Transfer Learning Experiments with Soar & the UCT

Nicholas Gorski

with John Laird and Taylor Lafrinere

Outline

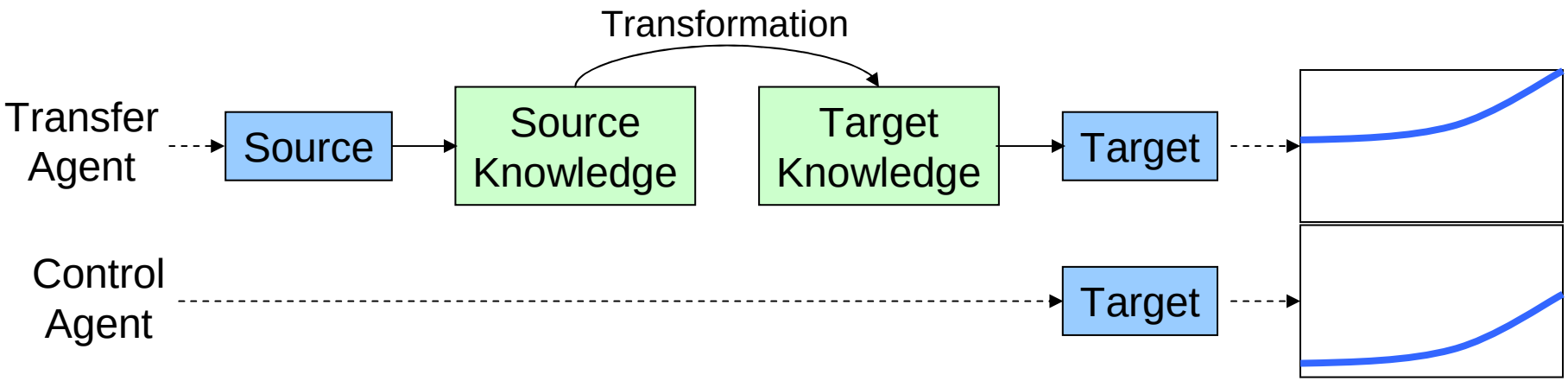
- Transfer Learning Background
- Urban Combat Testbed (UCT)
- Results
- Spatial Reasoning in the UCT
- Movies!

Project Background

- 3 year DARPA Transfer Learning (TL) initiative
- Soar grouped with ICARUS & Companion
  - Y1: Urban Combat Testbed (completed F06)
  - Y2: GGP (ongoing, evaluation in F07)
- Last workshop, reported some initial experiments

Transfer Learning (TL)

- Similarities to multi-task learning, inductive learning, and "learning to learn"
- Transfer Learning:
  - performs in source problem
  - applies learned knowledge to a target problem via transformation
  - performs in target problem, applying previoulsy learned knowledge



Urban Combat Testbed

- Software suite consisting of
  - First Person Shooter video game engine
  - Scenarios designed to test for specific types of transfer
- Complex domain
  - Large and continuous
  - Noisy actions
  - Many different objects & obstacles
    - Doors, windows, barriers, pits, water, electrical barriers, etc.

