

Towards Intelligent Collaborative Assistants

George Ferguson

**James Allen, Nate Blaylock, Donna Byron, Dave Costello,
Myroslava Dzikovska, Lucian Galescu, Peter Heeman, Marc Light,
Nat Martin, Paul McCarty, Brad Miller, Massimo Poesio, Eric
Ringer, Len Schubert, Amanda Stent, David Traum**

Computer Science Department

University of Rochester

Research Goals

- ◆ **Intelligent computer assistants that collaborate naturally with people**
- ◆ **Understand how people communicate**
 - **Build computers that “speak their language”**
- ◆ **Understand how people solve problems**
 - **Build computer assistants to help them**
- ◆ **New paradigms for human-computer interaction**

Human-Human Interaction

- ◆ **There is a long history of studying how people communicate**
 - linguistics, philosophy, psychology, physiology, etc.
- ◆ **People are highly optimized (physically and cognitively) for face-to-face verbal interaction**

Human-Computer Interaction

- ◆ **Traditionally hard to use**
 - Paper tape, punch cards, keyboards
 - Programming
 - Microsoft Office, etc.
- ◆ **Spend more time figuring out how to interact than actually interacting (ie., getting anything done)**

Technology for HCI

- ◆ **Recent advances in computer technology are opening up new ways of communicating with machines**
 - **Fast CPUs, huge memories**
 - **Speech recognition and generation**
 - **Real-time video I/O**
 - **Virtual reality**
 - **Haptics**

Our Approach

- ◆ **Make computers more like people**
 - **Conversational interaction**
 - ✦ **Unrestricted spoken natural language dialogue**
 - ✦ **Multimodal augmentation**
 - **Collaborative knowledge-based problem solving**
 - ✦ **Human in the loop (mixed-initiative)**
 - ✦ **Using, adapting, and developing new AI technology**

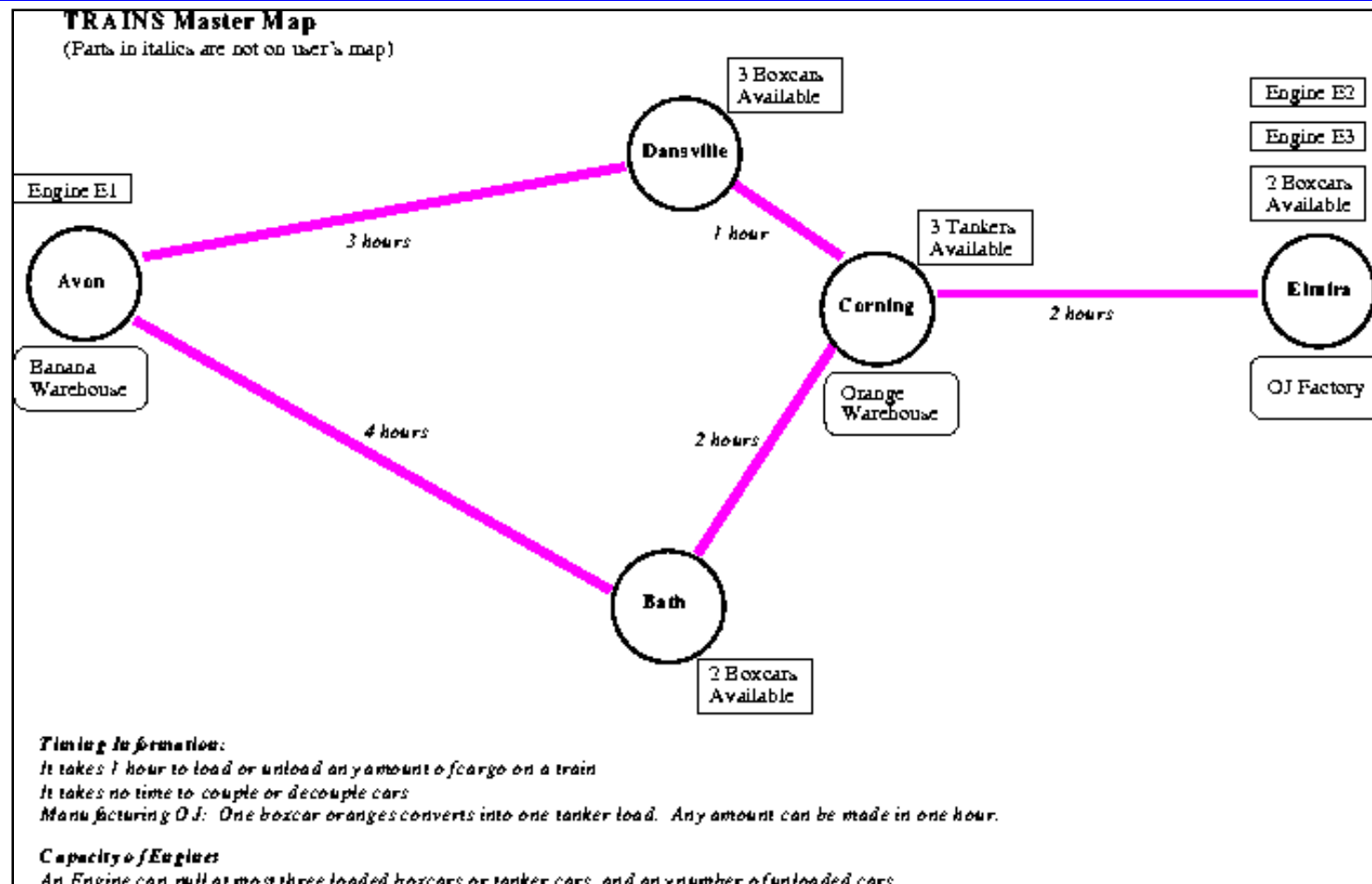
Towards Intelligent Collaborative Assistants

