Chapter 1
Introduction to Multimedia

1.1 What is Multimedia?

- When different people mention the term **multimedia**, they often have quite different, or even opposing, viewpoints.
  - A PC vendor: a PC that has sound capability, a DVD-ROM drive, and perhaps the superiority of multimedia-enabled microprocessors that understand additional multimedia instructions.
  - A consumer entertainment vendor: interactive cable TV with hundreds of digital channels available, or a cable TV-like service delivered over a high-speed Internet connection.
  - A Computer Science (CS) student: applications that use multiple modalities, including text, images, drawings (graphics), animation, video, sound including speech, and **interactivity**.
- Multimedia and Computer Science:
  - Graphics, HCI, visualization, computer vision, data compression, graph theory, networking, database systems.

Components of Multimedia

- Multimedia involves multiple modalities of text, audio, images, drawings, animation, and video. Examples of how these modalities are put to use:

  1. Video teleconferencing.
  2. Distributed lectures for higher education.
  3. Tele-medicine.
  4. Co-operative work environments.
  5. Searching in (very) large video and image databases for target visual objects.
  6. "Augmented" reality: placing real-appearing computer graphics and video objects into scenes.

7. Including audio cues for where video-conference participants are located.
8. Building searchable features into new video, and enabling very high- to very low-bit-rate use of new, scalable multimedia products.
9. Making multimedia components *editable*.
10. Building "inverse-Hollywood" applications that can recreate the process by which a video was made.
11. Using voice-recognition to build an interactive environment, say a kitchen-wall web browser.

Multimedia Research Topics and Projects

- To the computer science researcher, multimedia consists of a wide variety of topics:

  1. **Multimedia processing and coding**: multimedia content analysis, content-based multimedia retrieval, multimedia security, audio/image/video processing, compression, etc.
  2. **Multimedia system support and networking**: network protocols, Internet, operating systems, servers and clients, quality of service (QoS), and databases.
  3. **Multimedia tools, end-systems and applications**: hypermedia systems, user interfaces, authoring systems.
  4. **Multi-modal interaction and integration**: "ubiquity" — web-everywhere devices, multimedia education including Computer Supported Collaborative Learning, and design and applications of virtual environments.

Current Multimedia Projects

- Many exciting research projects are currently underway. Here are a few of them:

  1. **Camera-based object tracking technology**: tracking of the control objects provides user control of the process.
  2. **3D motion capture**: used for multiple actor capture so that multiple *real* actors in a *virtual* studio can be used to automatically produce realistic *animated* models with natural movement.
  3. **Multiple views**: allowing photo-realistic (video-quality) synthesis of virtual actors from several cameras or from a single camera under differing lighting.
  4. **3D capture technology**: allow synthesis of highly realistic facial animation from speech.

5. **Specific multimedia applications**: aimed at handicapped persons with low vision capability and the elderly — a rich field of endeavor.
6. **Digital fashion**: aims to develop smart clothing that can communicate with other such enhanced clothing using wireless communication, so as to artificially enhance human interaction in a social setting.
7. **Electronic Housecall system**: an initiative for providing interactive health monitoring services to patients in their homes
8. **Augmented Interaction applications**: used to develop interfaces between real and virtual humans for tasks such as augmented storytelling.

1.2 Multimedia and Hypermedia

- History of Multimedia:

1. **Newspaper**: perhaps the *first* mass communication medium, uses text, graphics, and images.
2. **Motion pictures**: conceived of in 1830's in order to observe motion too rapid for perception by the human eye.
3. **Wireless radio transmission**: Guglielmo Marconi, at Pontecchio, Italy, in 1895.
4. **Television**: the new medium for the 20th century, established video as a commonly available medium and has since changed the world of mass communications.

5. The **connection** between **computers** and ideas about **multimedia** covers what is actually only a short period:

1945 – Vannevar Bush wrote a landmark article describing what amounts to a hypermedia system called **Memex**.

⟶ Link to full V. Bush 1945 Memex article, "As We May Think"

1960 – Ted Nelson coined the term **hypertext**.

1967 – Nicholas Negroponte formed the **Architecture Machine Group**.

1968 – Douglas Engelbart demonstrated the **On-Line System** (**NLS**), another very early hypertext program.

1969 – Nelson and van Dam at Brown University created an early hypertext editor called **FRESS**.

1976 – The MIT Architecture Machine Group proposed a project entitled **Multiple Media** — resulted in the *Aspen Movie Map*, the first hypermedia videodisk, in 1978.

1985 – Negroponte and Wiesner co-founded the **MIT Media Lab**.