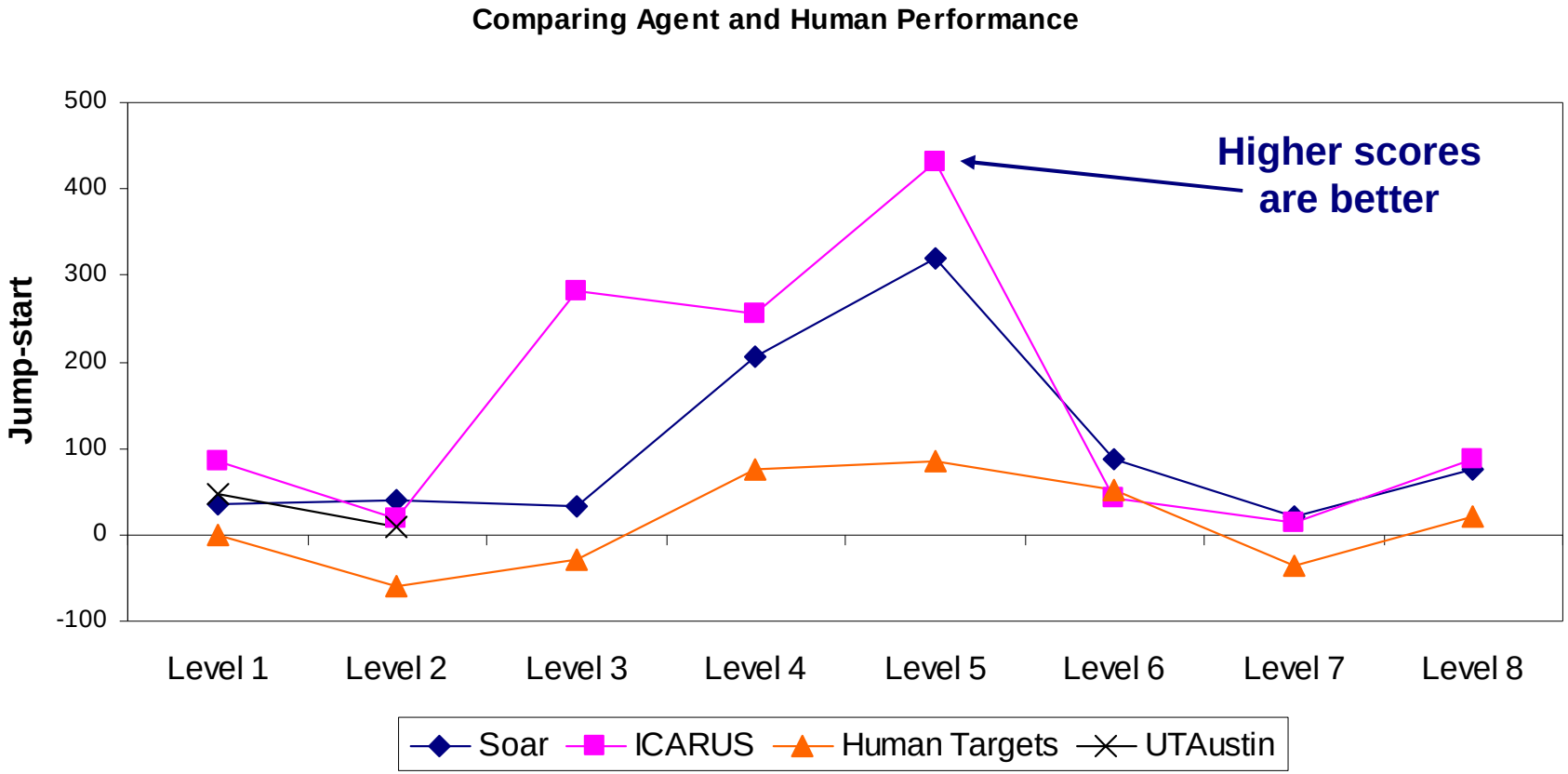
Urban Combat
Head
Torso
Arms
Legs
Fist
37 - 37

TL Scenarios in the UCT

Source Target

- Agent must navigate from start to goal
- To reach goal, it must climb ladder/drainpipe
- Generalization: drainpipe can be climbed because ladder was climbable