A Brief History

Human-Human Studies

- ◆ **Two people communicating via headsets**
 - **Common (but not shared) map**
- ◆ **Solving simple logistics problems in a transportation domain**
- ◆ **“Wizard of Oz” setup:**
 - **One person plays role of user**
 - **One person plays role of system**

TRAINS Dialogue Example



TRAINS Dialogue Corpus

- ◆ **Recording, transcription, annotation [Heeman 93]**
- ◆ **98 dialogues, 20 tasks, 34 speakers**
- ◆ **6.5 hours, 5900 turns, 55000 words**

First Steps From HHI To HCI

- ◆ **TDC dialogues are complex:**
 - **Complicated utterances (user & system)**
 - **Smart, perceptive, helpful “system”**
- ◆ **Abstract from data to simpler forms**
 - **Simpler language (in & out)**
 - **Less effective, less comprehensive, less perceptive problem solving**
 - **Attempt to indicate system’s limitations**

The TRAINS System

- ◆ **Initial prototype (1993-1995) of a collaborative planning assistant**
- ◆ **Led by Allen and Schubert**
- ◆ **Application of thesis work (Poesio, Traum, Ferguson) and others**
- ◆ **Feasibility demonstration**

TRAINS Map



TRAINS Shortcomings

- ◆ **Emphasis on written NLU**
 - **Speech recognition wasn't feasible!**
- ◆ **Not robust**
 - **Poor linguistic coverage**
 - **Little attempt to understand non-sentences**
 - **Limited reasoning capabilities**
- ◆ **Monolithic design**

TRIPS: The Rochester Interactive Planning System

TRIPS: Objectives

- ◆ **Robustness**
 - To both user and system errors
- ◆ **Multimodal interaction**
 - Speech and graphics
- ◆ **Incremental development and evaluation**
 - Scalable domain (from trivial to impossible)
- ◆ **Modular, agent-oriented design**

TRIPS

THE ROCHESTER INTERACTIVE PLANNING SYSTEM

**DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ROCHESTER**

TRIPS Example Summary

- ◆ **Intuitive interaction (conversational spoken language, GUI)**
- ◆ **Dialogue provides context for interpreting user inputs**
- ◆ **User control of problem solving process (top-down, bottom-up, what-if, revision)**
- ◆ **Fast, flexible specialized reasoners**

