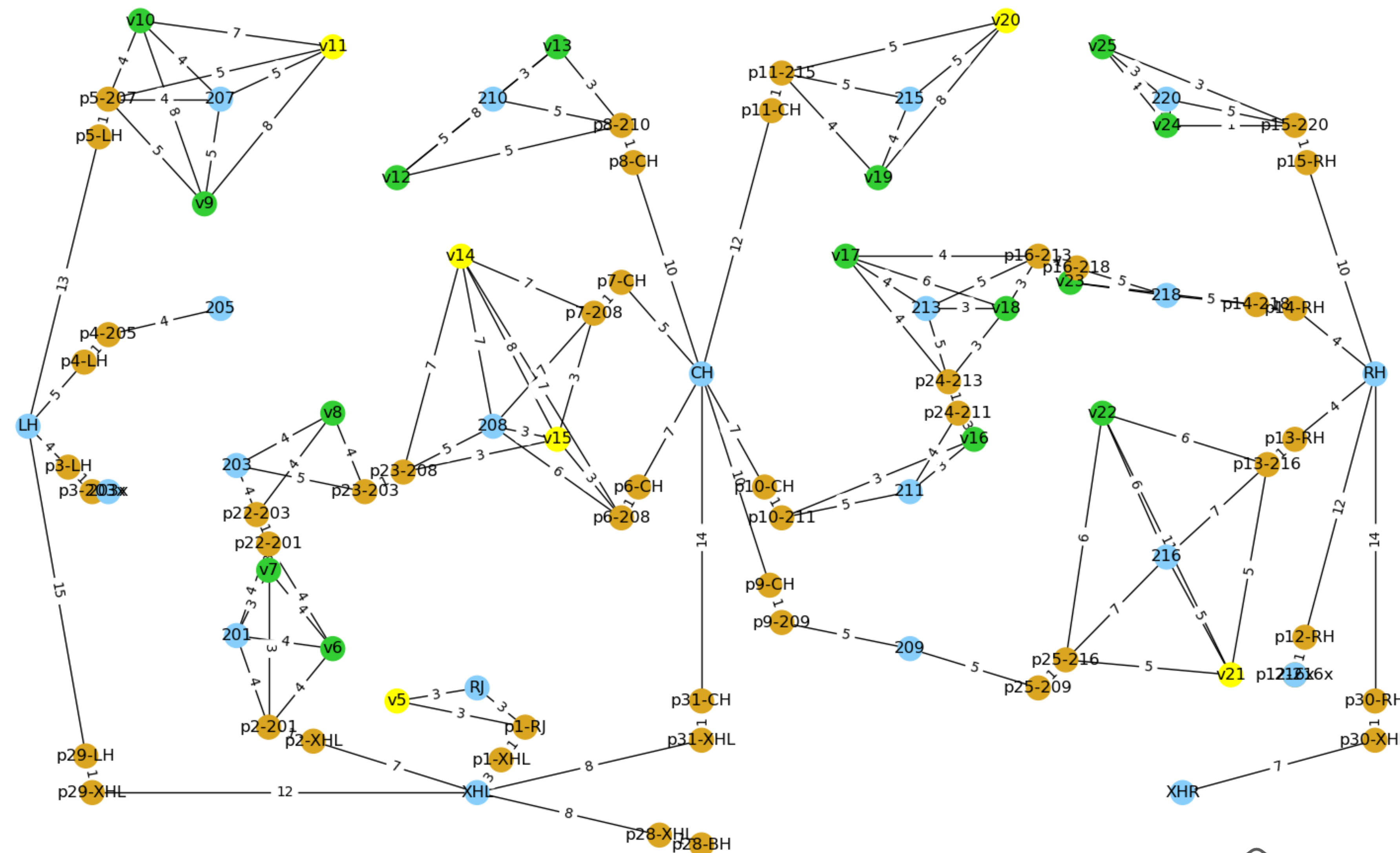# A Search & Rescue Graph for Building Artificial Social Intelligence

Arjun Voruganti, Yunzhe Wang, Jincheng Zhou, Volkan Ustun
Institute for Creative Technologies, University of Southern California

1. Search & Rescue (SAR)

The SAR scenario is a **team-based theory of mind** task. Humans observe and construct mental models of their environment, and then choose a best course of action to rescue victims in distress.