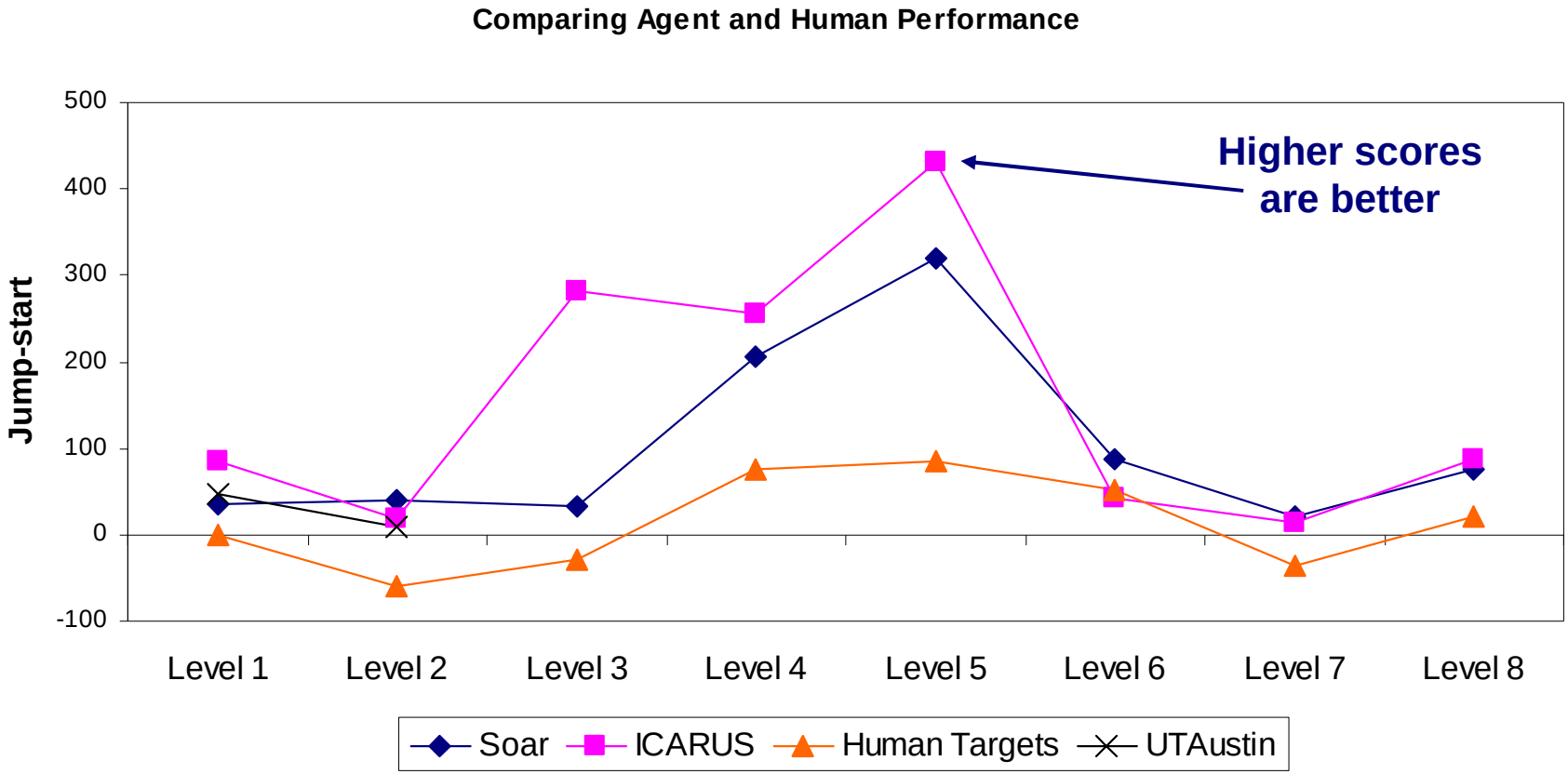
Y1 Results


Comparing Agent and Human Performance
Jump-start
500
400
300
200
100
0
-100
Higher scores are better
Level 1
Level 2
Level 3
Level 4
Level 5
Level 6
Level 7
Level 8
Soar
ICARUS
Human Targets
UTAustin

Jump-start Discussion

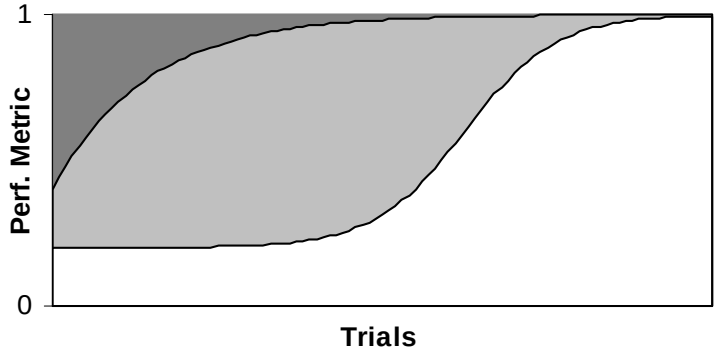
- Y1 results
  - Single methodology used by all teams
  - Used for go/no-go decision for Y2
- Used “Jump-start”
  - Magnitude of initial differences in performance (not normalized)
  - Rewards poorly performing agents

Y1 Results


Comparing Agent and Human Performance
Jump-start
500
400
300
200
100
0
-100
Higher scores are better
Level 1
Level 2
Level 3
Level 4
Level 5
Level 6
Level 7
Level 8
Soar
ICARUS
Human Targets
UTAustin

