MIT Press, Cambridge, Massachusetts, pp. 3-50, 1991

# Other Definitions

---

## ■ Third

- Jeff Conklin, *Hypertext: An Introduction and Survey*, IEEE Computer, Volume 20, Number 9 (1987), pp. 17-41

## ■ Fourth

- J. Fiderio, *A Grand Vision*, Byte Magazine, Volume 13, Number 10 (October 1988), pp. 237-244

## ■ Fifth

- J. McDaid, *Breaking Frames: Hyper-Mass Media* in E. Berk and J. Devlin (Eds.), *Hypertext/Hypermedia Handbook*, McGraw Hill Publishing Company, New York, pp. 445-458

# History

---

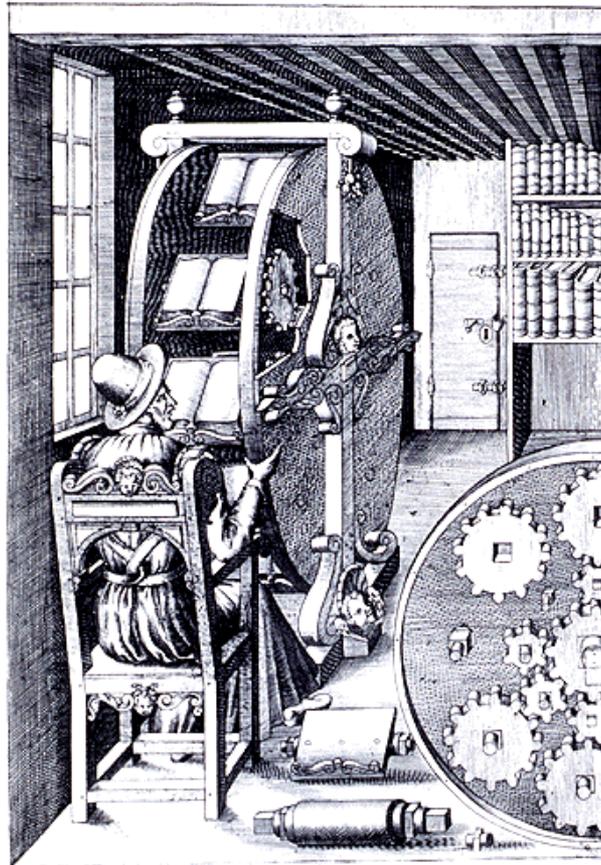
- 1588

- Book *Le diverse et artificiose machine del Capitano Agostino Ramelli*
  - *The Various and Artful Machines of Captain Agostino*

Agostino Ramelli  
*Le diverse et artificiose machine...* (1588)  
[The various and ingenious machines...]

DELL' ARTIFICIOSE MACHINE. 317

FIGURE CLXXXVIII.



© 1995 Smithsonian Institution

# History

---

## ■ 1945

- Vannevar Bush proposes **Memex** in the article “As We May Think”
  - Memory extender
- Never implemented
- Mechanized device which would enable user to view all sorts of written material and organize it arbitrarily, adding annotations and links

# History

---

## ■ 1945

- Bush invented MIT differential analyzer in 1931
- Bush knew computers as large and costly
  - Memex couldn't be implemented using computers
- Memex would store all information on microfilm, kept in one's desk

# History

---

- 1945

- Desk would have several microfilm projectors, enabling user to view several documents at once
  - User would add annotations in margin and they would be scanned into system

# History

---

## ■ 1945

- Ability to create links between items or documents
  - Combining links into **trails** of information relevant to given topics
- Building trails would be a new profession, the **trail blazer**
  - Trails would be shared

# History

---

## ■ 1965

- Ted Nelson introduces **Xanadu** and coins the term 'hypertext'
- A repository for everything ever written
- Announced its release in 1976, 1988, 1991, 1995
  - Byte magazine's first example of *vaporware*

# History

---

## ■ 1965

- User-interface (front-end) versus database (back-end)
  - Back-end available in UNIX
  - Simple front-end available for Sun workstations\
  - Work originated at Brown University, but later supported by Autodesk Company

# History

---

## ■ 1965

- Possible to address any substring of any document from any other location
  - Every byte in every document needs its own address
- Text is never deleted
  - All versions can be generated from latest version
- Author of every document is known and s/he gets royalties based on how many people read how many bytes of author's work

