Planning Multimedia Explanations Using Communicative Acts

Mark T. Maybury

The MITRE Corporation
Artificial Intelligence Center
MS A047, Burlington Road
Bedford, MA 01730
maybury@linus.mitre.org

Abstract[1]

A number of researchers have investigated the use of plan-based approaches to generate textual explanations (e.g., Appelt 1985; Hovy 1988; Moore 1989; Maybury 1990b). This paper extends this approach to generate multimedia explanations by defining three types of communicative acts: linguistic acts (illocutionary and locutionary speech acts), visual acts (e.g., deictic acts), and media-independent rhetorical acts (e.g., identify, describe). This paper formalizes several of these communicative acts as operators in the library of a hierarchical planner. A computational implementation is described which uses these plan operators to compose route plans in coordinated natural language and graphics in the context of a cartographic information system.

Introduction

The notion of communication as an action-based endeavor dates to Austin's (1962) view of language as purposeful behavior. Searle (1969) extended this view with his formalization of speech acts. Bruce's (1975) suggestion of a plan-based model of speech acts was followed by computational investigations into planning speech acts (Cohen 1978), planning referring expressions (Appelt 1985), and planning multisentential text to achieve particular communicative goals (e.g., Hovy 1988; Moore 1989; Maybury 1990b).

Related research has focused on orchestrating text and graphics (Neal et al. 1989; Wahlster et al. 1989; Feiner & McKeown 1990; Hovy & Arens 1990; Roth, Mattis & Mesnard 1991) and developing media-independent representations of intensions (Elhadad et al. 1989). Related to this is research into the psychoperception of verbal and pictorial elements (Guastello & Traut 1989).

This paper extends this integration of language and graphics using the notion of communicative acts. We first define several communicative acts, including linguistic and visual ones, in a common plan operator language. Next, we use these operators to plan coordinated texts and graphics which identify objects and convey route plans from the Map Display System (Hilton & Anken 1990), a knowledge-based cartographic information system. A final section identifies limitations and areas for further research.

Multimedia as Communicative Acts

Multimedia communication can be viewed as consisting of linguistic and visual acts that, appropriately coordinated, can perform some communicative goal such as describing an object, narrating a sequence of events, or explaining how a complex process functions. For example, when giving directions on how to get from one place to another, if possible, humans will often utilize maps, gestures, and language to explain a route.

Just as humans communicate using multiple media (i.e., language, graphics, gestures) in multiple modes (i.e., language can be written or spoken), we have implemented an explanation planner that represents and reasons about multimedia *communicative acts* (see Figure 1). Communicative acts include rhetorical, linguistic, and visual acts as well as non-linguistic auditory acts (e.g., snap, ring) and physical acts (e.g., gestures). A *rhetorical act* is a sequence of linguistic or visual acts which are used to achieve certain media-independent rhetorical goals such as identifying an entity, describing it, dividing it into its subparts or subtypes, narrating events and situations (Maybury 1990b), and arguing to support a conclusion.

In contrast, a *linguistic act* is a speech act (Searle 1969) such as INFORM or REQUEST which characterizes the illocutionary force of a single utterance. These illocutionary speech acts can be accomplished by *locutionary* or *surface speech acts* (Appelt 1985) such as ASSERT, ASK, and COMMAND which are associated with particular grammatical structures (declarative, imperative, and interrogative mood, respectively). While illocutionary speech acts are useful for plan abstraction (e.g., a REQUEST can be achieved by asking, commanding, recommending, etc.), we focus here on locutionary acts.

In contrast to linguistic acts, *visual acts* include graphical deictic gestures (e.g., pointing, highlighting, blinking, circling), display control (e.g., zooming, panning), and image depiction. In the current implementation deictic gestures are considered primitive acts. In contrast, depiction can include depictions of primitive images (e.g., a point or line), composite images (e.g., a tree with arcs and nodes), and complex images (e.g., a picture of a location). Thus, depiction itself can be viewed as a plan-based endeavor (e.g., composing and rendering a bar graph) (see Feiner 1985; Burger 1989). The next section details several of these communicative acts for identifying locations.

