USC Institute for Creative Technologies

University of Southern California

Intelligent Agents for Virtual Simulation of Human-Robot Interaction

Ning Wang, David V. Pynadath, Unni K.V., Santosh Shankar, Chirag Merchant

August 6, 2015

The work depicted here was sponsored by the U.S. Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

ARL

Background

- Increased capability of automated systems ≠ increased capability of human-machine teams

- One critical aspect of human-machine interaction is trust
  - Robot > human → robot performs the task
  - Robot < human → human performs the task
  - Lack of/over trust: Disuse and Misuse (Parasuraman & Riley, 1997)

- Hand-crafted explanation improves transparency and leads to trust (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003)

Robot Explanation and Testbeds

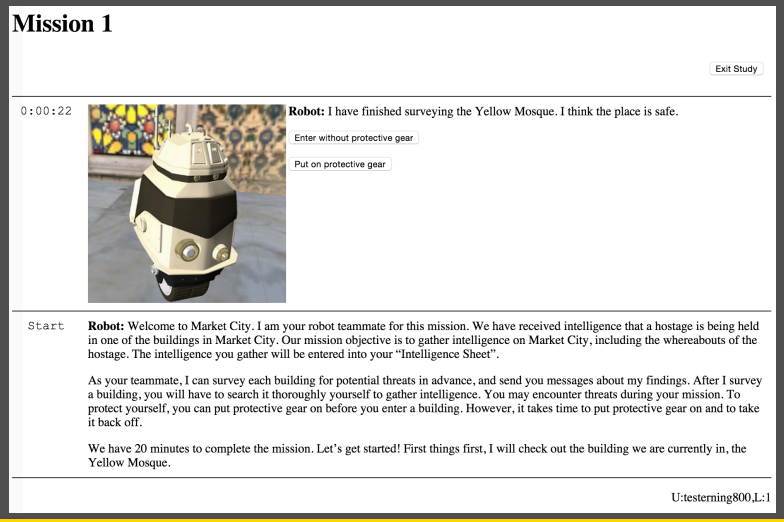
- Transparency about…
- Three aspects of trust: **Ability, Benevolence and Integrity** (Mayer, Davis, & Schoorman, 1995)
- Robot's abilities
  - Observe its surroundings
  - Understand the team/teammate's goals
  - Make decisions based on observations and goals

- Research testbeds to study the design of automatically generated robot explanation to influence trust in human-robot teams

Human-Robot Interaction Testbeds

- Unity 3D:

- HTML:



Testbed Requirements for HRI Trust Research

- Encourage human and robot to work together as a team
- Human working along-side the robot, not tele-operators
- Humans are assigned with their own tasks
- Encourage (verbal) communication between humans & robot
- Sources of distrust
- Behavioral measures of trust
- Cost/risks

Testbed Scenario for HRI Trust Research

Requirements

- Encourage human and robot to work together as a team
- Human working along-side the robot, not tele-operators
- Humans are assigned with their own tasks
- Encourage (verbal) communication between humans & robot
- Sources of distrust
- Behavioral measures of trust
- Cost/risks

Scenario

- Human teammate and Robot joint reconnaissance mission
- Robot serves as advanced scout for potential danger
- Relay its findings to the human teammate
- Teammate takes action based on robot's findings

Decision Theoretic Framework for Robot Modeling

- PsychSim – Social simulation framework
  - Recursive modeling gives agents a Theory of Mind
  - Boundedly rational, decision-theoretic reasoning
  - Agents generate behavior from explicit goals and beliefs
  - Authors can directly inspect and edit models
  - Agents can generate explanations of their behavior
- Robot as a PsychSim Agent
  - Observe the world
  - Form *beliefs* based on *observations*
  - Reason about what *actions* to choose in order to achieve its *goals* based on the *observations*
- PsychSim agent's observations, goals and reasoning process as basis for robot's explanations