# History

---

## ■ 1967

- Andries van Dam develops the **Hypertext Editing System** at Brown University
- Ran in 128K on an IBM/360 mainframe
- Supported by IBM, who sold to the Houston Manned Spacecraft Center
  - Used to produce documentation for the Apollo space program

# History

---

## ■ 1968

- van Dam develops **FRESS**, File Retrieval and Editing System
  - Timeshared version of previous system
  - Commercially available by Philips
  - Used by faculty and students for many years

# History

---

## ■ 1968

- Doug Engelbart of SRI developed **NLS**, On Line System
  - To store plans, designs, programs, documentation, reports
  - Invented mouse
  - System had video projectors and mice

# History

---

## ■ 1975

- Group at Carnegie-Mellon University developed ZOG
- Frame
  - Segment of ZOG database
  - Consisted of title, description, ZOG commands, and set of menu items leading to other frames
    - Mainly hierarchical with some cross-references
- In 1982, ZOG was installed on U.S. aircraft carrier to manage onboard information

# History

---

## ■ 1978

- Andrew Lippman of MIT Architecture Machine Group (now part of Media Lab) developed **Aspen Movie Map**
  - Simulated ride through Aspen, Colorado
- Videodisks containing photographs of all streets of Aspen
  - 4 cameras, each pointed in different direction, mounted on a truck
  - Photos taken every 3 meters

# History

---

## ■ 1978

- Each photo linked to others which supported user movement of straight ahead, backing up, moving left or right
- User could enter buildings
- System used 2 screens
  - One for video
  - One for map
- Could point to map and jump directly there with video

# History

---

- 1982

- Janet Walker of Symbolics devised the **Symbolic Document Editor**, the first hypertext system widely used
- 8,000 page document represented by a 10,000 node hyperdocument containing 23,000 links
  - 10 Mbytes of storage

# History

---

- 1982

- Authoring tool was separated from user interface

- Concordia

- Structure-oriented editor

- Templates for nodes with fields for standard information

- Hidden fields for authorization information

- Used a generic mark-up language, like SGML, to separate structure from appearance

- Concept of **bookmarks**

# History

## ■ 1985

### – **NoteCards** by Frank Halasz from Xerox PARC

- InterLisp programming environment
- Each node is a single *notecard*
  - Scrolling
- Destination node of a link can be displayed in a new window
- Over 50 specialized types of cards
  - Browser card shows graphical overview of hyperdocument
  - *FileBoxes* are special cards and can contain both *FileBoxes* and other notecards

# History

---

- 1985

- **Intermedia** by van Dam at Brown University

- Scrolling window model for node
- Links connect anchors, not nodes
  - Bidirectional
  - When following link, destination node scrolled so that destination anchor is visible
  - Other applications can be integrated into links

# History

---

- 1985

- **Intermedia** by van Dam at Brown University

- Overview nodes

- Display hyperdocument structure

- Manually constructed using a drawing package

- Web view

- Graphical overview of link structure

# History

---

## ■ 1986

- Office Workstations Ltd. (OWL) in England developed a version of **Guide** for the Macintosh
  - Originally research project at University of Kent
  - Now owned by Matsushita
  - First popular commercial general-purpose hypertext system
  - Link-mechanism usually based on *replacement*, not *pagination*
    - *Jumps* are based on pagination

# History

## ■ 1986

- Office Workstations Ltd. (OWL) in England developed a version of **Guide** for the Macintosh
  - **Pagination**
    - Currently displayed node replaced by destination of link
  - **Replacement**
    - When following link, anchor of link is replaced by contents of destination node
    - One can close destination node
      - Replaced again by anchor text
    - Hyperdocument structure must be hierarchical
  - *Pop-ups* for small annotations

# History

---

- 1987

- Apple introduced HyperCard

- Node object is the *card*
    - Collection of cards called a stack
    - Each card has a button to go to previous and next cards
    - Fields on card can be invisible

# History

---

## ■ 1987

- Apple introduced HyperCard
  - Can have buttons on screen associated with HyperTalk program
    - In most cases, will consist of simple goto statement
  - HyperTalk targeted for prototyping GUI's, not hypertext
- First ACM Conference on Hypertext

# Architecture

---

- Presentation level
  - User interface
- Hypertext Abstract Machine
  - Nodes and links
- Database level
  - Storage and network access