RHETORICAL ACT		LINGUISTIC ACT	VISUAL ACT
		ILLOCUTIONARY ACT	*DEICTIC ACT*
Identify		inform	highlight, blink, circle etc.
Define		request	indicate direction
Describe		warn	
Detail		concede	*DISPLAY CONTROL ACT*
Divide			display-region
Illustrate		*LOCUTIONARY ACT*	zoom (in, out)
Compare		assert (declarative)	pan (left, right, up, down)
Narrate		ask (interrogative)	
Explain		command (imperative)	*DEPICT ACT*
Argue		recommend ("should")	depict image
		exclaim (exclamation)	draw (line, arc, circle)

Figure 1. Communicative Acts: Rhetorical, Linguistic, and Visual

Multimedia Plans for Location Identification

Similar to physical actions, communicative acts (rhetorical, linguistic, and visual) can be formalized as plans. Communicative acts are represented as operators in the plan library of a hierarchical planner (Sacerdoti 1977). Each plan operator defines the *constraints* and *preconditions* that must hold before a communicative act applies, its intended *effects* (also known as postconditions), and the refinement or *decomposition* of the act into subacts. Preconditions and constraints encode conditions concerning both physical states (e.g., is an object too large to be displayed) as well as cognitive states (e.g., does the hearer believe some proposition). Constraints, unlike preconditions, cannot be achieved or planned for if they are false. The decomposition of a plan operator defines how higher level communicative acts (e.g., describing an object) are divisible into potentially coordinated lower level actions (e.g., describing it in natural language, depicting an image of it, or both).

NAME	Identify-location-linguistically
HEADER	Identify(*S*, *H*, *entity*)
CONSTRAINTS	Entity?(*entity*)
PRECONDITIONS	WANT(*S*, KNOW(*H*, Location(*entity*)))
EFFECTS	KNOW(*H*, Location(*entity*))
DECOMPOSITION	Assert(*S*, *H*, Location(*entity*))

Figure 2. Uninstantiated Linguistic Plan Operator

For example, the uninstantiated Identify-location-linguistically plan operator shown in Figure 2 is one of several methods of performing the communicative action Identify. As defined in the HEADER of the plan operator, the Identify act takes three arguments, the speaker (S), the hearer (H), and an entity. The English translation of Figure 2 is as follows: Provided the third argument is indeed an entity[2] (CONSTRAINTS) and the speaker wants the hearer to know about it (PRECONDITIONS), the speaker (S) will identify the location of the entity by informing the hearer (H) of its location (DECOMPOSITION), which has the intended effect that the hearer knows about it (EFFECTS).

[2]An entity is an object or event (e.g., a process or an action).

Plan operators are encoded in an extension of first order predicate calculus which allows for optionality within the decomposition. Predicates (which have true/false values (e.g., Entity?)), functions (which return values), and communicative acts (e.g., Identify, Assert, Blink) appear in lower-case type with their initial letter capitalized. Arguments to predicates, functions, and communicative acts include variables and constants. Variables are italicized (e.g., *S*, *H*, and *entity*) and constants appear in upper-case plain type.

Intensional operators, such as WANT, KNOW, and BELIEVE appear in capitals. KNOW details an agent's specific knowledge of the truth-values of propositions (e.g., KNOW(H, Red(ROBIN-1)) or KNOW(H, ¬Yellow(ROBIN-1))) where truth or falsity is defined by the propositions in the knowledge base. That is, KNOW(H, P) implies P ∧ BELIEVE(H, P). Agents can hold an invalid beliefs (e.g., BELIEVE(JOHN, Yellow(ROBIN-1))). KNOW-ABOUT is a predicate that is an abstraction of a set of epistemic attitudes of some agent toward an individual. An agent can KNOW-ABOUT an object or event (e.g., KNOW-ABOUT(H, ROBIN-1) or KNOW-ABOUT(H, EXPLOSION-445)) if they KNOW its characteristics, components, subtypes, or purpose. KNOW-HOW indicates an agent's ability to perform an action.

