gency dispatcher, cooperating with the system to dynamically allocate resources to and make plans for solving problems as they arise in the world. The setting, Monroe County, NY, is considerably more complex than our previous domains (e.g. Pacifica, TRAINS), and raises new issues in knowledge representation and reference. Emergencies include requests for medical assistance, car accidents, civil disorder, and larger problems such as flooding and snow storms. Resources at the user's disposal may include road crews, electric crews, ambulances, police units and helicopters. Some of the increase in mixed-initiative interaction comes from giving the system more knowledge of the tasks being solved. Some comes from the fact that the solution to one problem may conflict with the solution to another, either because of scheduling conflicts, scarce resources, or aspects of the physical world (e.g. an ambulance can't go down a road that has not been plowed). The range of tasks and complexity of the world allow for problem solving at different levels of granularity, making it possible for the system to take as much control over the task as the user permits.

4. Important Contributions

While a number of robust dialogue systems have been built in recent years, they mostly have operated in domains that require little if any reasoning. Rather, the task is hard-coded into the system operation. One of the major goals of the TRIPS project has been to develop dialogue models and system architectures that support conversational interaction in domains where complex reasoning systems are required. One goal has been to build a fairly generic model in which different domains can then be specified fairly easily. On this front, we are seeing some success as we have now constructed versions of TRIPS in three different domains, and TRIPS-911 will be the fourth. In developing the system for new domains, the bulk of the work by far has been in system enhancements rather than in developing the domain models.

The TRIPS-911 domain has forced a rethinking of the relationship between dialogue-management, problem-solving, the system's
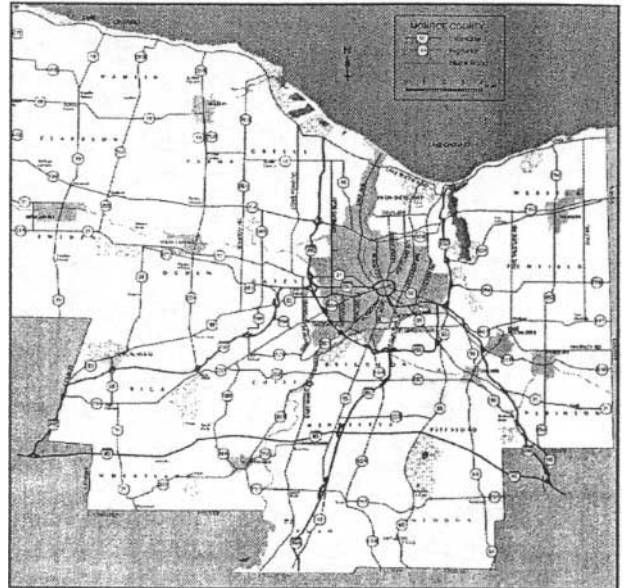

Figure 1: Monroe County map used in TRIPS-911

own goal-pursuit and generation. The new architecture is designed to support research into mixed-initiative interactions, incremental generation of content (in which the user might intervene before the system completes all it has to say), rich reference resolution models, and the introduction of plan monitoring and plan repair into the suite of plan management operations supported. The domain also can support longer and richer dialogues than in previous domains.

More complex domains mean even more complex dialogues. The complexity arises from many factors. First, more complex dialogues will involve topic progression, development and resumption, and more complex referential phenomena. On the problem solving front, there will be more complex corrections, elaborations and modifications—forcing us to develop richer discourse models. In addition, the complexity of the domain demonstrates a need for better grounding behavior and a need for incremental dialogue-based generation.

We have by no means solved these problems. Rather we have built a rich testbed, designed and implemented a plausible architecture, and have constructed an initial system to demonstrate basic capabilities in each of the problem areas.

5. Limitations

TRIPS-911 is a first attempt at handling a domain of this complexity. As such there are many capabilities that people have in such situations that are beyond the system's current capabilities. Some of the most important are:

- Scale – we can only handle small domains and the existing techniques would not extend directly to a realistic size 911 operation. To scale up we must face some difficult problems including reasoning about quantities and aggregates, planning in large-scale domains (i.e., the real domains are beyond the capabilities of current plan technology), and performing intention recognition as the number of options increases. In addition, for an effective dialogue system, all this must be done in real-time.

- Meta-talk – when faced with complex problems, people often first generally discuss the problem and possible strategies for solving it, and later may explicitly direct attention to specific subproblems. The current TRIPS system does not support such discussion.

- Time – in the 911 domain there are at least two temporal contexts that can be "used" by the conversants: there is the actual time (i.e., when they are talking), but there also is the time relative to a point of focus in a plan, or even simply talking about the past or the future. TRIPS-911 can currently interpret expressions with respect to the actual time.

- Interleaved generation – when people are discussing complex issues, they often have to plan to communicate their content across several different utterances. There is no guarantee that the other conversant will not "interrupt" (e.g., to clarify, correct, suggest alternatives, etc) before the entire content is conveyed. This requires a rethinking of current practice in generation to make it incremental and interactive.

- True interruptions - people may interrupt the system while it is talking. It is unclear at this stage what the system should assume was conveyed. The strategies of assuming nothing was conveyed, or that all was conveyed have obvious faults. We are pursuing alternatives based on knowing when speech was interrupted, but using this information successfully remains a difficult problem.

References

Allen, James et al, An Architecture for a Generic Dialogue Shell, to appear, *J. Natural Language Engineering*, 2000.

Ferguson, George and J. Allen, TRIPS: An Integrated Intelligent Problem-Solving Assistant, *Proc. National Conference on AI (AAAI-98)*, Madison, WI, 1998.