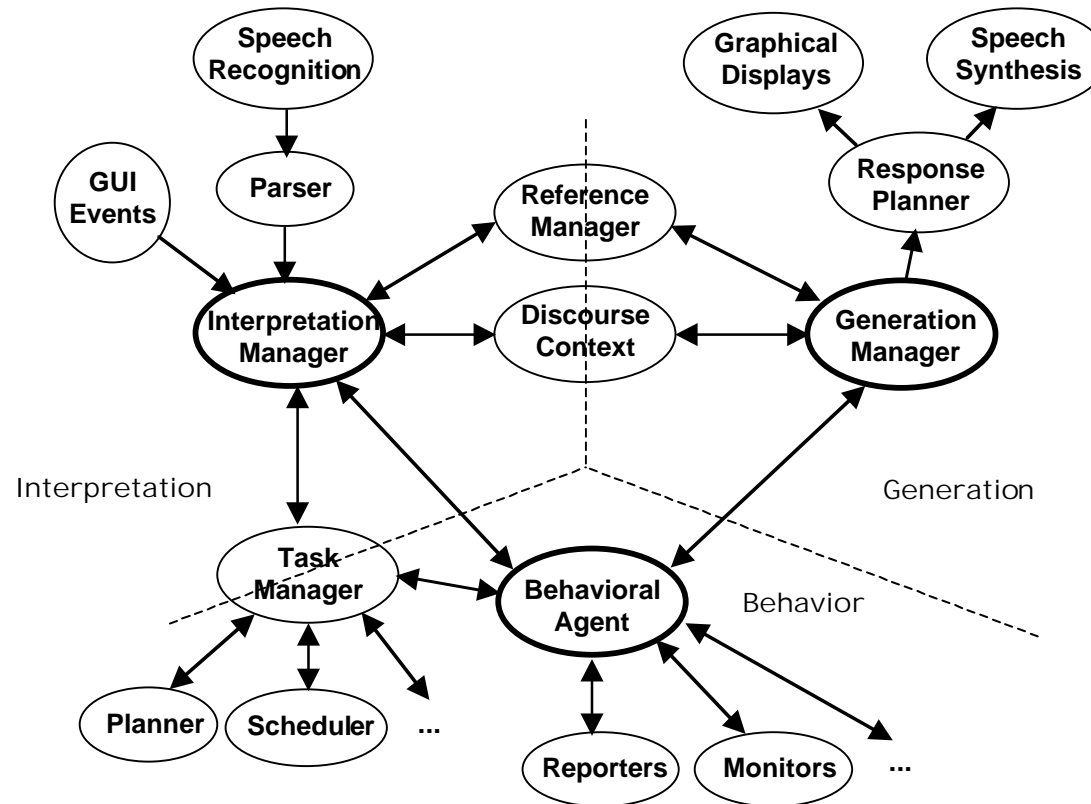
TRIPS:

Design and Implementation

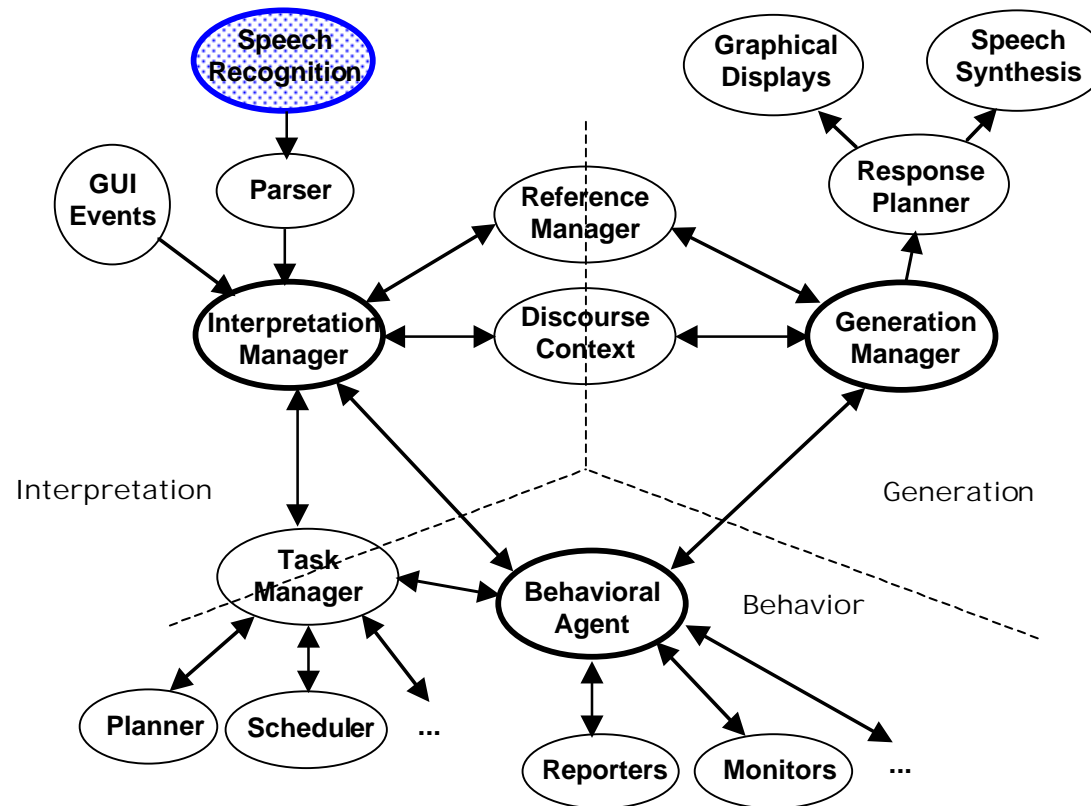
TRIPS: Functional Description

- ◆ **Distributed, heterogeneous, agent-oriented components**
- ◆ **Communicate using KQML message-passing**
- ◆ **Infrastructure supports: broadcast, content-based addressing, resource discovery and monitoring, replay, ...**