If the object we are identifying has an associated visual presentation in the backend cartographic display, we can augment natural language with visual identification. The Identify-location-linguistically-&-visually plan operator in Figure 3 is selected only if its constraints are satisfied (i.e., the given entity is a cartographic entity such as a town, road, lake, etc.). If these constraints are satisfied, the plan operator then ensures that the entity is visible. If the designated entity is out of the currently visible region or too small to be seen, this can be achieved by either panning, jumping, or zooming to the region around the designated entity. For example, Figure 4 illustrates the map display action, Make-entity-visible, which displays the region surrounding a given entity. Note that the precondition of this plan operator will ensure the entity is displayed. If it is not already displayed on the map, this will be planned for.

NAME	Identify-location-linguistically-&-visually
HEADER	Identify(*S, H, entity*)
CONSTRAINTS	Cartographic-Entity?(*entity*)
PRECONDITIONS	Visible(*entity*) ∧
	WANT(*S*, KNOW(*H*, Location(*entity*)))
EFFECTS	KNOW(*H*, Location(*entity*))
DECOMPOSITION	Indicate-Deictically(*S, H, entity*)
	Assert(*S, H*, Location(*entity*))

Figure 3. Plan Operator for Graphical/Textual Display

NAME	Make-entity-visible
HEADER	Make-Visible(*entity*)
CONSTRAINTS	Cartographic-Entity?(*entity*)
PRECONDITIONS	Displayed(*entity*)
EFFECTS	Visible(*entity*)
DECOMPOSITION	Display-Region(*entity*)

Figure 4 . Plan Operator for Map Display Control

After the entity is visible, the decomposition of the identify action of Figure 3 deictically indicates the entity and then describes its location in natural language (as above). There are several plan operators for deictic indication available including highlighting (a permanent indication of an entity), blinking (intermittent highlighting), and circling. These forms of visual deixis can be used to indicate individual objects (e.g., roads, towns, dams), groups of objects, or geographical regions. While the current implementation simply defaults to highlighting, the choice among different deictic techniques could be motivated by a number of considerations including the number and kind of entities visible in the region, their visual properties (e.g., size, color, shading) in order to maximize the distinction of the given entity and its background, and the kind of communication being generated (e.g., highlighting may be preferred when communicating route plans so that upon completion the entire route is visible). We next illustrate these plans in action.

Multimedia Identification Exemplified

To illustrate these and other communicative acts, we detail the planning of multimedia directions for the Map Display System (Hilton & Anken 1990), a knowledge-based cartographic information system which represents over 600 German towns, 227 NATO airbases, 40 lakes, 14 dams, as well as other objects. The road network in the map includes 233 roads (divided up into 4,607 road segments) and 889 intersections.

If the user queries the system "Where is Karl-Marx-Stadt?," this is simulated by posting the goal Identify(SYSTEM, USER, #<Karl-Marx-Stadt>) to the explanation planner. The planner then uses a unification algorithm to find all operators from the library whose HEADER portion matches the current goal. This includes the identification plan operators in Figures 2 and 3. Next all operators whose header matches this goal are found and instantiated with the bindings of the variables that match the header. Figure 5 shows the plan operator for linguistic and visual identification instantiated with bindings. When the action Identify(SYSTEM, USER, #<Karl-Marx-Stadt>) unifies against the header of the plan operator in Figure 3, the variable S is bound to SYSTEM, H is bound to USER, and *entity* is bound to the object #<Karl-Marx-Stadt>. These bindings are used to instantiate the entire plan operator to that shown in Figure 5.

Because there may be many methods of achieving a given goal, those operators that satisfy the constraints and essential preconditions are then prioritized using *preference metrics*. For example, operators that utilize both text and graphics are preferred over simply textual operators. Also, those operators with fewer subgoals are preferred (where this does not conflict with the previous preference). The preference metric prefers plan operators with fewer subplans (cognitive economy), with fewer new variables (limiting the introduction of new entities in the focus space of the discourse), those that satisfy all preconditions (to avoid backward chaining for efficiency), and those plan operators that are more common or preferred in naturally-occurring explanations (e.g., certain kinds of communicative acts occur more frequently in human-produced text or are preferred by rhetoricians over other methods). While the first three preferences are explicitly inferred, the last preference is implemented by the sequence in which operators appear in the plan library.