1989 – Tim Berners-Lee proposed the **World Wide Web**

1990 – Kristina Hooper Woolsey headed the **Apple Multimedia Lab**.

1991 – **MPEG-1** was approved as an international standard for digital video — led to the newer standards, **MPEG-2**, **MPEG-4**, and further **MPEGs** in the 1990s.

1991 – The introduction of **PDAs** in 1991 began a new period in the use of computers in multimedia.

1992 – **JPEG** was accepted as the international standard for digital image compression — led to the new JPEG2000 standard.

1992 – The first **MBone** audio multicast on the Net was made.

1993 – The University of Illinois National Center for Supercomputing Applications produced **NCSA Mosaic** — the first full-fledged browser.

1994 – Jim Clark and Marc Andreessen created the **Netscape** program.

1995 – The **JAVA** language was created for platform-independent application development.

1996 – **DVD video** was introduced; high quality full-length movies were distributed on a single disk.

1998 – **XML** 1.0 was announced as a W3C Recommendation.

1998 – **Hand-held MP3 devices** first made inroads into consumerist tastes in the fall of 1998, with the introduction of devices holding 32MB of flash memory.

2000 – WWW size was estimated at over **1 billion pages**.

Hypermedia and Multimedia

- A **hypertext** system: meant to be read nonlinearly, by following links that point to other parts of the document, or to other documents (Fig. 1.1)
- **HyperMedia**: not constrained to be text-based, can include other media, e.g., graphics, images, and especially the continuous media — sound and video.
  - The World Wide Web (WWW) — the best example of a hypermedia application.
- **Multimedia** means that computer information can be represented through audio, graphics, images, video, and animation in addition to traditional media.

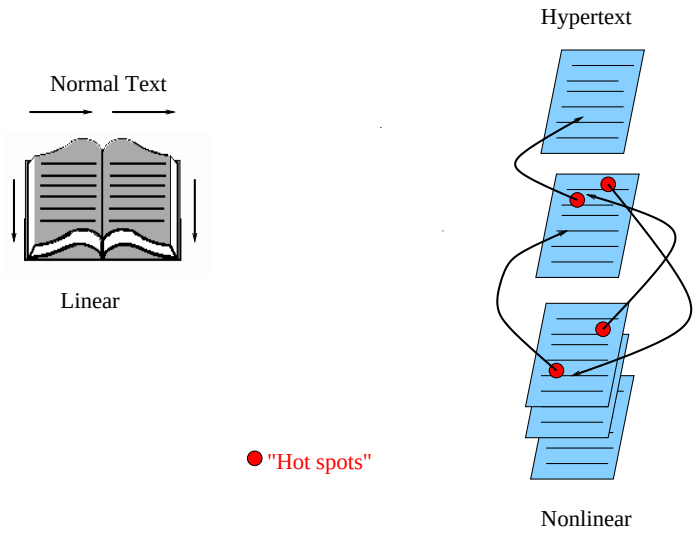

Fig 1.1: Hypertext is nonlinear

- Examples of typical present multimedia applications include:
  - Digital video editing and production systems.
  - Electronic newspapers/magazines.
  - World Wide Web.
  - On-line reference works: e.g. encyclopedias, games, etc.
  - Home shopping.
  - Interactive TV.
  - Multimedia courseware.
  - Video conferencing.
  - Video-on-demand.
  - Interactive movies.

1.3 World Wide Web

- The W3C has listed the following goals for the WWW:

  1. Universal access of web resources (by everyone everywhere).
  2. Effectiveness of navigating available information.
  3. Responsible use of posted material.

- History of the WWW

1960s- Charles Goldfarb et al. developed the Generalized Markup Language (**GML**) for IBM.

1986 – The ISO released a final version of the Standard Generalized Markup Language (**SGML**).

1990 – Tim Berners-Lee invented the HyperText Markup Language (**HTML**), and the HyperText Transfer Protocol (**HTTP**).

1993 – NCSA released an alpha version of **Mosaic** based on the version by Marc Andreessen for X-Windows — the first popular browser.

1994 – Marc Andreessen et al. formed Mosaic Communications Corporation — later the **Netscape** Communications Corporation.

1998 – The W3C accepted **XML** version 1.0 specifications as a Recommendation — the main focus of the W3C and supersedes HTML.

HTTP (HyperText Transfer Protocol)

- **HTTP**: a protocol that was originally designed for transmitting hypermedia, but can also support the transmission of any file type.
- HTTP is a **stateless** request/response protocol: no information carried over for the next request.
- The basic request format:

```
Method URI Version
Additional-Headers:
Message-body
```

- The **URI** (Uniform Resource Identifier): an identifier for the resource accessed, e.g. the host name, always preceded by the token "http://".

