General Board Game Playing for Education and Research in General AI Game Learning

Wolfgang Konen

TH Köln – University of Applied Sciences

Technology
Arts Sciences
TH Köln

What is GBG?

- Let various agents play on ‚all' board games
- Standardized interfaces: game states, agents, …

Games	N	det
TicTacToe	2	D
Hex (scalable)	2	D
2048	1	ND
Connect-4	2	D
Nim (scalable)	2	D
Othello	2	D
Rubik's Cube (beta)	1	D
Sim (beta)	2,3	D

Table 1: Games in GBG (N: # player, D: deterministic, ND: non-deterministic)

Agents	Description
Max-N	‚Minimax' for N player
Expectimax-N	… for ND games
MC	Monte Carlo
MCTS	Monte Carlo Tree Search
MCTSE	… for ND games
TD	Temporal Difference
TD n-tuple	… with n-tuple features
Sarsa	Sarsa with n-tuple

Table 2: Agents in GBG

Why? We have GGP!

GGP with GDL: learns an unknown game at *run time*. This is a tough task and fantastic endeavor in logic reasoning!

But as a consequence there are also some limitations [Swiechowski2015] of GGP:

- Simulations in GDL are slow
- Cannot compare & compete with the best game-specific agents (e.g. Othello → Edax, Hex → Hexy)
- CI agents (TD and deep learning) are difficult to integrate (slow learning and combinatorial explosion)

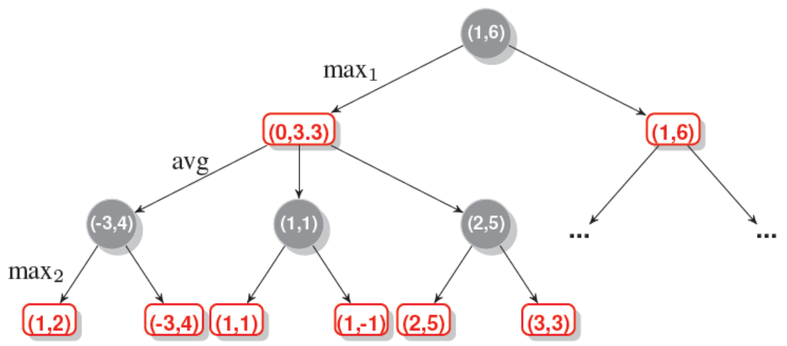
GBG Advantages

- Easier for education: The complicated agents can be re-used for new games
- First ***generic*** implementation of **TD-n-tuple agents**: Arbitrary games, arbitrary number of players
- GBG allows fast game simulation (10.000 – 90.000 moves per second) for CI agents
- Comparison with game-specific agents (strong or perfect players, e.g. Othello → Edax) is possible
- Generic inclusion of game **symmetries**
- Game-specific visualization and inspection → get deeper insights
- Human – agent play

Disadvantage of GBG:

- Does not allow new games at *run time*. But new games can be added at *compile time*.

Expectimax-N



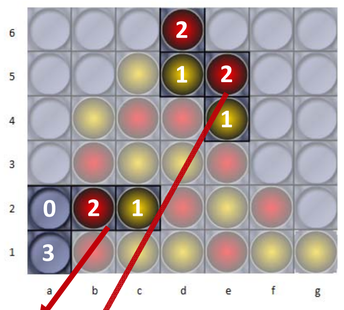
Expectimax-N tree for N-player games: extension of Max-N for nondeterministic games. Shown is an example for N = 2 and depth d = 3.

- 1st level: maximizes the tuple entry of the 1st player,
- 2nd level: expectation value of all child nodes (grey circles), each having a certain probability of occurrence,
- 3rd level: maximizes the tuple entry of the 2nd player.