Working from this prioritized list of operators, the planner ensures preconditions are satisfied and tries to execute the decomposition of each until one succeeds. This involves processing any special operators (e.g., optionality is allowed in the decomposition) or quantifiers (∀ or ∃) as well as distinguishing between subgoals and primitive acts. For example, if the planner chooses the plan operator in Figure 5 from those that satisfy their constraints, it first ensures its preconditions hold (i.e., by making sure the entity is visible through other visual acts).

NAME	Identify-location-linguistically-&-visually
HEADER	Identify(SYSTEM, USER, #<Karl-Marx-Stadt>)
CONSTRAINTS	Cartographic-Entity?(#<Karl-Marx-Stadt>)
PRECONDITIONS	Visible(#<Karl-Marx-Stadt>) ∧
	WANT(SYSTEM, KNOW(USER, Location(#<Karl-Marx-Stadt>)))
EFFECTS	KNOW(USER, Location(#<Karl-Marx-Stadt>))
DECOMPOSITION	Indicate-Deictically(SYSTEM, USER, #<Karl-Marx-Stadt>)
	Assert(SYSTEM, USER, Location(#<Karl-Marx-Stadt>))

Figure 5. Instantiated identify Plan Operator

Next, the planner attempts to execute the two subacts in its decomposition, `Indicate-Deictically(SYSTEM, USER, #<Karl-Marx-Stadt>)` and `Assert(SYSTEM, USER, Location(#<Karl-Marx-Stadt>)`. `Assert` is a primitive act and so decomposition halts here. In contrast, `Indicate-Deictically` is not a primitive act and so the planner is reinvoked. As indicated in the previous section, in the current implementation deictic indication defaults to highlighting, which is also a primitive act.
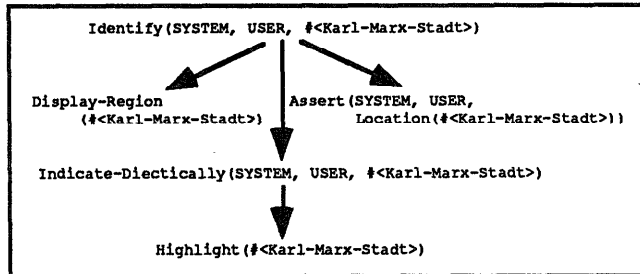

Figure 6. Hierarchical Multimedia Plan to identify Karl-Marx-Stadt

Thus, our original simulated user query, "Where is Karl-Marx-Stadt?", results in the hierarchical decomposition shown in Figure 6. This tree is linearized by a depth-first search and the resulting sequence of linguistic and visual primitive actions is executed. The surface speech act, `Assert`, together with the `Location` predicate and its argument, `#<Karl-Marx-Stadt>`, are passed to the linguistic realization component. Using this information, the realizer fills a semantic case role associated with the `Location` predicate to yield a semantic specification `#<Assert location-predicate Karl-Marx-Stadt>`, which contains the following information:

```
ACTION:     #<be-copula>
AGENT:      #<Karl-Marx-Stadt>
PATIENT:    #<town>
MODIFIERS:
  (location (latitude 50.82) (longitude 12.88))
```

This specification is used to build grammatical relations (subject, object), then syntactic constituents (noun, verb, adverbial, and prepositional phrases), and finally a surface tree which is realized as (see Maybury, forthcoming):

Karl-Marx-Stadt is a town located at 50.82° latitude 12.88° longitude.

This is uttered after the map displays the region around Karl-Marx-Stadt and highlights its icon.

Extended Multimedia Directions

While the above visual and linguistic identification of Karl-Marx-Stadt may satisfy the user's query, often a cartographic information system must communicate a route between distant points. This can be accomplished in

```
NAME      Explain-route-linguistically-and-visually
HEADER    Explain-Route(S, H, from-entity, to-entity)
CONSTRAINTS
  Cartographic-entity?(from-entity) ∧
  Cartographic-entity?(to-entity) ∧ path
PRECONDITIONS
  visible(from-entity) ∧
  WANT(S, KNOW-HOW(H, Go(from-entity, to-entity)))
EFFECTS
  KNOW-HOW(H, Go(from-entity, to-entity)) ∧
  ∀segment ∈ path
      KNOW(H, Subpath(segment, path))
DECOMPOSITION
  ∀segment ∈ path
      Indicate-Deictically(S, H, source(segment ))
      Command(S, H, Do(H, Go(source(segment ),
                              link(segment),
                              destination(segment ))))
      Indicate-Deictically(S, H, link(segment ))
      Indicate-Direction(S, H, source(segment ),
                              link(segment),
                              destination(segment ))
  Identify(S, H, to-entity)
WHERE  path = cartographic-path(from-entity, to-entity)
```