Robot as a PsychSim Agent

PsychSim: Social Simulation Framework from ICT

- **State, *S:*** True state of the world
  - e.g. a gunmen is present in the building
- **Observations, *Ω:*** observations of the world,
  - e.g. readings from robot sensors.
  - **Obs. function, *O*** uses probability of receiving an accurate/inaccurate observation to simulate robot's noisy sensor readings
- **Actions, *A*:** Possible decisions
  - e.g. declare a building safe/dangerous
- **Transition Probability, *P*:** Effects of actions
  - e.g. declaring a building safe when it's not injures human teammate
- **Reward, *R*:** Quantitative model of goals
  - Keep human teammate unharmed
  - Minimize time to complete mission

Robot as a PsychSim Agent

PsychSim: Social Simulation Framework from ICT

- **Beliefs:** Incorporate **Observations, *Ω*** (sensor readings) and **Obs. function, *O*** (noisy sensor model) to form probabilistic beliefs about true **State, *S***

- **Policy, *π*:** Robot's decision-making using scenario-independent **POMDP** (Partially Observable Markov Decision Problem) algorithms
  - Consider effects of each **Action, *A***
  - Starting from current Beliefs about **State, *S***
  - Projected through **Transition Probability, *P***
  - Choose ***A*** that maxes resulting **Reward, *R***

Robot's Explanations Automatically Generated from…

PsychSim Agent's Decision-making process

To build trust with human teammate

Trust

Ability → ***Ω***: Does the robot get accurate observations?

Explanation: "My sensors have detected traces of dangerous chemicals."

Benevolence → ***R***: Does the robot value the teammate's life correctly?

Explanation: "I think it will be dangerous for you to enter the building without protective gear."