# Architecture

---

- Reference models

- Hypertext Abstract Machine (HAM)

- B. Campbell and J.M. Goodman, 'HAM: A General Purpose Hypertext Abstract Machine,' *CACM*, Volume 31, Number 7 (1988), pp. 856-861

- Trellis

- P.D. Stotts and R. Furuta, 'Petri-Net-Based Hypertext: Document Structure with Browsing Semantics,' *ACM Transactions on Information Systems*, Volume 7, Number 1 (1989), pp. 3-29

# Architecture

- Reference models

- Dexter

- F. Halasz and M. Schwartz, 'The Dexter Hypertext Reference Model,' *NIST Hypertext Standardization Workshop*, February 1990, pp. 94-133

- Written in Z

- Formal model of B. Lange

- D.B. Lange, 'A Formal Model of Hypertext,' *NIST Hypertext Standardization Workshop*, February 1990, pp. 145-166

- Written in the specification language VDM

# Architecture

---

- Reference models

- Tower model

- P. De Bra, G.J. Houben, and Y. Kornatzky, 'An Extensible Data Model for Hyperdocuments,' *Proceedings of the Fourth ACM Conference on Hypertext*, Milan, Italy, December 1992, pp. 222-231

# Navigation

---

## ■ Book

- You can flip pages and read material in any order you like
- You always know where you are
- Author assumes you have read preceding pages for understanding

# Navigation

---

- Hypertext

- You should be able to follow links and never encounter information that relies on information you haven't read

# Navigation

---

- Users of a hypertext may become disoriented
  - Easy to get lost
  - Even in small documents, users experience the 'lost in hyperspace' phenomenon

# Navigation

---

- Navigation of the user through a hyperdocument is influenced by
  - Hyperdocument structure
  - Navigation aids provided by hypertext system
  - Browsing strategy employed by user

# Navigation

---

- Lost in hyperspace
  - An interesting node may be hard to find again in the future
    - Bookmarks

# Navigation

---

- Lost in hyperspace
  - While browsing, you get confused about where you are
    - No directions in hyperspace
    - Fish-eye views
      - Shows only a limited part of a hyperdocument in detail
    - Birds-eye views
      - Detailed maps
      - May be too large to view at one time

# Structural Analysis

---

- Browsing through hypertext versus exploring a city
  - Grid patterns make life easier

# Structural Analysis

---

## ■ Hierarchies

- Hierarchical structure of hyperdocument can be compared to grid structure of a city
- Exceptions to the hierarchy, the cross-reference links, can be compared to non-grid exceptions in city geography, such as Broadway in Manhattan

# Structural Analysis

---

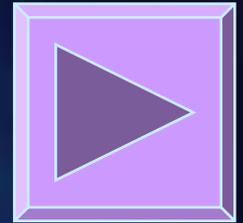
- Identifying hierarchies
  - In order to view a hyperdocument like a book with chapters, sections, subsections, etc., a hierarchical structure must be found
    - The root must be identified
    - Hierarchical and cross-reference links must be distinguished

# Structural Analysis

---

- Identifying hierarchies
  - Root (central node)
    - Every, or almost every, node must be reachable from the root
    - Distance from root to any other node should not be too large
    - Root should have a 'reasonable' number of children

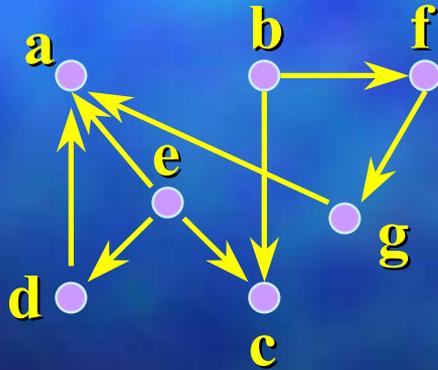
# Structural Analysis



## ■ Identifying hierarchies

– Distance matrix  $D = [d_{i,j}]$

- $d_{i,j}$  is the minimum number of links that are necessary to go from node  $i$  to node  $j$



	a	b	c	d	e	f	g
a	0	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$
b	3	0	1	$\infty$	$\infty$	1	2
c	$\infty$	$\infty$	0	$\infty$	$\infty$	$\infty$	$\infty$
d	1	$\infty$	$\infty$	0	$\infty$	$\infty$	$\infty$
e	1	$\infty$	1	1	0	$\infty$	$\infty$
f	2	$\infty$	$\infty$	$\infty$	$\infty$	0	1
g	1	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$	0

