Monte Carlo Approaches to Parameterized Poker Squares

Todd W. Neller, Zuozhi Yang, Colin M. Messinger,
Gettysburg College (tneller@gettysburg.edu)

Calin Anton, Karo Castro-Wunsch, William Maga
MacEwan University (antonc@macewan.ca)

Steven Bogaerts, Robert Arrington, and Clay Langley
DePauw University (stevenbogaerts@depauw.edu)

Outline

- Motivation: to promote undergraduate AI research

- Challenge: Parameterized Poker Squares

- Contest Structure and Results

- Overview of Top 3 Players

- Conclusions

What is the EAAI NSG Challenge?

- EAAI: Educational Advances in Artificial Intelligence (@AAAI)
- NSG: NSF has its Grand Challenges. Ours are **N**ot **S**o **G**rand. ☺
- Goal: to teach the craft of research to undergraduates
  - "One must learn by doing the thing; for though you think you know it, you have no certainty, until you try." - Sophocles
  - Enable high-quality, faculty-mentored undergraduate research experiences that include an option for peer-reviewed paper publication and presentation at a major CS conference.
- EAAI-16's NSG Challenge: **Parameterized Poker Squares**

Poker Squares (Original)

- Materials: shuffled standard (French) 52-card card deck

- Each turn, a player draws a card and places it into an empty 5-by-5 grid position. (No card may be moved after placement.)

- After 25 turns, the grid is full.

- Each 5-card row and column is classified as a Poker hand and is scored according to a point system.

- The final score is the sum of the 10 hand scores.

American Point System

Poker Hand	Points	Description	Example
Royal Flush	100	A 10-J-Q-K-A sequence all of the same suit	10♣, J♣, Q♣, K♣, A♣
Straight Flush	75	Five cards in sequence all of the same suit	A♦, 2♦, 3♦, 4♦, 5♦
Four of a Kind	50	Four cards of the same rank	9♣, 9♦, 9♥, 9♠, 6♥
Full House	25	Three cards of one rank with two cards of another rank	7♠, 7♣, 7♦, 8♥, 8♠
Flush	20	Five cards all of the same suit	A♥, 2♥, 3♥, 5♥, 8♥
Straight	15	Five cards in sequence; Aces may be high or low but not both	8♣, 9♠, 10♥, J♦, Q♣
Three of a Kind	10	Three cards of the same rank	2♠, 2♥, 2♦, 5♣, 7♠
Two Pair	5	Two cards of one rank with two cards of another rank	3♥, 3♦, 4♣, 4♠, A♣
One Pair	2	Two cards of one rank	5♦, 5♥, 9♣, Q♠, A♥
High Card	0	None of the above	2♦, 3♣, 5♠, 8♥, Q♦

Scoring Example

Parameterization of Poker Squares

- The American Point System (0, 2, 5, 10, 15, 20, 25, 50, 75, 100) is based on hand rank in Poker.
- The British Point System (1, 3, 6, 12, 5, 10, 16, 30, 30) is based on the difficulty of forming the hands in Poker Squares. (a.k.a. English Point System)
- For our challenge, AI players are *given the scoring system at play time* with points in the range [-128, 127]. Examples:
  - Ameritish point systems: random variations on American and British systems
  - Single Hand: 1 point for one hand type, 0 points otherwise
  - Hypercorners: all 1 or -1 score values
  - Random: random score system in range [-128, 127]

Structure of the Game

- The game is structured as an alternating sequence of *chance nodes* and player *choice nodes*.

  - Each card draw is a probabilistic event where any remaining card is drawn with equal probability.

  - Each player *action* is a commitment to a card placement.
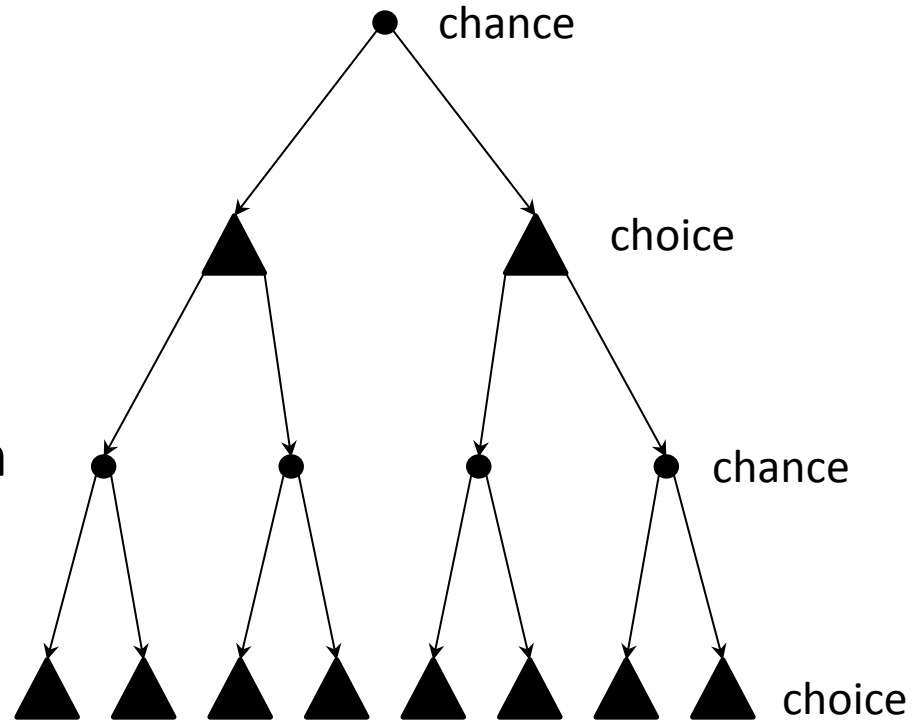
Game Tree Size

