

Experiments with SARSA

Eric B Baum
Dennis Horte
Chick Markley
ebaum@fastmail.fm
Azure Sky Research Inc
1238 Campus Drive
Berkeley CA 94708

We have been SARSA and gradient-descent training neural nets to address some interesting problems. We built/trained a program that entered into the AI Challenge, sponsored by Google, which was a contest to produce programs to control hundreds of ants in a complex video game. The game requires response in half a second, so moving the ants intelligently is a problem well beyond human capabilities. Although we had major bugs at the time of the tournament submission that greatly hampered its learning, our trained program still beat almost 90% of human coded entrants in the programming competition. You can see our ants play at <http://aichallenge.org/profile.php?user=11208>. We believe we could produce a much stronger program, and may come back to this before the Snowbird conference but have moved on to No Limit Texas Holdem after the Ants competition deadline. After a few weeks of development and learning, our nets are currently playing Hold'Em at a level that beats two hand coded expert systems of our design (that we haven't invested great effort in but are not terrible), and seem to Eric's strong amateur judgement to be better than some humans he's played.

We have achieved these results by a judicious combination of self-play, play against other programs, learning from expert traces, reward engineering, hand coding of features, and training sub-nets and functions on subproblems. Our goal is a methodology and tool kit that allows rapid production of strong programs to control complex systems or to make useful predictions.