Figure 7. `Explain-Route` Plan Operator

language alone, or by coordinating language and map displays. The communicative act `Explain-Route`, formalized in Figure 7, does the latter, the former being a simplification thereof. The constraints of this operator first test if both objects are cartographic ones and that there exists a path between them in the underlying Map Display System (Hilton & Anken 1990). The function `cartographic-path` which is used in the plan operator takes as arguments two objects from the cartographic knowledge base and, using a branch and bound search strategy, explores the road network to return the "best" route between the two points (if one exists). The path returned by the function is an ordered list of roads, intersections, and towns indicating the preferred route from one entity to another, as defined by the rewrite rules:

```
path      ->  segment + (path)
segment   ->  point + road-segment + point
point     ->  intersection | city | town | bridge
```

where "()" indicates optionality and "|" indicates logical disjunction. For any given segment, the functions `source`, `link`, and `destination` return the source and destination point and the link that connects them (i.e., a road segment).

If the constraints on the `Explain-Route` action are satisfied, then the planner attempts to achieve its preconditions. The first precondition requires the source location to be visible. If not currently the case, this can be achieved using visual actions like the make-visible act defined in Figure 4. If the constraints and preconditions can be satisfied, then the decomposition first visually identifies the source of the next segment, next linguistically requests the hearer to move from the source to the destination of that

segment, then visually identifies the link of the next segment, and lastly visually indicates the direction of the movement along the link between the two (using an arrow). (The initial source location is not *linguistically* identified because we assume the hearer is travelling from that location and thus is familiar with it.) After repeating this for all segments, the plan concludes by identifying the ultimate destination using actions like those of Figures 2 and 3. The effect of explaining the route is that the hearer knows how to get from origin to destination and the hearer knows the segments of that route.
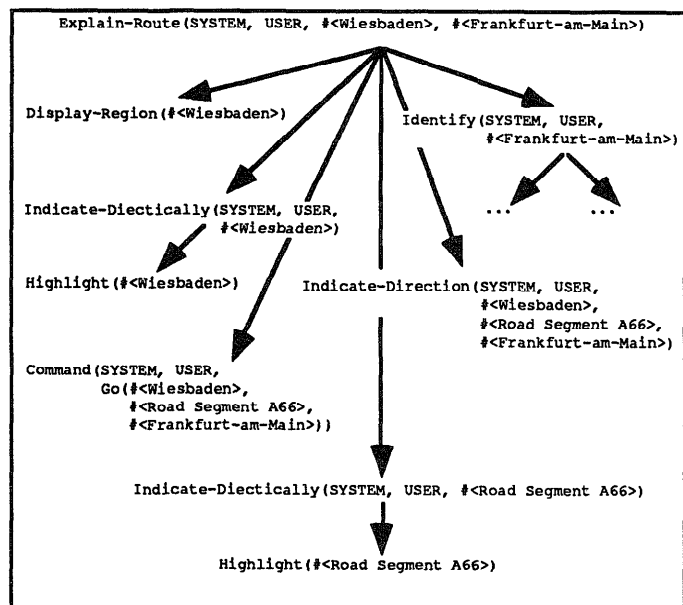

Figure 8. Hierarchical Plan for Locational Instructions

For example, assume the user asks "How do I get from Wiesbaden to Frankfurt?", simulated by posting the discourse goal `Explain-Route(SYSTEM, USER, #<Wiesbaden>, #<Frankfurt-am-Main>)`. The planner uses the `Explain-Route` act of Figure 7 to build the explanation plan shown in Figure 8. This plan is realized as (visual acts indicated parenthetically in italics):

(Display map region around Wiesbaden) (highlight Wiesbaden) From Wiesbaden take Autobahn A66 Northeast for thirty-one kilometers to Frankfurt-am-Main. *(highlight Autobahn A66) (indicate direction with blinking arrow) (highlight Frankfurt-am-Main)* Frankfurt-am-Main is located at 50.11° latitude and 8.66° longitude.