- How big is the Poker Squares game tree?
  - Root chance node: 52 possible cards
  - 52 depth-1 choice nodes: 25 possible placements
  - 52x25 depth-2 chance nodes: 51 possible cards
  - 52x25x51 depth-3 choice nodes: 24 possible placements
  - …
  - $52!/27! \times 25! = 52!/(27 \times 26) \cong 1.15 \times 10^{65}$ nodes
  - Although:
    - Different draw/play sequences can lead to the same state.
    - Rows/columns may be reordered without affecting score.
  - Still, we cannot evaluate entire expectimax trees except for much smaller end-game situations.

Evaluation

- Players were evaluated using 12 point systems. For each point system:
  - Players had 5 minutes to process the point system, form strategy, etc.
  - Players then played 100 games with 30 seconds of decision time per game.
  - Total scores were linearly scaled between 0 (min. total score) and 1 (max. total score).
- The player with the maximum sum of scaled total scores is the winner.

Results

1. Score: **11.821**; Player: **BMO_V2**; Students: Karo Castro-Wunsch, William Maga; Faculty mentor: Calin Anton; School: MacEwan University

2. Score: **11.763**; Player: **GettysburgPlayer**; Students: Colin Messinger, Zuozhi Yang; Faculty mentor: Todd Neller; School: Gettysburg College

3. Score: **11.334**; Player: **Tiger**; Students: Robert Arrington, Clay Langley; Faculty mentor: Steven Bogaerts; School: DePauw University

4. Score: 11.170; Player: JoTriz; Student: Kevin Trizna; Faculty mentor: David Mutchler; School: Rose-Hulman Institute of Technology

5. Score: 7.149; Player: SRulerPlayer; Student: Zachary McNulty ; Faculty mentor: Timothy Highley; School: La Salle University

6. Score: 0.192; Player: MonteCarloTreePlayer; Student: Isaac Sanders; Faculty mentor: Michael Wollowski; School: Rose-Hulman Institute of Technology

7. Score: 0.190; Player: DevneilPlayer; Student: Adam Devigili; Faculty mentor: Brian O'Neill; School: Western New England University

Players **Mean Scores by Point System**

Players	American	Ameritish	British	Hypercorner	Random	High Card	One Pair	Two Pair	3 of a Kind	Straight	Flush	Full House
BMO_V2	125.27	105.54	54.50	1.10	437.77	9.37	9.12	4.46	3.20	2.97	3.43	1.82
DevneilPlayer	14.36	15.27	7.51	-9.52	-86.92	5.22	4.10	0.45	0.21	0.04	0.05	0.03
Gettysburg	123.94	110.28	53.38	1.24	429.89	9.37	9.17	4.47	3.02	2.71	3.46	1.93
SRulerPlayer	51.83	55.39	30.29	-5.10	242.85	9.34	8.84	4.04	2.10	1.58	1.98	0.61
JoTriz	116.75	109.03	53.59	-0.78	351.07	9.31	9.15	4.59	3.03	2.59	3.36	1.67
Tiger	116.12	111.26	53.92	-2.20	411.78	9.35	9.16	4.52	2.89	2.94	3.41	1.82
MonteCarloTreePlayer	15.47	15.31	7.61	-9.30	-86.83	4.80	4.53	0.45	0.20	0.05	0.02	0.00
RandomPlayer	14.25	15.67	7.71	-9.66	-106.80	5.20	4.31	0.42	0.23	0.01	0.01	0.01
Max	125.27	111.26	54.50	1.24	437.77	9.37	9.17	4.59	3.20	2.97	3.46	1.93
Min	14.25	15.27	7.51	-9.66	-106.80	4.80	4.10	0.42	0.20	0.01	0.01	0.00