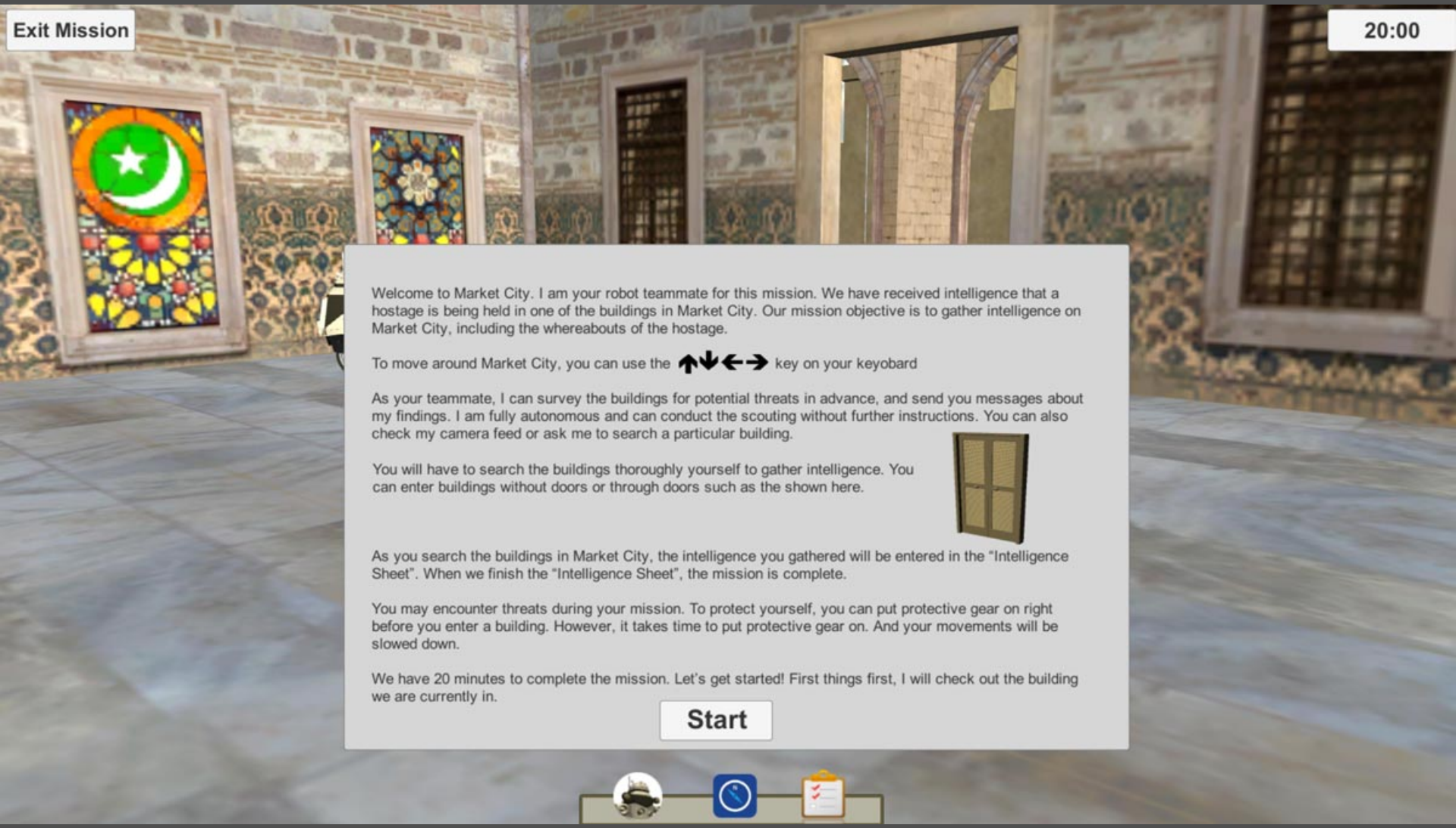
Integrity

Example Explanation Text

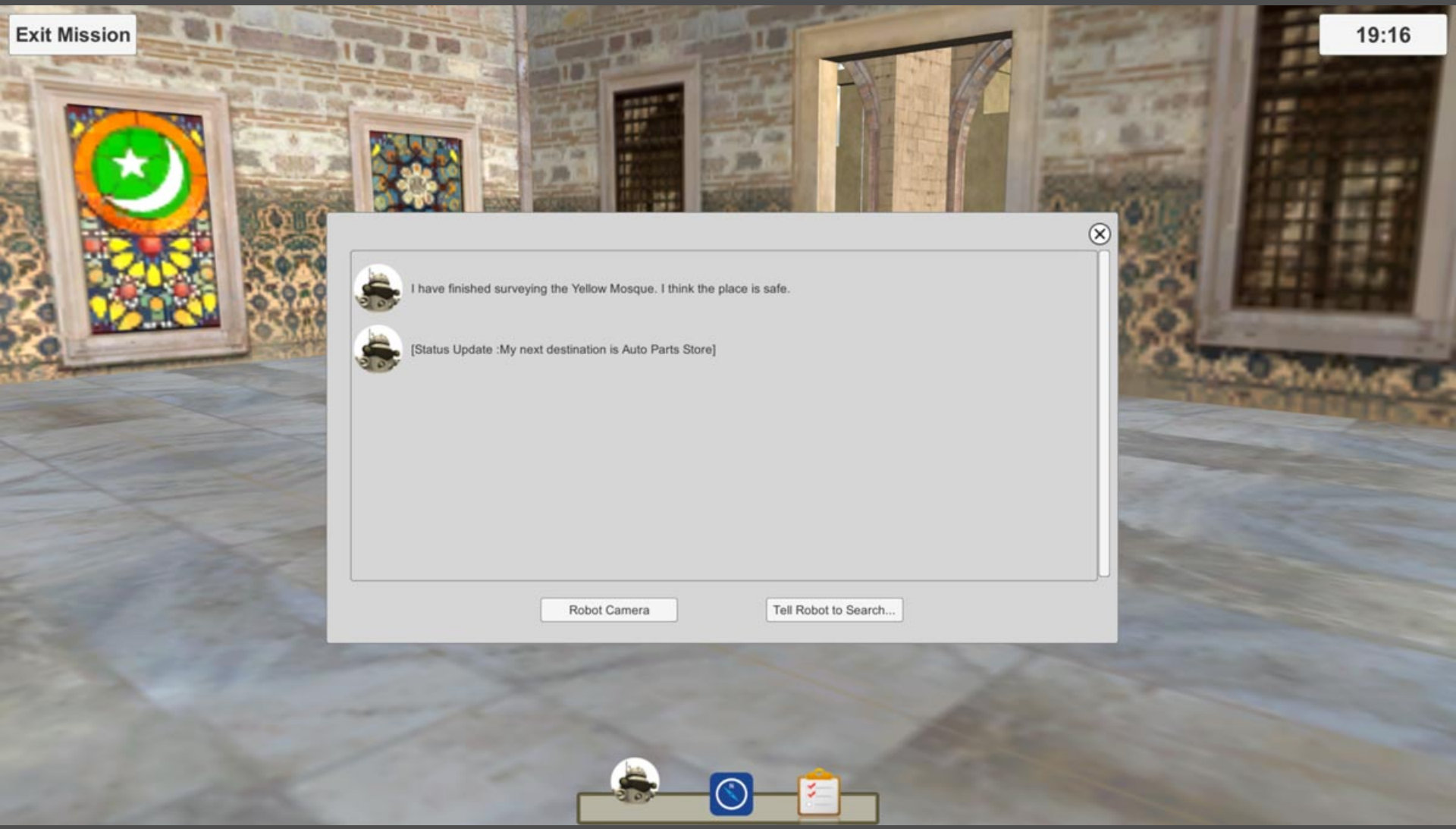
- “I have finished surveying the Cafe. I think the place is dangerous. My sensors have detected traces of dangerous chemicals. From the image captured by my camera, I have not detected any armed gunmen in the Cafe. I think it will be dangerous for you to enter the Cafe without protective gear. The protective gear will slow you down a little.”