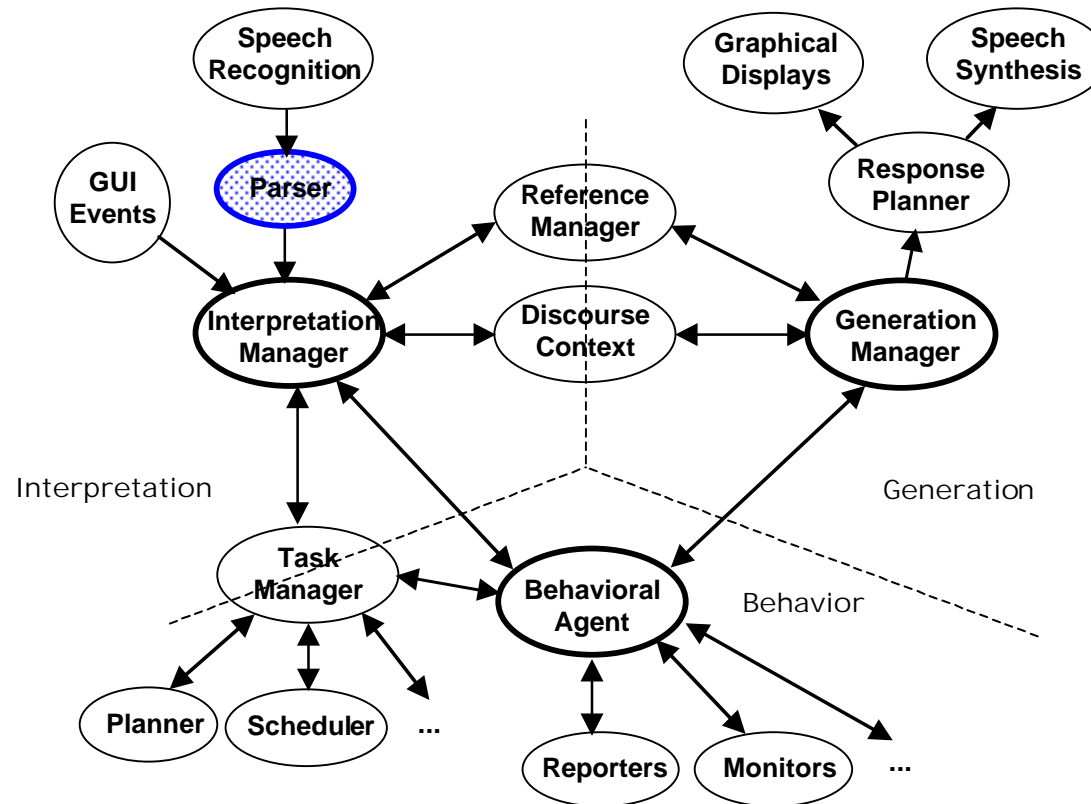
TRIPS Architecture



TRIPS Architecture: Speech Recognition



TRIPS Architecture: Parsing



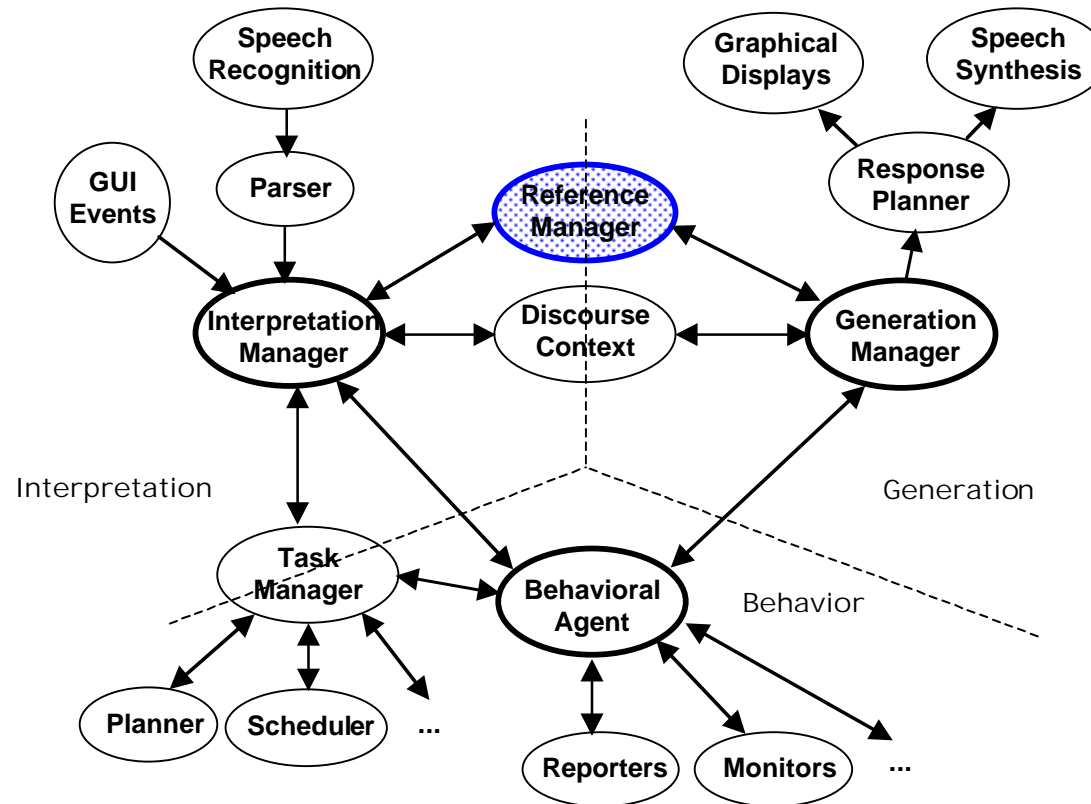
Natural Language Understanding

- ◆ **Syntactic analysis using incremental chart parser [Allen NLU 2nd ed.]**
- ◆ **Semantics attached to rules that add constituents to the chart**
- ◆ **Logical form in terms of (possible) communicative acts**

Parsing Challenges (Myrosia Dzikovska)

- ◆ **Managing ambiguity**
 - **“I saw a bird with binoculars”**
