Social Tag: Finding the Person with the Pink Hat

Carl DiSalvo and **Didac Font** and **Laura Hiatt** and
Nik Melchior and **Marek Michalowski** and **Reid Simmons**
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15217

At the AAAI 2005 Robot Exhibition, the robot GRACE (Graduate Robot Attending a ConferencE, Fig. 1) will be playing a game that involves human-robot social interaction, navigation, and interface design. The task is for Grace to locate and rendezvous with one of our team members, who will be wearing a pink hat. The game can be seen as a social version of "tag" or "Marco Polo," where the the robot finds the target not through the modalities of sight or sound, but rather through social interactions with strangers in the environment.

The task has four phases (Fig. 2), which are repeated until the pink hat is located visually:

1. *Identification* of approachable humans.

2. *Approach* toward a human with whom the robot would like to interact.

3. *Asking* for directions to the person with the pink hat.

4. *Following* those directions.

In typical robot tasks involving detection of a visual target, distinctively colored objects such as pink hats are used to simplify the vision problem. However, in this case, the pink hat is as much for the benefit of other people as for Grace herself. The team member should be a prominent individual who is easily recognized and remembered by conference participants, so that it will be easy for them to help a wandering robot. Accordingly, it is not the completion of the goal (finding the pink hat) in which we are most interested, but rather in Grace's journey and her social interactions along the way.

The phases of the task, as mentioned above, are described here.

1. **Identification.** Grace is equipped with a laser scanner near human knee-height and a camera near human face-height. The laser scanner clusters short range readings, labels those that appear to be human beings, and tracks those humans over time using a Kalman filter. The camera locates faces using appearance-based frontal face detectors and tracks them using skin color models. Data from these two sensors is combined to determine more reliably

Figure 1: The robot GRACE.

where there are people, and to determine whether a person is approachable based on the direction they are facing and their proximity to other people.

2. **Approach.** Grace moves toward the person and attempts to begin a conversation. She is equipped with an animated face and a text-to-speech engine that generates audible speech to greet the person she is approaching. During this approach, she must be especially careful to observe societal norms such as speed, direction of approach, and personal space. Without these behaviors, people who are not accustomed to interacting with robots may not realize that the robot is attempting to engage them, and they may move away from the robot rather than stand in her path.

3. **Asking.** Grace first verbally asks whether the person is able to help. She is equipped with a touch-screen that simultaneously displays the request, along with Yes/No buttons. If No is pressed, the robot thanks the person and looks for someone else. If Yes is pressed, Grace asks where she should go, and the touch-screen displays arrows that the person can press to indicate a good direction in which Grace should continue her search for the person in the pink hat.

4. **Following.** Grace follows the suggested direction for some minimum distance, avoiding obstacles found by the
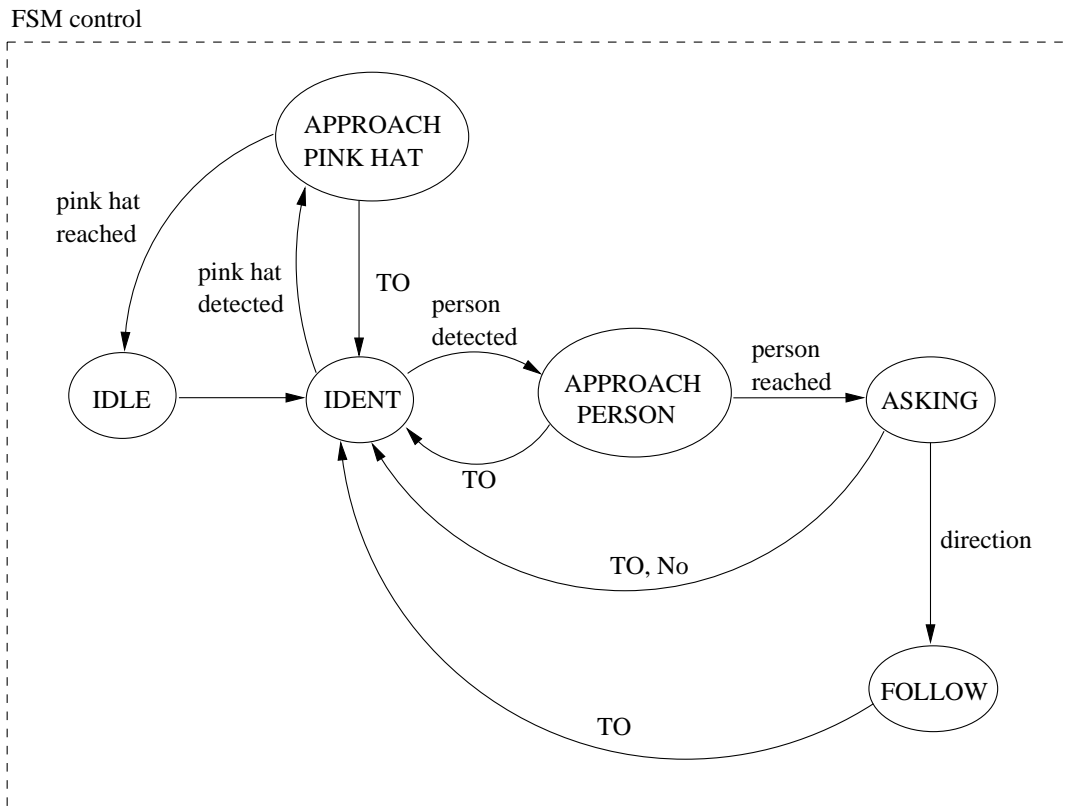
Figure 2: The finite state machine that comprises Grace's control task for Social Tag.

laser scanner. Meanwhile, she is looking for the pink hat. If it is not found after traversing some distance, she returns to the identification phase and looks for another person to approach.

This task explores issues in human-robot interaction that involve shared space and the negotiation of an environment filled with dynamic, untrained humans. The sensory abilities of robots are severely limited compared to those of humans. While research in computer vision and speech recognition continues to improve robot senses, Grace relies on the assistance of humans with fully developed senses of sight and hearing. Grace depends on her own sense of vision to complete this task, but her primary mode of gathering information is by asking people for help in the most intuitive and socially acceptable manner possible.

Additionally, at points during the conference, Grace will be stationary and will act as an information kiosk with a keyboard interface. This functionality is derived from Grace's sister robot, Valerie the Roboceptionist (http://www.roboceptionist.com).