- Explanations that impacts *ability*, *benevolence* and *integrity* aspects of the trust

Immersive Unity-based Online HRI Testbed: Instructions of Joint Reconnaissance Mission



Immersive Unity-based Online HRI Testbed:
Robot's Message

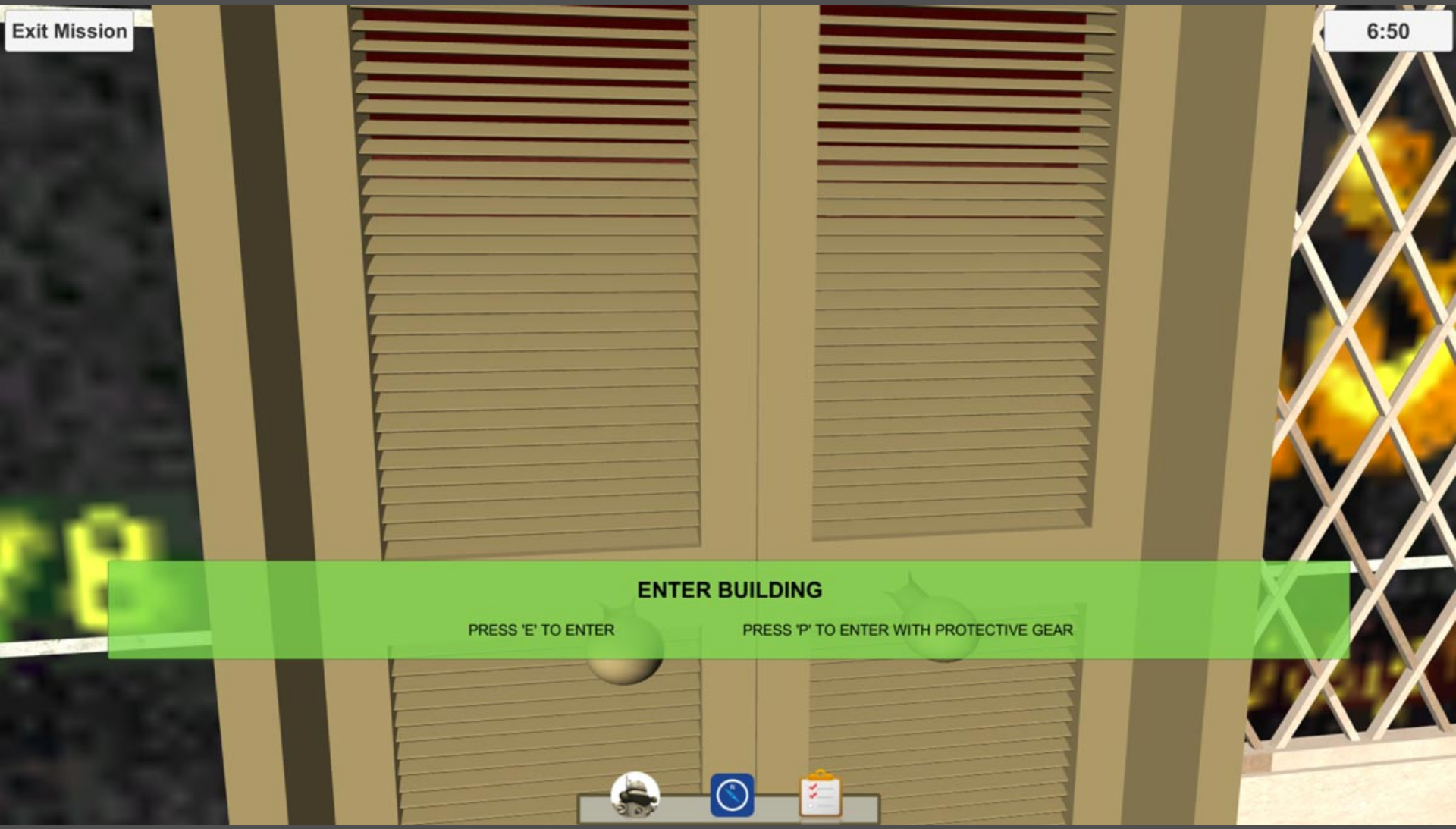


Immersive Unity-based Online HRI Testbed: Robot's Camera

Immersive Unity-based Online HRI Testbed:
Map of City

Immersive Unity-based Online HRI Testbed: Intelligence Sheet

Immersive Unity-based Online HRI Testbed:
Entering a building



Immersive Unity-based Online HRI Testbed:
Equipping Protective Gear

Immersive Unity-based Online HRI Testbed:
Safe building without protective gear

Immersive Unity-based Online HRI Testbed:
Same Building with Protective Gear

Immersive Unity-based Online HRI Testbed:
Enter dangerous building without protective gear

Agile HTML-based Online HRI Testbed

Instructions Screen

Mission 1

Exit Study

0:00:22

Robot: I have finished surveying the Yellow Mosque. I think the place is safe.

Enter without protective gear

Put on protective gear

Start

Robot: Welcome to Market City. I am your robot teammate for this mission. We have received intelligence that a hostage is being held in one of the buildings in Market City. Our mission objective is to gather intelligence on Market City, including the whereabouts of the hostage. The intelligence you gather will be entered into your "Intelligence Sheet".

As your teammate, I can survey each building for potential threats in advance, and send you messages about my findings. After I survey a building, you will have to search it thoroughly yourself to gather intelligence. You may encounter threats during your mission. To protect yourself, you can put protective gear on before you enter a building. However, it takes time to put protective gear on and to take it back off.

We have 20 minutes to complete the mission. Let's get started! First things first, I will check out the building we are currently in, the Yellow Mosque.

U:testerning800,L:1

Testbed Demo

- HTML
- Unity 3D

Discussion and Future Work

- Focus on research, balance between simulation and game
- Since publication…
  - User study with over 200 AMT participants with HTML testbed
  - Pilot studies with cadets with Unity 3D testbed
  - Explanations of robot's sensing capability, confidence level
  - Robot's of varied ability/reliability
- Future work
  - Explanation about decision-making process
  - Repair trust
  - Physical robot
- Collaborations

More info: nwang@ict.usc.edu

USC Institute for Creative Technologies

ARL

University of Southern California

Pilot studies with Unity 3D Testbed

- Six cadets from West Point
- Robot ability: high vs. low
  - High ability: always make correct decisions
  - Low ability: occasionally encounter errors in observing the environment resulting in incorrect decisions (robot decision-making process is intact)
  - Note: in low ability condition, we fixed the error on camera failure

- Robot explanation: decision followed by two types of explanations
  - "I have finished surveying the Cafe. I think the place is dangerous. My sensors have detected..."
  - "I have finished surveying the Cafe. I think the place is dangerous. I am 60% confident about this assessment."

Pilot studies with Unity 3D Testbed

- Variance in reconnaissance strategies
- Variance in trust in the low ability (unreliable) robot
  - Trust
  - Distrust: use of camera and protective gear
- Combination of decision and confidence level: additional ability of the robot being self-aware?