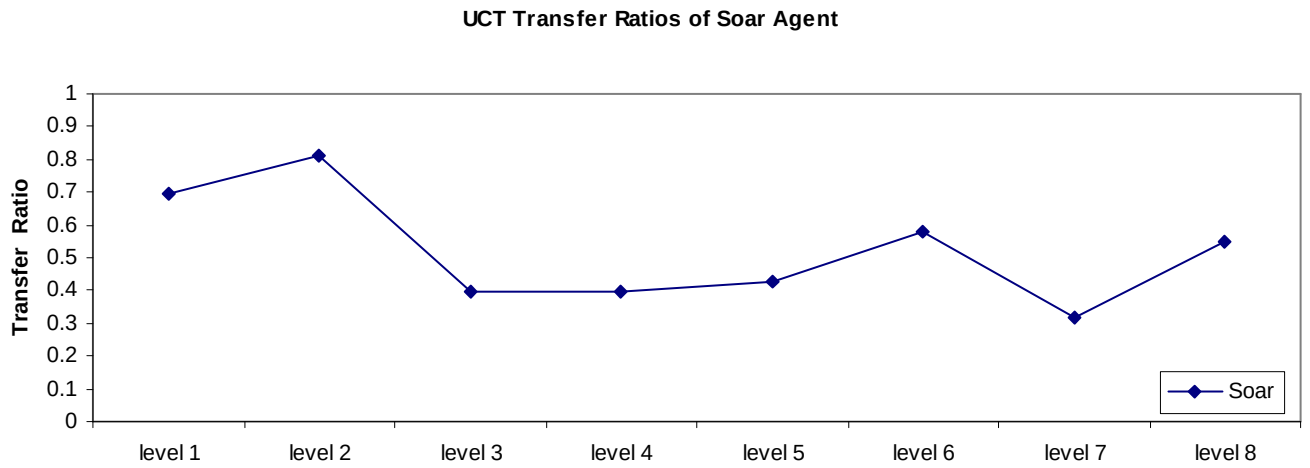
Calibrated Transfer Ratio (CTR)

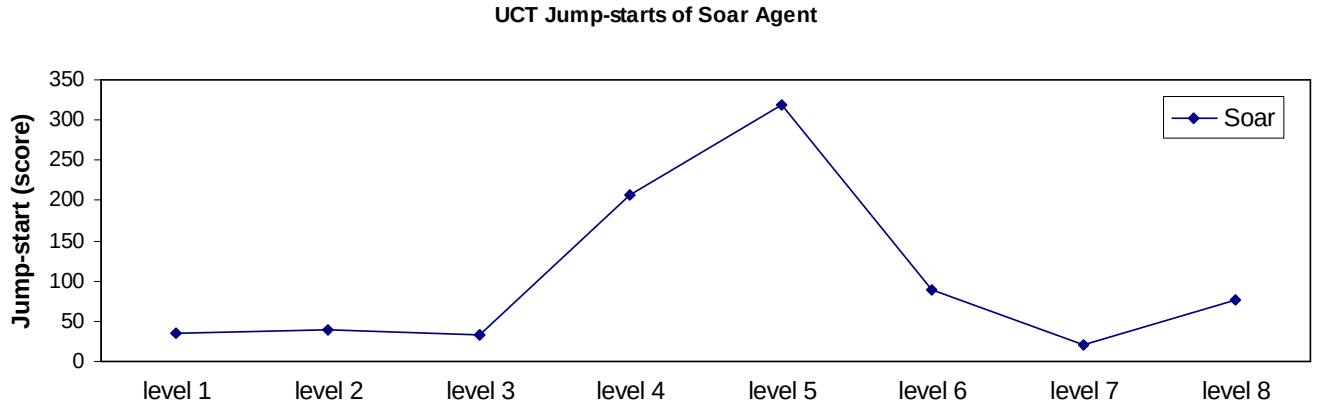
$$1 - \frac{\int_{t_0}^{t^*} OPT(t) - AB(t)\,dt}{\int_{t_0}^{t^*} OPT(t) - B(t)\,dt}$$