- Two popular methods: **GET** and **POST**.
- The basic response format:

```
Version Status-Code Status-Phrase
Additional-Headers
Message-body
```

- Two commonly seen **status codes**:
  1. **200 OK** — the request was processed successfully.
  2. **404 Not Found** — the URI does not exist.

HTML (HyperText Markup Language)

- **HTML**: a language for publishing Hypermedia on the World Wide Web — defined using SGML:

  1. HTML uses ASCII, it is portable to all different (possibly binary incompatible) computer hardware.
  2. The current version of HTML is version 4.01.
  3. The next generation of HTML is XHTML — a reformulation of HTML using XML.

- HTML uses **tags** to describe document elements:
  - `<token params>` — defining a starting point,
  - `</token>` — the ending point of the element.
  - Some elements have no ending tags.

- A very simple HTML page is as follows:

```
<HTML> <HEAD>
 <TITLE>
 A sample web page.
 </TITLE>
 <META NAME = "Author" CONTENT = "Cranky Professor">
</HEAD> <BODY>
 <P>
 We can put any text we like here, since this is
 a paragraph element.
 </P>
</BODY> </HTML>
```

- Naturally, HTML has more complex structures and can be mixed in with other standards.

XML (Extensible Markup Language)

- **XML**: a markup language for the WWW in which there is modularity of data, structure and view so that user or application can be able to define the tags (structure).
- Example of using XML to retrieve stock information from a database according to a user query:
  1. First use a global Document Type Definition (**DTD**) that is already defined.
  2. The server side script will abide by the DTD rules to generate an XML document according to the query using data from your database.
  3. Finally send user the *XML Style Sheet* (XSL) depending on the type of device used to display the information.

- The current XML version is XML 1.0, approved by the W3C in Feb. 1998.
- XML syntax looks like HTML syntax, although it is much more strict:
  - All tags are in lower case, and a tag that has only inline data has to terminate itself, i.e., `<token params />`.
  - Uses name spaces so that multiple DTDs declaring different elements but with similar tag names can have their elements distinguished.
  - DTDs can be imported from URIs as well.

- An example of an XML document structure — the definition for a small XHTML document:

```
<?xml version="1.0" encoding="iso-8859-1"?> <!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
                          transition.dtd">
 <html xmlns="http://www.w3.org/1999/xhtml">
 ... [html that follows
     the above mentioned
     XML rules]
 </html>
```


- The following XML related specifications are also standardized:
  - **XML Protocol**: used to exchange XML information between processes.
  - **XML Schema**: a more structured and powerful language for defining XML data types (tags).
  - **XSL**: basically CSS for XML.
  - **SMIL**: synchronized Multimedia Integration Language, pronounced "smile" — a particular application of XML (globally predefined DTD) that allows for specification of interaction among any media types and user input, in a temporally scripted manner.

SMIL (Synchronized Multimedia Integration Language)

- **Purpose of SMIL**: it is also desirable to be able to publish multimedia presentations using a markup language.
- A multimedia markup language needs to enable scheduling and synchronization of different multimedia elements, and define their interactivity with the user.
- The W3C established a Working Group in 1997 to come up with specifications for a multimedia synchronization language — SMIL 2.0 was accepted in August 2001.
- SMIL 2.0 is specified in XML using a *modularization* approach similar to the one used in xhtml:

1. All SMIL elements are divided into modules — sets of XML elements, attributes and values that define one conceptual functionality.
2. In the interest of modularization, not all available modules need to be included for all applications.
3. **Language Profiles**: specifies a particular grouping of modules, and particular modules may have integration requirements that a profile must follow.
   - SMIL 2.0 has a main language profile that includes almost all SMIL modules.

- Basic elements of SMIL as shown in the following example:

```
<!DOCTYPE smil PUBLIC "-//W3C//DTD SMIL 2.0"
"http://www.w3.org/2001/SMIL20/SMIL20.dtd">
<smil xlmns=
"http://www.w3.org/2001/SMIL20/Language">
<head>
    <meta name="Author" content="Some Professor" />
</head>
<body>
    <par id="MakingOfABook">
        <seq>
        <video src="authorview.mpg" />
        <img src="onagoodday.jpg" />
        </seq>

        <audio src="authorview.wav" />
        <text src="http://www.cs.sfu.ca/mmbook/" />
    </par>
</body>
</smil>
```


