The Gin Rummy AI Challenge

Todd W. Neller

Gin Rummy Rules

- 2 players, 52 cards, Aces low
- **Object**: to be the first player to score 100 points or more.
- **Meld**: a set of 3 or more cards of the (1) same rank (e.g. 7♣,7♦,7♥) or (2) same suit in sequence (e.g. A♠, 2♠, 3♠, 4♠)
- **Deadwood**: cards not in melds
- **Card points**: 10 for face cards, A=1, number value for number ranks
- **Deadwood points**: sum of card points for all cards not in disjoint melds. (Melds must not share cards.)

Gin Rummy Play

- The dealer alternates. The dealer deals 10 cards to each player and turns the top card of the remaining draw pile face up to form a discard pile.
- Each turn, player may draw the top face-up card from the discard pile or the top face-down card of the draw pile.
  - First turn exception: If the dealer **declines** the top face-up card, the opponent may begin the deal play by drawing that card, or may also decline. If the opponent declines, the dealer begins by drawing from the draw pile.
- After drawing, the player must discard.
- A player who would have less than or equal to 10 deadwood points after discard may end the deal’s play by **knocking**, sometimes signalled by discarding face-down.

Gin Rummy Scoring

- After a player knocks,
  - The knocking player lays down melds face-up and reveals deadwood cards.
  - The opponent lays down any melds.
  - If the knocking player has any deadwood, the opponent may then “lay off” opponent deadwood cards to knocking player melds. Any remaining opponent deadwood is revealed.
- If the knocking player has no deadwood, they are said to have “gin”. That player scores 25 points + opponent deadwood points.
- If the knocking player has deadwood that is…
  - … less than opponent deadwood, the knocking player scores the deadwood point difference.
  - … greater than or equal to the opponent deadwood, the opponent “undercuts” and scores 25 points + the deadwood point difference.

Let’s Play

- Decks for local play
- Android: Gin Rummy Free by AI Factory Limited https://play.google.com/store/apps/details?id=uk.co.aifactory.ginrummyfree&hl=en_US
- Web: https://www.gin-rummy-online.com/
- iOS: Gin Rummy Plus by Zynga https://apps.apple.com/us/app/gin-rummy-plus-card-game/id1068095192

Research Topics Overview

- Opponent Hand Estimation
  - Given:
    - Your knowledge of which cards are discarded and unavailable, in your hand, and in the opponents hand
    - Your observations of which face-up cards were refused by the opponent, or discarded by the opponent
  - Estimate: the probability / relative likelihood of the opponent holding a particular card
- Optimal Play
  - Given: above game state knowledge and opponent hand estimation
  - Choose: a draw/discard action that maximizes the probability of winning

Opponent Hand Estimation

- Some basic opponent hand estimation is described by Jeff Rollason (AI Factory) at https://www.aifactory.co.uk/newsletter/2007_02_imperfect_info.htm
  - An opponent *drawing* a face-up card or *discarding* a card respectively *increases* or *decreases* the probability of having same rank or adjacent suit cards.

Example Estimation

- What can we learn if we observe a player:
  - Draw a face-down card with 4H face-up?
  - Discard 6D?
  - Draw a face-down card with 6H face-up?
  - Discard 7C?
  - Draw a face-up QD?