A slightly more complex locational instruction results if the user asks how to get from Mannheim to Heidelberg, initiated by posting the discourse goal `Explain-Route(SYSTEM, USER, #<Mannheim>, #<Heidelberg>)`. The resulting multimedia explanation is realized as:

(Display map region around Mannheim) (highlight Mannheim) From Mannheim take Route 38 Southeast for four kilometers to the intersection of Route 38 and Autobahn A5. *(highlight Route 38) (indicate direction with blinking arrow) (highlight intersection of Route 38 and Autobahn A5)* From there take Autobahn A5 Southeast for seven kilometers to Heidelberg. *(highlight Autobahn A5) (indicate direction with blinking arrow) (highlight Heidelberg)* Heidelberg is located at 49.39° latitude and 6.68° longitude, four kilometers Northwest of Dossenheim, six kilometers Northwest of Edingen, and five kilometers Southwest of Eppelheim.

The linguistic realization component keeps track of the relationship of the current *spatial focus* (the current visited segment) to the previous spatial focus (the previously visited segment). This relationship constrains the choice of surface choices (Maybury 1990b) such as demonstrative pronouns ("this" versus "that"; "here" versus "there") as well as the generation of spatial directionals (e.g., "Southeast", "West") and durationals (e.g., "seven kilometers"). This focus-based choice contrasts with the use of heuristic approaches based on rules (e.g., describe an entity using a demonstrative noun phrase if there is no proper name for that entity (Neal et al. 1989)).

Conclusion and Future Directions

This paper proposes a number of communicative acts -- linguistic, visual, and rhetorical -- that can be exploited to plan and coordinate multimedia explanations. We first formalize several linguistic acts and visual acts as plan operators. These are abstracted into higher level, media-independent actions called rhetorical acts. A computational implementation is described which identifies locations and composes route plans in coordinated natural language text and graphics in the context of a cartographic information system.

We are currently extending the implementation to incorporate other types of visual acts. For example, the system is able to divide an entity linguistically in two ways: by detailing its constituents or subparts (e.g., "The United Kingdom contains England, Scotland, Wales, and Northern Ireland.") or if the entity is an abstract concept, by indicating its subtypes or subclasses (e.g., "There are three Baltic languages: Old Prussian, Lithuanian, and Latvian."). Visually, subpart division can be accomplished, for example, by depicting subcomponents or hierarchical trees.

Similarly, subtype division can be accomplished visually using trees (which indicate parent/child relations) or Venn diagrams (indicating set relationships). Also, while the system can linguistically characterize an entity (e.g., "The pancreas is a long, soft, irregular shaped gland located behind the stomach."), entities which have visual attributes such as size, shape, color and location can be depicted, perhaps with greater effect than the corresponding linguistic description. Finally, the system can generate paragraph-length comparisons of entities, and we intend to compose tabular comparisons of attributes and values, although this will require planning of more sophisticated composite graphs (Feiner 1985; Burger 1989). Other composite visual

acts also require further investigation (e.g., circling a group of objects and indicating their movement with an arrow).

There are several issues which require further investigation. These include the relationship of deictic and display control acts to the model of the user's attention (i.e., salient objects, events, and regions). Another important issue concerns coordinating visual and linguistic acts at the clausal and lexical level (e.g., referring expressions coordinated with deixis). One approach would be to extend paragraph planning below the sentence level (Appelt 1985). Much more difficult is how to narrate events and situations in multiple media, which requires communication of temporal, spatial, and causal information (i.e., story telling coupled with visual animation).

Finally, we need to investigate the relation of linguistic and visual acts to other non-speech audio acts. For example, there are analogs between mediums such as linguistic, visual, and auditory warnings (exclaiming, flashing, and beeping), visual and auditory icons (e.g., using sirens to indicate danger), and graphical and auditory motion (e.g., using the perception of Doppler effects to indicate motion). These remain interesting avenues for future research.

Acknowledgements

References

Austin, J. 1962. *How to do Things with Words.* editor J. O. Urmson. England: Oxford Univ. Press.