## ■ Identifying hierarchies

- Distance matrix  $D = [d_{i,j}]$ 
  - To define the centrality of a node, we sum the distances from that node to all other nodes
    - Instead of  $\infty$ , we use a large number,  $K$ , called the **conversion constant**
    - Result is called the **converted distance matrix**

# Structural Analysis

- Identifying hierarchies
  - In an  $n$  node hypertext, can let  $K = n$
  - Converted distance matrix

	a	b	c	d	e	f	g										
a	0	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$	a	0	7	7	7	7	7	7	42	5.52
b	3	0	1	$\infty$	$\infty$	1	2	b	3	0	1	7	7	1	2	21	11.05
c	$\infty$	$\infty$	0	$\infty$	$\infty$	$\infty$	$\infty$	c	7	7	0	7	7	7	7	42	5.52
d	1	$\infty$	$\infty$	0	$\infty$	$\infty$	$\infty$	d	1	7	7	0	7	7	7	36	6.44
e	1	$\infty$	1	1	0	$\infty$	$\infty$	e	1	7	1	1	0	7	7	24	9.67
f	2	$\infty$	$\infty$	$\infty$	$\infty$	0	1	f	2	7	7	7	7	0	1	31	7.48
g	1	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$	0	g	1	7	7	7	7	7	0	36	6.44

# Structural Analysis

- Identifying hierarchies

- Converted distance matrix

- Nodes with small row sums have the first two properties of being a root (a central node)

- Row sum of node  $i$  = **Converted Out Distance** for node  $i$

- = **COD <sub>$i$</sub>**  =  $\sum_j d_{i,j}$

# Structural Analysis

- Identifying hierarchies

- Converted distance matrix

- Define the **relative out centrality** for node  $i$  (**ROC<sub>i</sub>**) as  $CD/COD_i$ , where  $CD$ , the **converted distance** of the hypertext is defined by

$$\sum_i \sum_j d_{i,j}$$

- When  $COD_i$  is small,  $ROC_i$  is large
- This measure allows for meaningful comparisons of node centrality for different hypertexts
- For previous example,  $CD = 232$

# Structural Analysis

---

- Identifying hierarchies
  - Index node
    - Node that can be used as an index or guide to many other nodes
    - As in a book, an index node is not a good starting point for the reader
      - Not a good root (central) node

# Structural Analysis

- Identifying hierarchies

- Index node

- Points to many other nodes

- Has high ROC value

- But has many children

- Definition

- Let  $\mu$  be the mean of the outdegrees of the nodes of the hypertext

- Let  $\sigma$  be the standard deviation of the outdegrees of the nodes of the hypertext

- Let  $\tau$  be a threshold value, typically given by  $3\sigma$

- An index node is a node whose outdegree  $> \mu + \tau$

# Structural Analysis

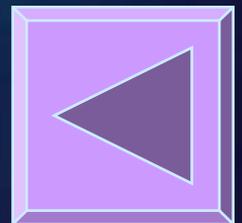
- Identifying hierarchies

- Index node

- For the previous example

$$\mu = (0 + 2 + 0 + 1 + 3 + 1 + 1) / 7 = 8/7 = 1.14$$

$$\begin{aligned}\sigma &= \sqrt{\overline{\text{out}^2} - \overline{\text{out}}^2} = \sqrt{\text{out}^2 - \mu^2} \\ &= \sqrt{\frac{(0^2 + 2^2 + 0^2 + 1^2 + 3^2 + 1^2 + 1^2)}{7} - 1.14^2} \\ &= \sqrt{2.29 - 1.30} \\ &= \sqrt{0.99} \\ &= 0.99\end{aligned}$$



# Structural Analysis

---

- Identifying hierarchies

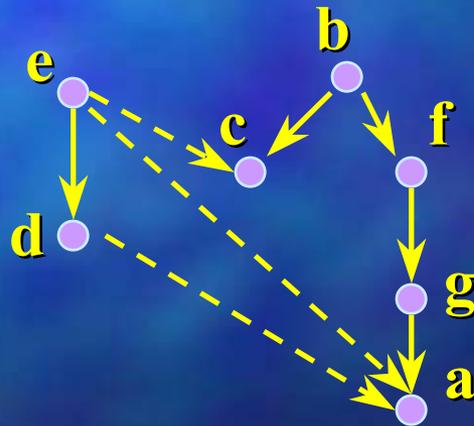
- Index nodes

- So  $\mu + \tau = 4.11$

- No index nodes, though b and e are closest to being them
      - Nodes b and e are good roots