 - **Use knowledge about how objects are related**
- ◆ **“Knowledge bottleneck”**
 - **Hand-coded: labor-intensive**
 - **Automatically acquired: poor coverage**
 - **Hybrid: use learning to augment a hand-coded core of knowledge**

TRIPS Architecture: Reference Manager



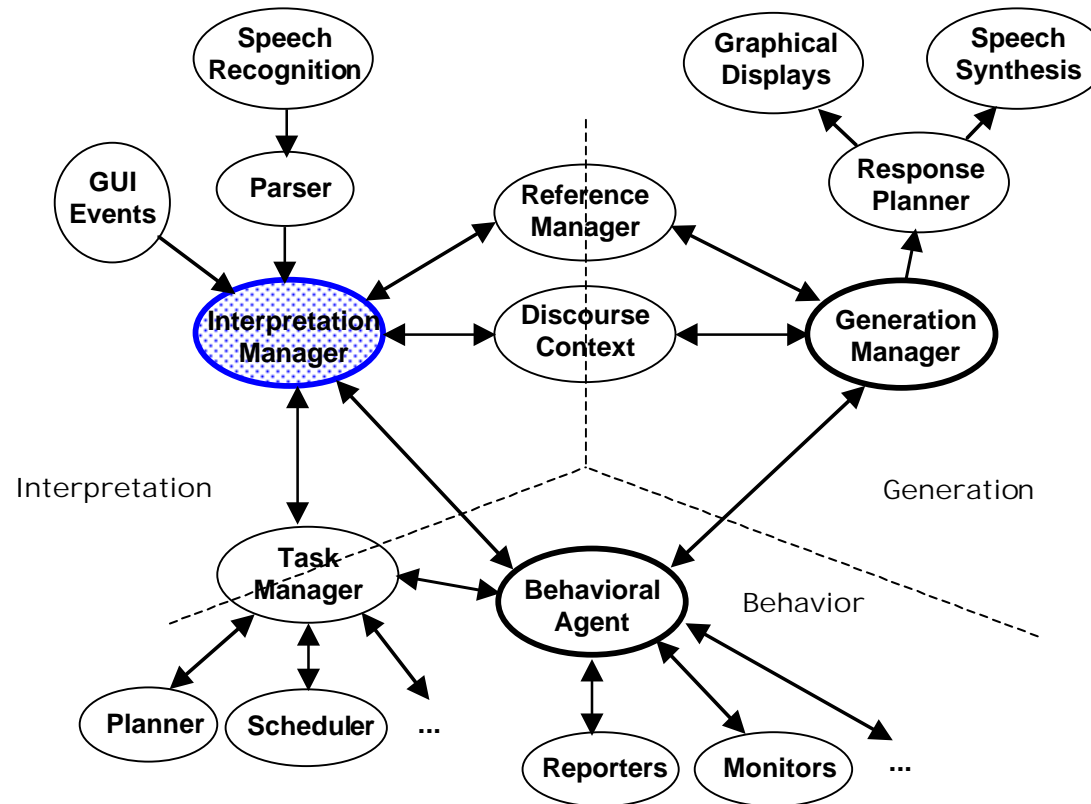
Definite Reference

- ◆ People use indexical terms in language for efficiency
- ◆ Pronouns: “it”, “they”, “that”, ...
- ◆ Definite descriptions: “the truck”, “the other truck”, “the people we picked up in Abyss”, ...