1.4 Overview of Multimedia Software Tools

- The categories of software tools briefly examined here are:

  1. **Music Sequencing and Notation**
  2. **Digital Audio**
  3. **Graphics and Image Editing**
  4. **Video Editing**
  5. **Animation**
  6. **Multimedia Authoring**

Music Sequencing and Notation

- **Cakewalk**: now called Pro Audio.
  - The term **sequencer** comes from older devices that stored sequences of notes ("events", in MIDI).
  - It is also possible to insert WAV files and Windows MCI commands (for animation and video) into music tracks (MCI is a ubiquitous component of the Windows API.)
- **Cubase**: another sequencing/editing program, with capabilities similar to those of Cakewalk. It includes some digital audio editing tools.
- **Macromedia Soundedit**: mature program for creating audio for multimedia projects and the web that integrates well with other Macromedia products such as Flash and Director.

Digital Audio

- **Digital Audio** tools deal with accessing and editing the actual sampled sounds that make up audio:
  - **Cool Edit**: a very powerful and popular digital audio toolkit; emulates a professional audio studio — multitrack productions and sound file editing including digital signal processing effects.
  - **Sound Forge**: a sophisticated PC-based program for editing audio WAV files.
  - **Pro Tools**: a high-end integrated audio production and editing environment — MIDI creation and manipulation; powerful audio mixing, recording, and editing software.

Graphics and Image Editing

- **Adobe Illustrator**: a powerful publishing tool from Adobe. Uses vector graphics; graphics can be exported to Web.
- **Adobe Photoshop**: the standard in a graphics, image processing and manipulation tool.
  - Allows layers of images, graphics, and text that can be separately manipulated for maximum flexibility.
  - **Filter factory** permits creation of sophisticated lighting-effects filters.
- **Macromedia Fireworks**: software for making graphics specifically for the web.
- **Macromedia Freehand**: a text and web graphics editing tool that supports many bitmap formats such as GIF, PNG, and JPEG.

Video Editing

- **Adobe Premiere**: an intuitive, simple video editing tool for **nonlinear** editing, i.e., putting video clips into any order:
  - Video and audio are arranged in "tracks".
  - Provides a large number of video and audio tracks, superimpositions and virtual clips.
  - A large library of built-in transitions, filters and motions for clips $\Rightarrow$ effective multimedia productions with little effort.
- **Adobe After Effects**: a powerful video editing tool that enables users to add and change existing movies. Can add many effects: lighting, shadows, motion blurring; layers.
- **Final Cut Pro**: a video editing tool by Apple; Macintosh only.

Animation

- **Multimedia APIs**:
  - **Java3D**: API used by Java to construct and render 3D graphics, similar to the way in which the Java Media Framework is used for handling media files.
    1. Provides a basic set of object primitives (cube, splines, etc.) for building scenes.
    2. It is an abstraction layer built on top of OpenGL or DirectX (the user can select which).
  - **DirectX** : Windows API that supports video, images, audio and 3-D animation
  - **OpenGL**: the highly portable, most popular 3-D API.

- **Rendering Tools**:
  - **3D Studio Max**: rendering tool that includes a number of very high-end professional tools for character animation, game development, and visual effects production.
  - **Softimage XSI**: a powerful modeling, animation, and rendering package used for animation and special effects in films and games.
  - **Maya**: competing product to Softimage; as well, it is a complete modeling package.
  - **RenderMan**: rendering package created by Pixar.
- **GIF Animation Packages**: a simpler approach to animation, allows very quick development of effective small animations for the web.

Multimedia Authoring

- **Macromedia Flash**: allows users to create interactive movies by using the score metaphor, i.e., a timeline arranged in parallel event sequences.
- **Macromedia Director**: uses a movie metaphor to create interactive presentations — very powerful and includes a built-in scripting language, **Lingo**, that allows creation of complex interactive movies.
- **Authorware**: a mature, well-supported authoring product based on the **Iconic/Flow-control** metaphor.
- **Quest**: similar to Authorware in many ways, uses a type of flowcharting metaphor. However, the flowchart nodes can encapsulate information in a more abstract way (called **frames**) than simply subroutine levels.

1.5 Further Exploration

⟶ Link to Further Exploration for Chapter 1.

- In Chapter 1 of the Further Exploration directory, the website provides links to much of the history of multimedia.
- Other links in the text website include information on:
  - Ted Nelson and the Xanadu project.
  - Nicholas Negroponte's work at the MIT Media Lab.
  - Douglas Engelbart, and the history of the "On-Line System".
  - The MIT Media Lab
  - Client-side execution.