# Structural Analysis

- Identifying hierarchies
  - After root is found, find hierarchical and cross-reference links
    - Breadth-first spanning tree



# Structural Analysis

---

- Identifying hierarchies
  - Maybe some links are missing
    - 2 roots

# Structural Analysis

- Identifying hierarchies
  - Reference node
    - Inverse of index node
      - Many other nodes point to it
    - Definition
      - Let  $\mu^*$  ( $= \mu$ ) be the mean of the indegrees of the nodes of the hypertext
      - Let  $\sigma^*$  be the standard deviation of the indegrees of the nodes of the hypertext
      - Let  $\tau$  be a threshold value, typically given by  $3\sigma^*$
      - A reference node is a node whose indegree  $> \mu^* + \tau^*$

# Structural Analysis

- Identifying hierarchies

- Reference node

- Reference nodes have high values of **Relative In Centrality**,  $RIC_i = CD/CID_i$ , where  $CID_i$ , the **Converted In Distance** for node  $i$  = column sum of node  $i$  =

$$\sum_j d_{j,i}$$

# Structural Analysis

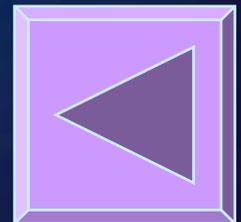
- Identifying hierarchies

- Reference node

- For the previous example

$$\mu^* = (3 + 0 + 2 + 1 + 0 + 1 + 1) / 7 = 8/7 = 1.14$$

$$\begin{aligned}\sigma^* &= \sqrt{\overline{in^2} - \bar{in}^2} = \sqrt{\overline{in^2} - \mu^{*2}} \\ &= \sqrt{\frac{(3^2 + 0^2 + 2^2 + 1^2 + 0^2 + 1^2 + 1^2)}{7} - 1.14^2} \\ &= \sqrt{2.29 - 1.30} \\ &= \sqrt{0.99} \\ &= 0.99\end{aligned}$$



# Structural Analysis

---

- Identifying hierarchies

- Reference node

- So  $\mu^* + \tau^* = 4.11$

- No reference nodes, though a and c are closest to being them

# Structural Analysis

---

- Global Metrics

- Compactness

- High compactness means that each node can easily reach any other node in the hypertext
      - Might be intended
      - Might indicate a poorly structured hypertext that can lead to disorientation

# Structural Analysis

- Global Metrics

- Compactness

- Low compactness may indicate an insufficient number of links and that parts of the hypertext are disconnected

$$C_p = \frac{\text{Max} - \sum_i \sum_j C_{i,j}}{\text{Max} - \text{Min}}$$

# Structural Analysis

- Global Metrics

- Compactness

- Max is the maximum value that the total converted distance can be
      - $\text{Max} = (N^2 - N) K$  in a hypertext of  $N$  nodes
    - Min is the minimum value that the total converted distance can be
      - $\text{Min} = (N^2 - N)$  in a hypertext of  $N$  nodes

# Structural Analysis

---

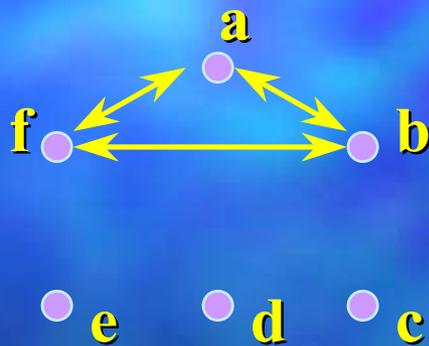
- Global Metrics

- Compactness

- $C_{ij}$  is the converted distance between nodes  $i$  and  $j$
    - When hypertext is fully connected,  $C_p = 1$
    - When hypertext is completely disconnected,  $C_p = 0$

