Markov Games
vs.
Game Tree Search

Ronald Parr
Duke University

(Contains joint work with Michail Lagoudakis)

Type Mismatch?

- Markov Games: A problem class

- Game Tree Search: An algorithm

Dynamic Programming
vs.
Game Tree Search

- Different algorithms for the same problem
- Both are impractical in most cases
- Limitations *may* be complementary

Game Tree Search

- Exponential growth with depth
- Discrete action space
- Tractable only for shallow trees

- In practice:
  - Depth/resource limited search
  - Heuristic evaluation used for “terminal” values

Dynamic Programming

- Polynomial in state space (good and bad)
- Discrete actions or mixed strategies
- Tractable only for small state spaces

- In practice:
  - State aggregation
  - Value approximation
    - Alternating games: Samuels, Tesauro, etc.
    - Simultaneous move 0-sum games: Lagoudakis & Parr, Savagoankar, Chong & Givan

Observations

- Search is necessarily myopic
  - Depth limited
  - Needs good terminal values
- DP is necessarily approximate
  - Can propagate values across long distances
  - Can only do so coarsely due to VFA

Combining DP and Search

- Use DP to devise a good evaluation function
- Natural combination:
  - DP = slow, global, off-line (but evaluates quickly)
  - Search = relatively fast, local, on-line
- Done by TD-Gammon
  - ~1 million off line training games played
  - Shallow search in play
- Rarely done though obviously *the right thing*

Why not more DP Eval. Fns.?

- People are good at hand-tweaking evaluation functions (for popular games)

- Approximate DP:
  - Hard to do well (requires good features)
  - Requires LOTS of data
  - Violation of Markov assumption (from value function approximation) introduces bias into approximation

Prospects #1

- Improved approximation methods may help
  (Plug #1: Try LSPI)

- Approximation more important for simultaneous move games
  - Hard to prune
  - On-line solutions are expensive (large LP)

  (Plug #2: Try LSPI for this too!)

About our work…

- Show that value function approximation works the same way for zero-sum stochastic games as MDPs

- Combined with LSPI
  - Low data requirements
  - Used for generalization of Littman's soccer
  - Used for router optimization

- Applied approximation to team vs. team games (Address large action spaces within function approximation framework.)

Prospects #2

- DP for partial information games? (Hansen, Bernstein & Zilberstein)
- Can do DP on your information state
- Use approximate POMDP techniques?