Similar extension: MCTS → MCTS-Expectimax

N-Tuple Systems

An n-tuple is a sequence of board cells. Example Connect-4:
Each cell can have one out of P=4 states:
0=empty and not reachable, 1=Yellow, 2=Red,
3=empty and reachable



$T_1 = (a1, a2, b2, c2)$
$\xi_1 = 3 \cdot 4^0 + 0 \cdot 4^1 + 2 \cdot 4^2 + 1 \cdot 4^3 = 99$

$T_2 = (d6, d5, e5, e4)$
$\xi_2 = 2 \cdot 4^0 + 1 \cdot 4^1 + 2 \cdot 4^2 + 1 \cdot 4^3 = 102$

Educational Benefits

Benefits for students:

- Agents readily available
- First results within days or weeks
- Code better re-usable

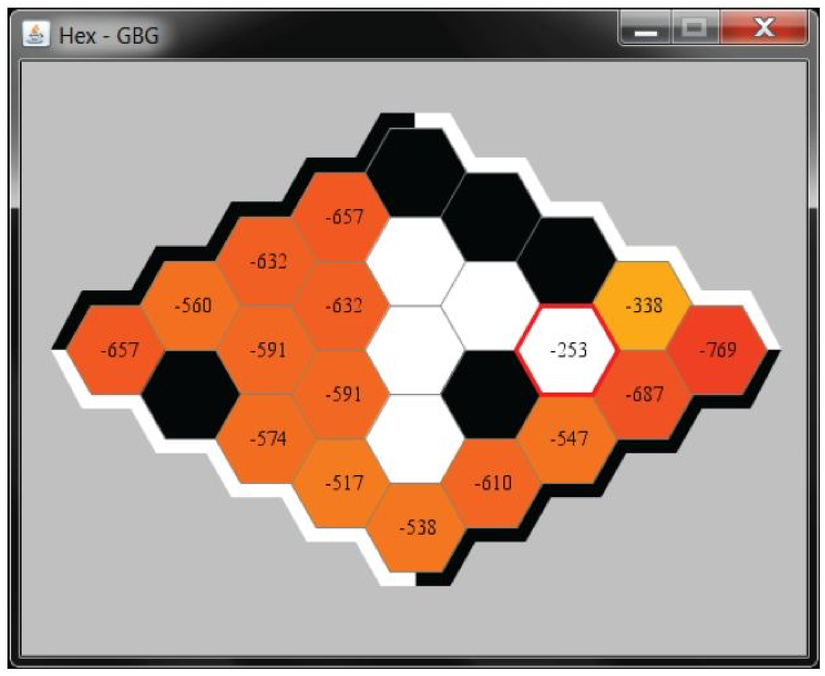
Evaluation (questionnaire, 3 students):

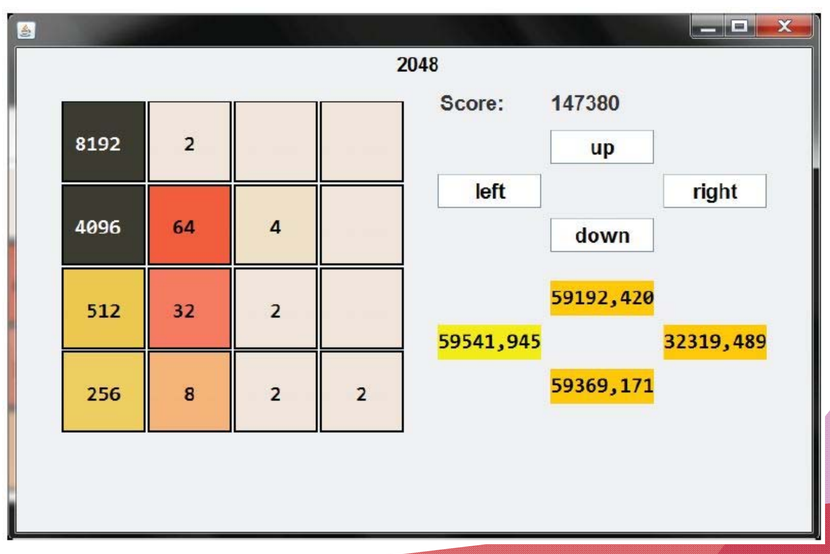
- "Better than starting from scratch?" – Strong agree
- "Time to get familiar with GBG?" – 2 days (median)
- "Enough documentation?" – Strong agree