We propose a graph environment in order to **develop AI-based agents that can assist humans**:

- Removes complexities of the real world – easier to develop AI models and algorithms
- Still maintains the conceptual basis of SAR

2. Minecraft Environment

A **1st level abstraction** from real world SAR:

- 3D graphics and first-person perspective
- Blocks represent semantic objects: "door," "victim," "light switch," "furniture," etc.
- Objective: Rescue as many victims as possible within 10 minutes

Humans play a SAR scenario in Minecraft and data is collected on their mental models and decision-making patterns.
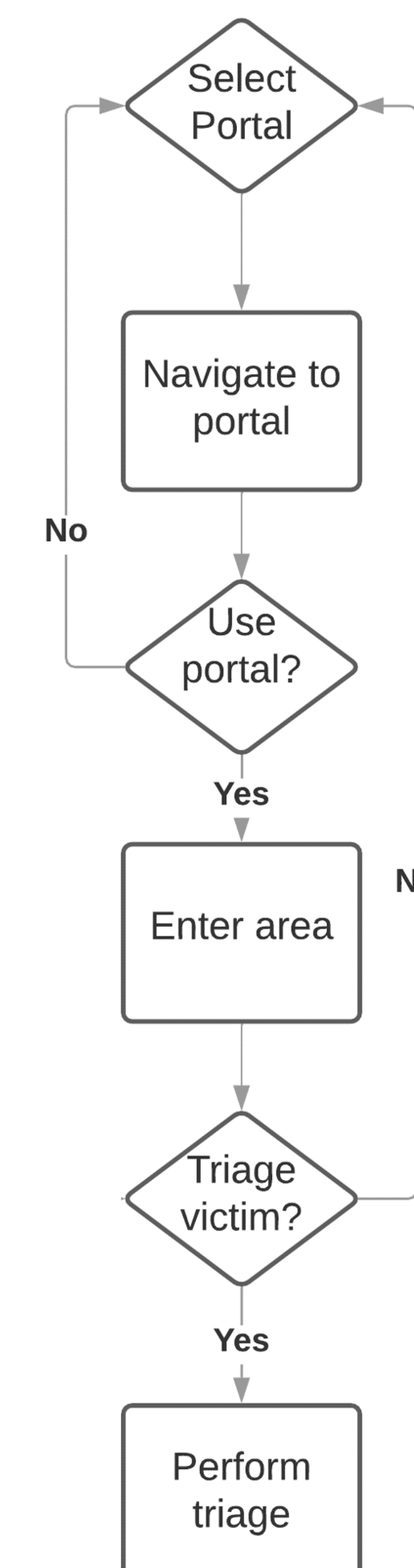
3. Graph Environment

A **2nd level abstraction** from real world SAR:

- Nodes represent semantic objects: "**room**," "**portal**" (door), and "**victim**"
- Edges represent distance between nodes
- Objective: Rescue as many victims as possible within 10 minutes

4. Reinforcement Learning (RL)

RL experiments are run in the graph environment. Goal: **Train AI-based models that match the mental models of humans** when they play Minecraft SAR.

Like humans, RL algorithms must make observations of their environment and choose appropriate sequences of actions.

Action sequence in the graph environment.

5. Project Malmo

The Malmo interface allows **for RL experimentation in Minecraft**.

RL models trained in the graph can be applied to an AI-controlled agent in Minecraft to assess their performance side-by-side with human players.

6. Artificial Social Intelligence

Long term goal: Build AI-based models that resemble theory of mind in humans. Implement them in real world AI systems for better **human-robot teamwork**.