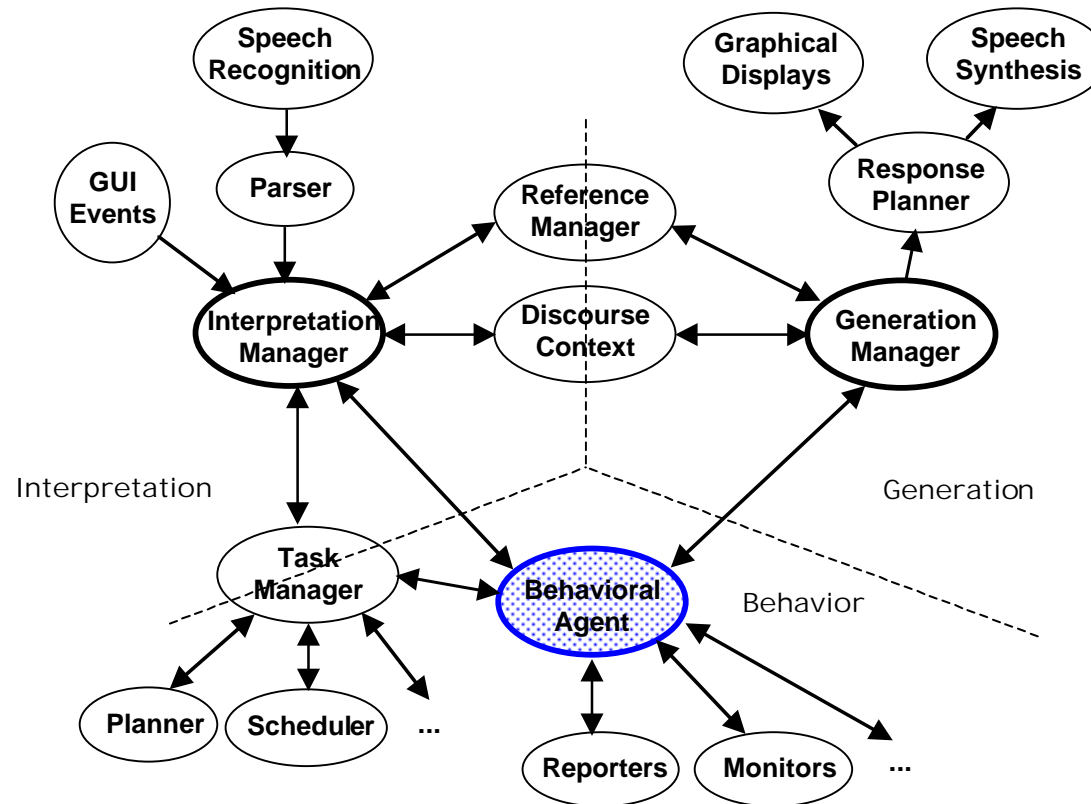
Reference Resolution (Donna Byron)

- ◆ **Resolve pronominal referents by searching an ordered set of semantically-rich salience contexts**
 - **Not simply prior linguistic mention**
 - **Handle abstract entities**
 - ✦ **“The other truck is closer to Abyss.”**
“Does that help at all?”
 - ✦ **“Do we have both trucks going to Calypso?”**
“No, I don’t think that’s happening.”

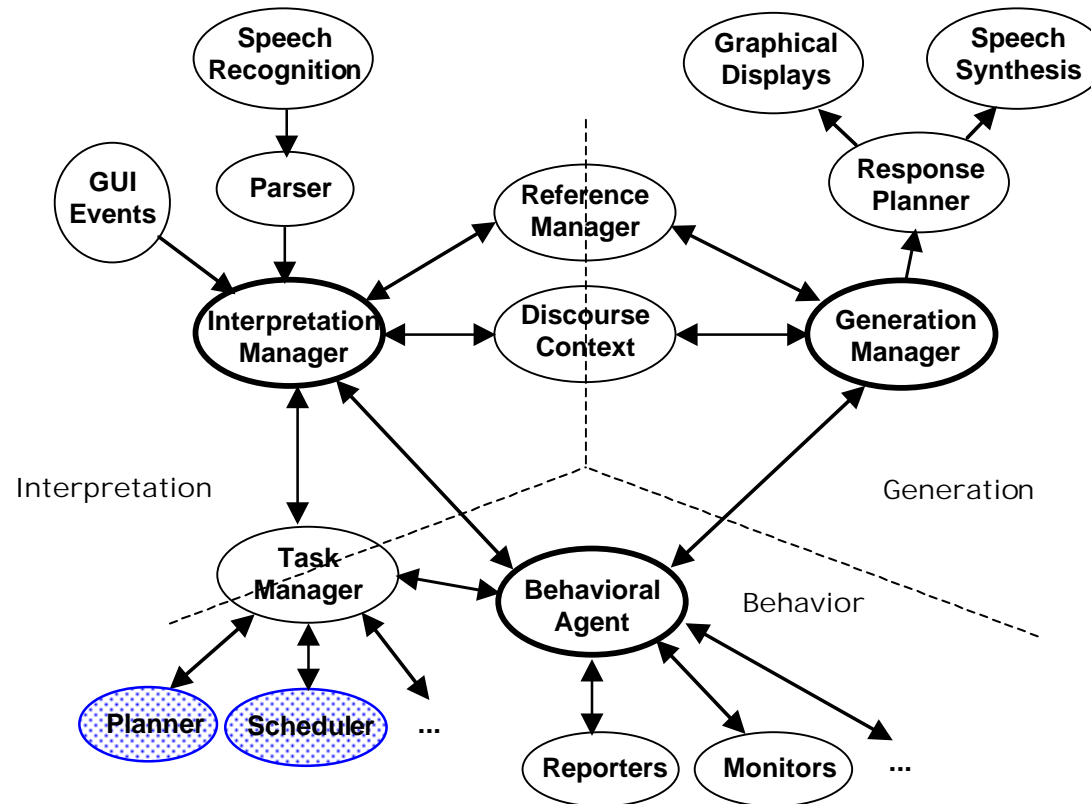
TRIPS Architecture: Interpretation Manager