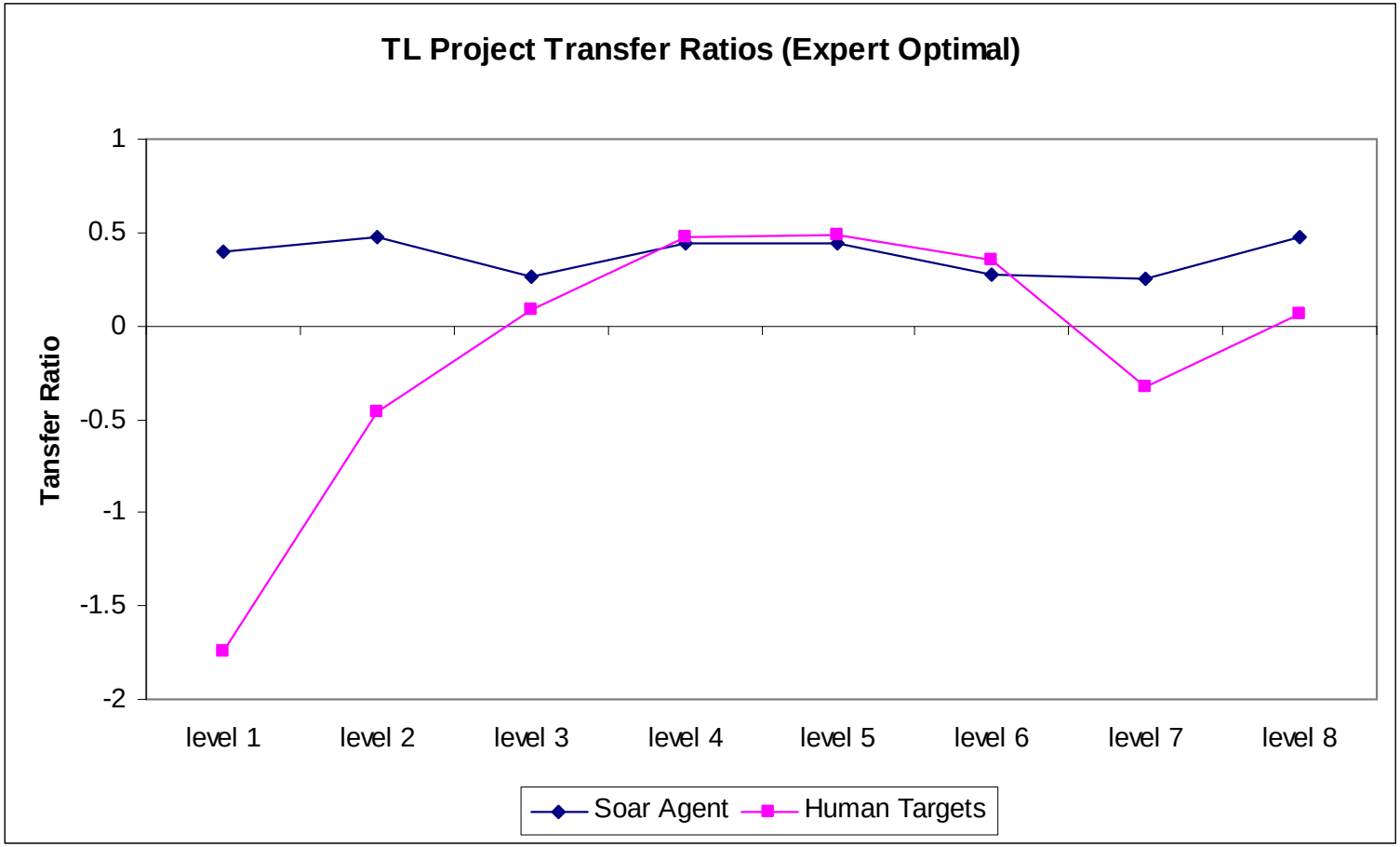
- Interpretation: “the amount of available improvement achieved”
- Disadvantage: requires knowledge of optimality
- Advantage: more meaningful

UM Evaluation: CTR vs Jump-start





UM Evaluation: Soar vs Human



Outline

- Transfer Learning Background
- Urban Combat Testbed (UCT)
- Results
- **Spatial Reasoning in the UCT**
- Movies!

Navigation in UCT

- Agent perceives 3D space as set of positive and negative convex polyhedrons
  - Mapped to 2D convex polygons by SML middleware
  - “Gateways” are intersections of free space regions
- UCTBot navigates from region to region by:
  - Moving to a gateway
  - Moving through a gateway
  - Suboptimal navigation
    - Doesn’t cut close to corners
    - Uses partitioning even when moving in wide open terrain