Appelt, D. 1985. *Planning English Sentences.* England: Cambridge Univ. Press.

Burger, J. 1989. User Models for Intelligent Interfaces. IJCAI-89 Workshop: A New Generation of Intelligent Interfaces, Detroit, MI, August 22, 1989. 17-20.

Bruce, B. C. 1975. Generation as a Social Action. Proceedings of Theoretical Issues on Natural Language Processing (ACL), 64-67.

Cohen, P. R. 1978. On Knowing What to Say: Planning Speech Acts. Univ. of Toronto TR-118.

Elhadad, M.; Seligmann, D. D.; Feiner, S.; and McKeown, K. 1989. A Common Intentional Description Language for Interactive Multi-Media Systems. IJCAI-89 Workshop: A New Generation of Intelligent Interfaces, Detroit, MI, August 22, 1989. 46-52.

Feiner, S. November, 1985. APEX: An Experiment in the Automated Creation of Pictorial Explanations. *IEEE Computer Graphics and Application* 5(11):29-37.

Feiner, S. and McKeown, K. 1990. Coordinating Text and Graphics in Explanation Generation. Proceedings of AAAI-90, Boston, MA. 442-449.

Geller, J. July, 1988. A Knowledge Representation Theory for Natural Language Graphics. Ph.D. diss., State Univ. of New York at Buffalo.

Guastello, S. J. and Traut, M. 1989. Verbal versus pictorial representations of objects in a human-computer interface. *International Journal of Man-Machine Studies* 31(1989):99-120.

Hilton, M. L. and Anken, C. S. February, 1990. Map Display System: An Object-Oriented Design and Implementation. Rome Air Development Center Technical Report 90-54.

Hovy, E. 1988. Planning Coherent Multisentential Text. Proceedings of the 26th Annual Meeting of the ACL, Buffalo, NY, June 7-10, 1988. 163-169.

Hovy, E. and Arens, Y. 1990. When is a Picture Worth a Thousand Words? -- Allocation of Modalities in Multimedia Communication. AAAI Spring Symposium on Human-Computer Communication, Stanford, CA, March, 1990.

Maybury, M. 1990a. Custom Explanations: Exploiting User Models to Plan Multisentential Text. Proceedings of the Second International Workshop on User Models, Univ. of Honolulu, Hawaii, 30 March - 1 April, 1990.

Maybury, M. 1990b. Using Discourse Focus, Temporal Focus, and Spatial Focus to Plan Narrative Text. Proceedings of the 5th International Workshop on Natural Language Generation, Dawson, PA, 3-6 June, 1990.

Maybury, M. forthcoming. Planning Multisentential English Text Using Communicative Acts. Ph.D. diss., Univ. of Cambridge, England.

McKeown, K. 1985. *Text Generation.* Cambridge Univ. Press.

Moore, J. D. November, 1989. A Reactive Approach to Explanation in Expert and Advice-Giving Systems. Ph.D. diss., Univ. of California at Los Angeles.

Neal, J. G.; Thielman, C. Y.; Dobes, Z.; Haller, S. M.; and Shapiro, S. C. 1989. Natural Language with Integrated Deictic and Graphic Gestures. *Proc. of the 1989 DARPA Workshop on Speech and Natural Language,* Harwich Port, MA, Morgan Kaufmann. 410-423.

Roth, S.; Mattis, J.; and Mesnard X. 1991. Graphics and Natural Language Generation as Components of Automatic Explanation. To appear in J. Sullivan and S. Tyler (eds.) *Architectures for Intelligent Interfaces: Elements and Prototypes,* Reading, Mass: Addison-Wesley.

Sacerdoti, E. D. 1977. *A Structure for Plans and Behavior.* New York: Elsevier North-Holland.

Searle, J. R. 1969. *Speech Acts.* Cambridge Univ. Press.

Wahlster, W.; Jameson, A.; and Hoeppner, W. 1978. Glancing, Referring and Explaining in the Dialogue System HAM-RPM. *Computational Linguistics* (microfiche 77):53-67.

Wahlster, W.; André, E.; Hecking, M.; and Rist, T. May, 1989. Knowledge-based Presentation of Information (WIP). Project Proposal.