TRIPS Architecture: Behavioral Agent



TRIPS Architecture: Reasoners



TRIPS Planner

- ◆ **Hybrid model trading completeness for efficiency**
 - **Domain-dependent HTN-like task decomposition**
 - **Domain-independent underlying reasoner based on interval temporal logic**

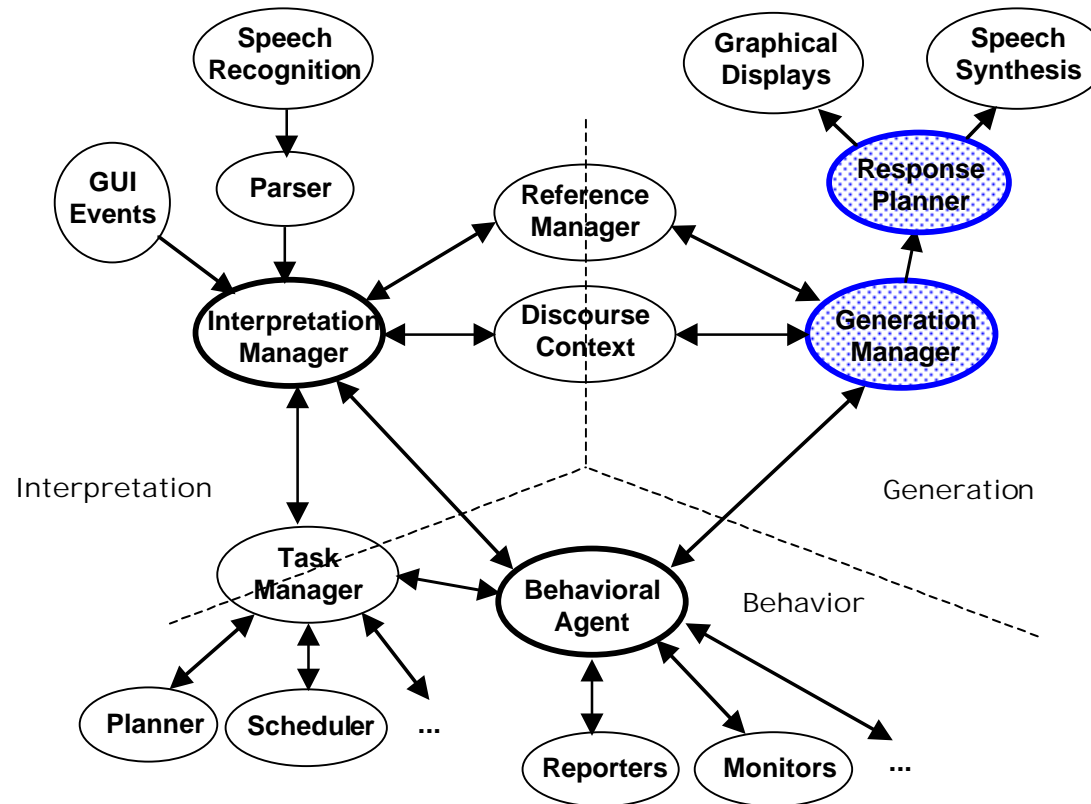
TRIPS Router

- ◆ **Path planner based on Dijkstra's algorithm**
- ◆ **Supports complex constraint specifications (e.g. "via Calypso," "along the coast," "avoiding highways," "as fast as possible")**

TRIPS Scheduler

- ◆ **Assigns nominal start and end times to actions in plan based on interval temporal constraints**
- ◆ **Known NP-complete problem (currently use $O(N^3)$ approximation)**
- ◆ **Crucial to have concrete course of action (“straw plan”) for presentation to user**