# Structural Analysis

- Global Metrics
  - Compactness

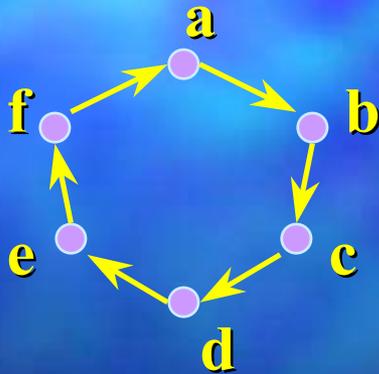


	a	b	c	d	e	f	COD
a	0	1	1	6	6	6	20
b	1	0	1	6	6	6	20
c	1	1	0	6	6	6	20
d	6	6	6	0	6	6	30
e	6	6	6	6	0	6	30
f	6	6	6	6	6	0	30
CID	20	20	20	30	30	30	150

$$C_p = 0.2$$

# Structural Analysis

- Global Metrics
  - Compactness



	a	b	c	d	e	f	COD
a	0	1	2	3	4	5	15
b	5	0	1	2	3	4	15
c	4	5	0	1	2	3	15
d	3	4	5	0	1	2	15
e	2	3	4	5	0	1	15
f	1	2	3	4	5	0	15
CID	15	15	15	15	15	15	90

$$C_p = 0.6$$

# Structural Analysis

---

- Global Metrics

- Stratum

- Captures the linear ordering of the hypertext

- Linear hypertext has stratum = 1

- Can start in only one place

- If one can start anywhere and read everything, stratum = 0

- **Status** of a node

- Sum of finite values on corresponding row of distance matrix

# Structural Analysis

---

- **Global Metrics**

- **Stratum**

- **Contrastatus** of a node

- Sum of finite values on corresponding column of distance matrix

- **Prestige** of a node

- $\text{status}(\text{node}) - \text{contrastatus}(\text{node})$

# Structural Analysis

## ■ Global Metrics

### – Stratum

- Total prestige of a hypertext is always 0
  - Total status of the nodes = total contrastatus of the nodes
- **Absolute prestige** of a hypertext is sum of absolute values of prestige for each node
- **Linear absolute prestige (LAP)** of a hypertext with N nodes is the absolute prestige of a linear hypertext with N nodes
- **Stratum** of a hypertext is the absolute prestige of the hypertext divided by LAP

# Structural Analysis

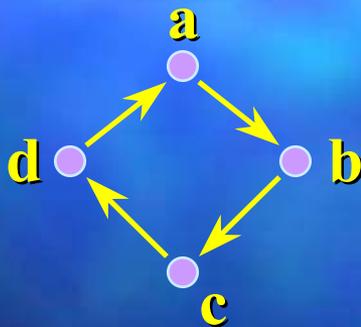
- Global Metrics
  - Stratum



	a	b	c	d	Status	Prestige
a	0	1	2	3	6	6
b		0	1	2	3	2
c			0	1	1	-2
d				0	0	-6
Contrastatus	0	1	3	6		16

# Structural Analysis

- Global Metrics
  - Stratum



	a	b	c	d	Status	Pr estige
a	0	1	2	3	6	0
b	3	0	1	2	6	0
c	2	3	0	1	6	0
d	1	2	3	0	6	0
Contrastatus	6	6	6	6		0

# Navigation Aids

---

- **Backtracking**

- In most hypertext systems, links are unidirectional
- **Back** button
- **Forward** button

# Navigation Aids

---

- Sneak preview

- In Hyperties, a short description of the destination node is given when the cursor is moved over the anchor

# Navigation Aids

---

- Highlighting links
  - Links pointing to 'old' versus 'new' nodes
- Unique anchors
  - Same anchor text must point to same node
- Bread crumbs
  - Bread crumb trail
  - Recognize nodes which were previously visited

# Navigation Aids

---

- History list
  - List of previously visited nodes
  - Can directly jump to them
- Bookmarks
  - Place bookmark on a node
  - Can jump directly there

# Navigation Aids

---

- Birds-eye views
  - Overview of hypertext
  - One approach is to view the hypertext as a tree or forest with cross-reference links as exceptions
    - Won't fit on screen
    - Scrolling window
    - Zoom in and out

# Navigation Aids

---

- Fish-eye views

- Planar graph which shows the structure around the current node in detail, and which shows less and less detail as the distance from the current node gets larger
- Difficulty in deciding which details to leave out

- Guided tours

- Hyperlink

# Navigation Aids

---

- Interest determination based on user navigation history