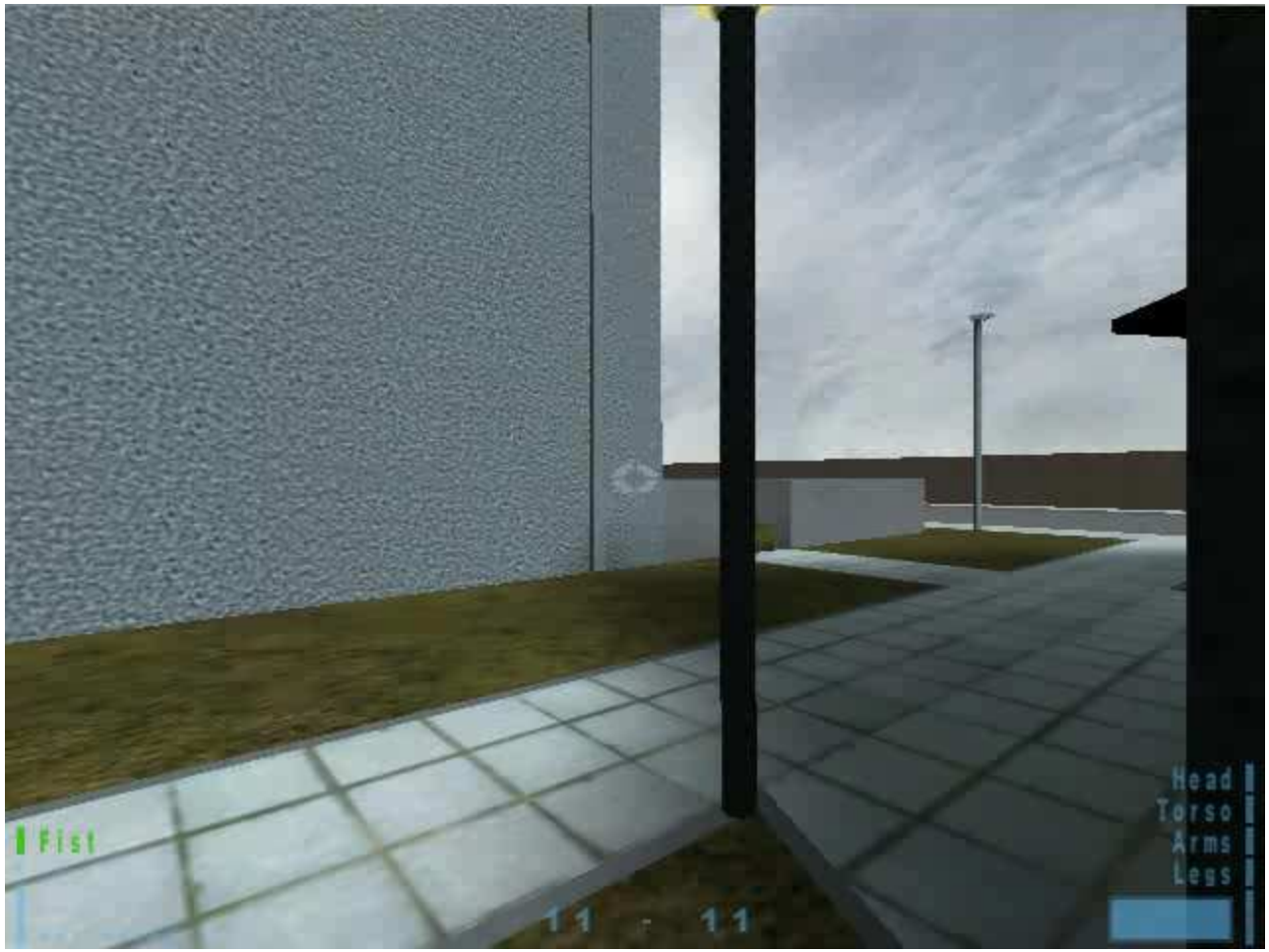
Obstacle Detection & Avoidance

- UCTBot is “blind” to obstacles and some gateways
- Detection: relies on velocity
- Avoidance:
  - Some obstacles can be surmounted
    - Test all available actions
    - Robust for most obstacle types
  - For blocking obstacles, find paths around them
    - (Mostly) robust
- Learning: which obstacle/gateway is blocking?

(Partially) Motivated SRS

- Better navigation improves performance
  - In untrained agents
  - In transfer agents
  - To honestly evaluate TL, must optimize both transfer & control cases
- Possible applications:
  - Route finding, obstacle avoidance
  - But would also allow for multi-agent tasks

Learning to climb



Searching indoors

Using weapons

Gold Nuggets & Lumps of Coal

- Met Y1 goals
- Laid groundwork for more motivating TL experiments in Y2
- Motivated SRS
- Developed CTR

- UCT cut from Y2
- Scenarios and domain lacked motivating transfer
- Didn’t get to use SRS
- CTR not adopted for internal evaluations