Normalized Scores

Players	American	Ameritish	British	Hypercorner	Random	High Card	One Pair	Two Pair	3 of a Kind	Straight	Flush	Full House	Total
BMO_V2	1.00	0.94	1.00	0.99	1.00	1.00	0.99	0.97	1.00	1.00	0.99	0.94	**11.821**
DevneilPlayer	0.00	0.00	0.00	0.01	0.04	0.09	0.00	0.01	0.00	0.01	0.01	0.02	**0.190**
Gettysburg	0.99	0.99	0.98	1.00	0.99	1.00	1.00	0.97	0.94	0.91	1.00	1.00	**11.763**
SRulerPlayer	0.34	0.42	0.48	0.42	0.64	0.99	0.93	0.87	0.63	0.53	0.57	0.32	**7.149**
JoTriz	0.92	0.98	0.98	0.81	0.84	0.99	1.00	1.00	0.94	0.87	0.97	0.87	**11.170**
Tiger	0.92	1.00	0.99	0.68	0.95	1.00	1.00	0.98	0.90	0.99	0.99	0.94	**11.334**
MonteCarloTreePlayer	0.01	0.00	0.00	0.03	0.04	0.00	0.08	0.01	0.00	0.01	0.00	0.00	**0.192**
RandomPlayer	0.00	0.00	0.00	0.00	0.00	0.09	0.04	0.00	0.01	0.00	0.00	0.01	**0.153**

BeeMo Overview

- 1st Place: 11.821/12 (98.5% of max)
- Parallel Flat Monte Carlo (FMC) Search
- Heuristic Evaluation:
  - Assume row/column independence of expected partial hand scores.
  - Partial hands are represented as bitstrings of salient features.
  - Heuristic evaluation is the sum of expected bitstring scores.
  - Bitstring score expectations are initially learned with MC reinforcement learning (RL).
- Parallelized to maximize use of available cores.

BeeMo 16-Bit Hand Encoding

- 1 bit – if the hand contains a flush
- 1 bit – if the hand contains a straight
- 3 bits – number of cards in the hand without a pair
- 2 bits – number of pairs in the hand
- 1 bit – if the hand contains three of a kind
- 1 bit – if the hand contains four of a kind
- 1 bit – if the hand is a row hand
- 2 bits – number of undealt cards of primary rank
- 2 bits – number of undealt cards of 2nd-ary rank
- 2 bits – not/exactly/more than enough cards for flush

GettysburgPlayer Overview

- 2nd Place: 11.763/12 (98.0% of max)
- Expectimax search to depth 2 (only 9% of allotted time)
- Heuristic Evaluation:
  - Assume independence of expected partial hand scores.
  - Partial hands represented as a string of salient features.
  - Heuristic evaluation is the sum of expected partial hand scores.
  - Partial hand score expectations learned with MC RL with epsilon-greedy simulation and exponentially decaying epsilon across the 5 minutes.

Naïve Partial Hand Abstraction

- Why naïve? – Assume that each row/column is independent
- Key Features:
  - number of cards played so far
  - row ("-") or column ("|")?
  - sorted rank counts with how many cards undealt in the rank
  - flush ("f") is achievable? If so, how many cards remaining in flush suit?
  - straight ("s") is achievable?
  - royal flush ("r") is achievable?
- For example, "14|1(3)1(2)1(2)f3(8)*s*" represents a column hand abstraction after the 14th move.

Tiger Overview

- 3rd Place: 11.334/12 (94.5% of max)
- Monte Carlo Tree Search (UCT) from Kyle Hughart (code.google.com/p/uct-for-games) with
  - High UCB1 parameter C_p = 200 / $\sqrt{2}$
  - Pruning of moves with lowest heuristic score
  - Tuned heuristic based on weighted partial hand patterns

Tiger Heuristic

- Each partial hand maps to a possibility vector consisting of 0, 1, and 2, encoding impossibility, possibility, and achievement of each hand classification.

- For each 4 card hand, a [0, 1] weight approximates the probability of hand achievement on next draw.

- For fewer cards, relative likelihoods of hand achievements are hand-estimated from English scoring system and then weighted according to tuned parameters for 1-, 2-, and 3- cards.

Conclusion

- Parameterized Poker Squares is a fun, simple, accessible, and interesting game of chance suitable for game AI research with undergraduates.
- While this first challenge attracted only 7 submissions, it is a modest step forward in growing mentored undergraduate game AI research.
- All top performers, while abstracting the state space differently for heuristic evaluation
  - made a row/column independence assumption about expected partial hand scores, and
  - refined decisions to account for dependencies through Monte Carlo techniques.