

# MULTIMEDIA 93

*SIGGRAPH 93  
20th International Conference  
on Computer Graphics and  
Interactive Techniques*

*Anahelm Convention Center  
1 to 6 August 1993*

COURSE NOTES 52

Multimedia and Multimodal Parsing

ORGANIZER/LECTURER  
Kent Wittenburg  
*Bellcore*



---

# *Multimedia and Multimodal Parsing*

**Kent Wittenburg  
Bellcore, MRE 2A-347  
445 South St  
Morristown, NJ 07962-1910  
email kentw@bellcore.com  
phone (201) 829-4382  
fax (201) 829-5981**

*Tutorial Notes  
ACM Multimedia '93  
Anaheim, California, August 3, 1993*

## *Abstract*

---

The problems of parsing and interpreting interface languages are generally well understood as long as the languages are unambiguous and consist of a single linear channel of information. Use of context and multiple information sources to interpret locally ambiguous input such as speech is an ongoing research topic. As new information channels and input devices arrive on the scene, the spectrum of possibilities for interface languages increases. Pen based gestural input along with handwriting and drawing provide one set of examples requiring extensions to string based parsing methods. Speech eye tracking and 3D devices such as a spaceball provide others, particularly when coupled with input from more than one medium in separate channels. This tutorial will consider the problem of characterizing multimodal expressions separately and concurrently as languages and then surveying techniques for parsing and interpreting them. The emphasis will be placed on those cases amenable to rule based grammatical approaches.

Bringing a parsing perspective to the domain of interpreting multimedia interface languages carries with it a set of assumptions about the decomposition of the problem into a number of separate but related subproblems, namely:

- segmentation of input into discrete symbol vocabularies
- grammatical representation of words and sentences composed from the symbols and relations between them
- mapping between syntax and semantics representations
- use of discourse models and other state information for context dependent interpretation

The tutorial will initially include a brief characterization of current research and practice in interpreting nonlinear and multimodal input along these lines. Then it will turn to a survey of grammatical frameworks including graph and array grammars as well as current research from the visual languages community. Various parsing methodologies will also be addressed, classified according to different application needs and different characteristics of the multimodal languages.

*Biography of Instructor*

---

Kent Wittenburg currently a member of technical staff in the Computer Graphics and Interactive Media research group at Bellcore received his Ph D from the University of Texas at Austin in the area of computational linguistics He was previously at the MCC Human Interface Lab in Austin Texas where he was project leader of the Interface Languages group and also on the faculty of the University of Texas at Austin His publications are in the areas of grammatical formalisms and parsing for natural languages visual languages and multimodal interfaces Current interests include extensions to grammatical models and parsing techniques in order to interpret pen-based drawing and handwritten math expressions and to parse multidimensional data in support of design visualization and data integrity tasks He is presently coordinator of the Association for Computational Linguistics special interest group on Multimedia Language Processing

*Table of Contents*

---

<b>Abstract</b>	<b>2</b>
<b>Biography of Instructor</b>	<b>3</b>
<b>Introduction</b>	<b>7</b>
<b>Problem Definitions</b>	<b>17</b>
<b>Survey of Applications</b>	<b>25</b>
<b>Theory and Methods</b>	<b>52</b>
<b>Case Studies Matching Parsers to Applications</b>	<b>83</b>
<b>Concluding Remarks</b>	<b>90</b>
<b>Multimedia and Multimodal Parsing</b>	
<b>A Bibliography</b>	<b>95</b>
<b>References</b>	<b>96</b>
<b>Subject Index</b>	<b>109</b>
<b>Article Reprints</b>	<b>111</b>