TRIPS Architecture: Generation



Generation for Conversational Systems

- ◆ **Goal: Have system generate natural utterances in conversation**
 - **Handles rhetorical structure, speech acts, grounding acts and turn-taking behaviors in human-like ways**
 - **Permits flexible task and discourse initiative**
 - **Based on study of real human-human dialogue (rather than text)**

Generation Architecture (Amanda Stent)

- ◆ **Generation Manager (GM)**
 - Plans discourse contributions for different modalities
- ◆ **Response Planner (RP)**
 - Turn-taking
 - Grounding
 - Speech acts
- ◆ **Multi-modal coordinator**

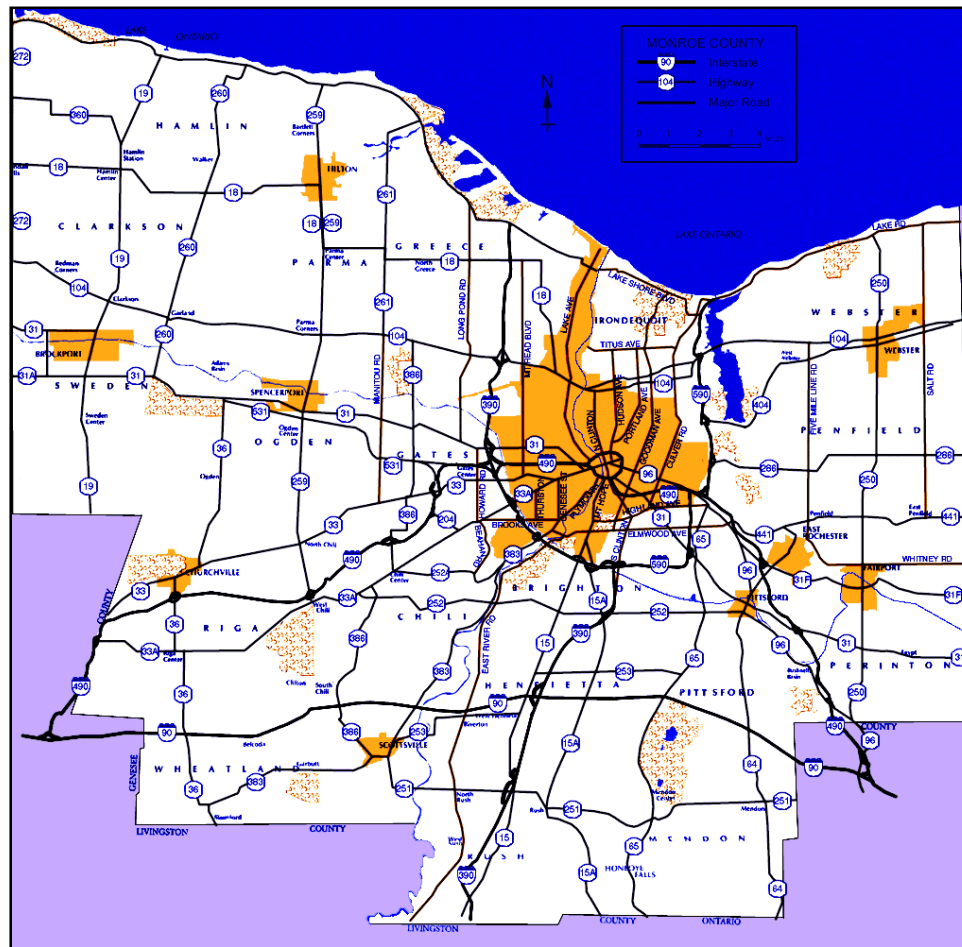
TRIPS Summary

- ◆ **TRIPS is a “system of systems”**
 - **Input modalities**
 - **Language understanding**
 - **Intention recognition**
 - **Collaborative problem-Solving**
 - **Knowledge-based reasoners**
 - **Multi-level generation**
 - **Multi-modal synthesis**

Future Directions for TRIPS

- ◆ **Improved robustness**
- ◆ **More complex domains (Monroe County)**
- ◆ **Dynamic world (online planning)**
- ◆ **Use of simulation**
- ◆ **Visual avatar to enhance interaction**
- ◆ **CAVE: Conversational Agents in Visual Environments**

TRIPS Monroe County



Towards Intelligent Collaborative Assistants

- ◆ **Robust understanding for effective communication**
 - Recognition of user intention
- ◆ **Support for human problem-solving**
 - Expressive representations
 - Incremental reasoning
- ◆ **Natural, incremental generation**
- ◆ **Modular system architecture**

End

ferguson@cs.rochester.edu

<http://www.cs.rochester.edu/research/trips>