Their wishlist:

- More GUI-elements for game-related settings
- More "How-To"-cases in documentation

Example Games



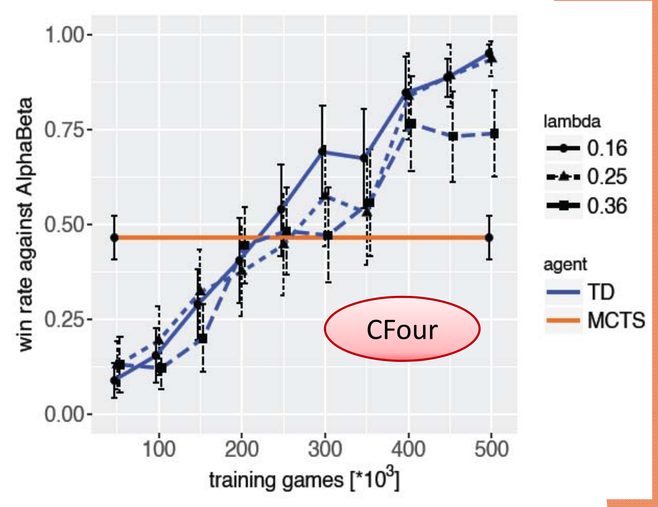


Speed

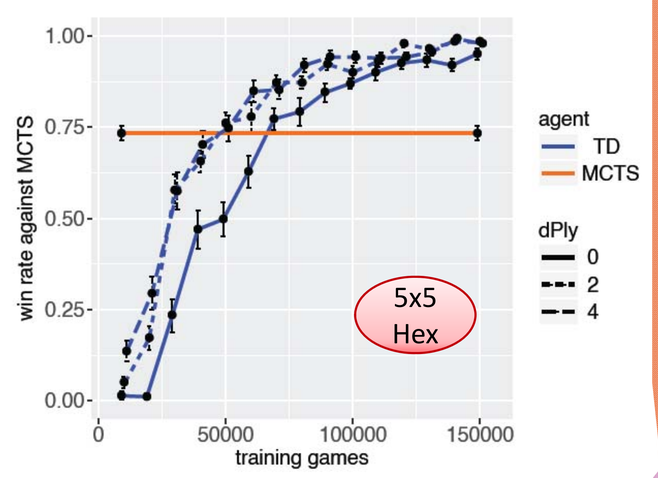
- GBG allows fast game simulation / game play
- The table shows moves/second for various agents and games on a single core.

Agent	Game	Moves/second during … … game learning	… game play
TD-n-tuple [0-ply]	2048	66.000	94.000
TD-n-tuple [2-ply]	2048	66.000	5.000
MCTSE (1000 iter)	2048	–	120
TD-n-tuple [0-ply]	Connect-4	7.900	40.400
TD-n-tuple [2-ply]	Connect-4	7.900	5.100
MCTS (1000 iter)	Connect-4	–	54
TD-n-tuple [0-ply]	5x5 Hex	17.600	20.500
TD-n-tuple [2-ply]	5x5 Hex	17.600	700
MCTSE (1000 iter)	5x5 Hex	–	31

Some Results



Win rates on winnable episodes



Conclusion

GBG benefits

Educational perspective:

- Much easier for students to use complex CI agents
- Standardization of code development
- Attracts students to fascinating area of game learning

Research perspective:

- **Fast** simulation of CI agents
- TD-n-tuple successful on diverse games: 2048, Connect-4, Hex, Othello, Nim
- Success against **strong** or **perfect-playing** agents (Connect-4, Hex7x7, Nim) → advantage over GGP
  - Not yet: Othello, Hex above 7x7

CIOP

Technology
Arts Sciences
TH Köln