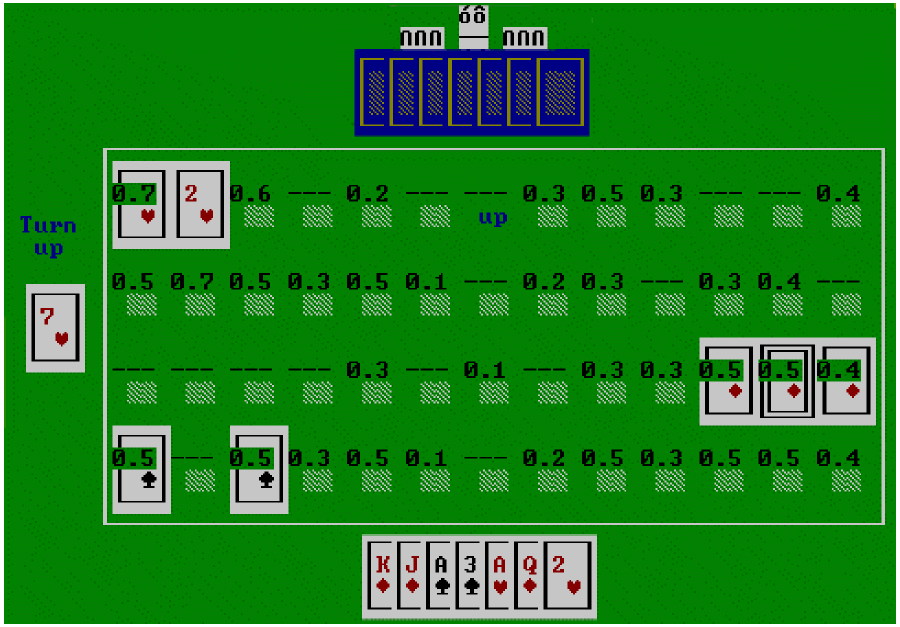


Image source: https://www.aifactory.co.uk/newsletter/2007_02_imperfect_info.htm

Bayes' Theorem

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

where A and B are events and $P(B) \neq 0$.

- $P(A \mid B)$ is a conditional probability: the likelihood of event A occurring given that B is true.
- $P(B \mid A)$ is also a conditional probability: the likelihood of event B occurring given that A is true.
- $P(A)$ and $P(B)$ are the probabilities of observing A and B independently of each other; this is known as the marginal probability.

Source: https://en.wikipedia.org/wiki/Bayes%27_theorem

Bayes' Theorem

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Let: A = holding card C in hand
B = observing draw/discard of card D

where A and B are events and $P(B) \neq 0$.

- $P(A \mid B)$ is a conditional probability: the likelihood of event A occurring given that B is true.
- $P(B \mid A)$ is also a conditional probability: the likelihood of event B occurring given that A is true.
- $P(A)$ and $P(B)$ are the probabilities of observing A and B independently of each other; this is known as the marginal probability.

Source: https://en.wikipedia.org/wiki/Bayes%27_theorem

Bayes' Theorem

$$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$$

Let: A = holding card C in hand
B = observing draw/discard of card D

where A and B are events and $P(B) \neq 0$.

- $P(A \mid B)$ is a conditional probability: the likelihood of event A occurring given that B is true.
- $P(B \mid A)$ is also a conditional probability: the likelihood of event B occurring given that A is true.
- $P(A)$ and $P(B)$ are the probabilities of observing A and B independently of each other; this is known as the marginal probability.

Source: https://en.wikipedia.org/wiki/Bayes%27_theorem

Observe:

- P(B) doesn't change for different A, so we can focus on the fact that P(A|B) $\propto$ P(B|A)P(A)
- For a given player strategy, we can simulate a lot of play, collect a lot of frequency data for P(B|A) and P(A), and estimate the relative P(A|B) for different possibly held cards when we make a draw/discard observation

Experimental Simple Gin Rummy Player

- Poor, simple strategy:
  - Ignore opponent actions and cards no longer in play.
  - Draw face up card only if it becomes part of a meld. Draw face down card otherwise.
  - Discard a highest ranking unmelded card from among the deadwood of melds that minimize deadwood points (without regard to breaking up pairs, etc.)
  - Knock as early as possible.

Simple Hand Estimation Experiment

- Simulate 10,000 games.
- Collect frequency data on draw/discard events for each card with respect to :
  - Whether or not a card was drawn face-up
  - Rank of the face-up card
  - Rank of the card discarded
  - Whether the card was suited with the face-up card and/or discarded card
- Given such frequency data (with non-zero event probabilities), apply Bayes' Theorem to adjust estimations of likelihoods of cards held.
- Demonstration: SimpleGinRummyPlayer2HETest

Example Output

Rank	A	2	3	4	5	6	7	8	9	T	J	Q	K
C	0.4894	0.5650	0.6587	FALSE	FALSE	FALSE	FALSE	FALSE	0.3366	0.4660	FALSE	0.1442	FALSE
H	0.4867	FALSE	FALSE	FALSE	0.2741	0.1352	FALSE	FALSE	0.2413	0.3678	FALSE	0.1727	FALSE
S	0.5438	0.5429	0.6530	0.3665	0.4975	0.2655	0.3018	0.0944	0.3211	0.4566	FALSE	FALSE	0.1257
D	0.4720	FALSE	FALSE	0.3248	FALSE	FALSE	0.1043	FALSE	0.1903	0.2938	0.1083	FALSE	FALSE

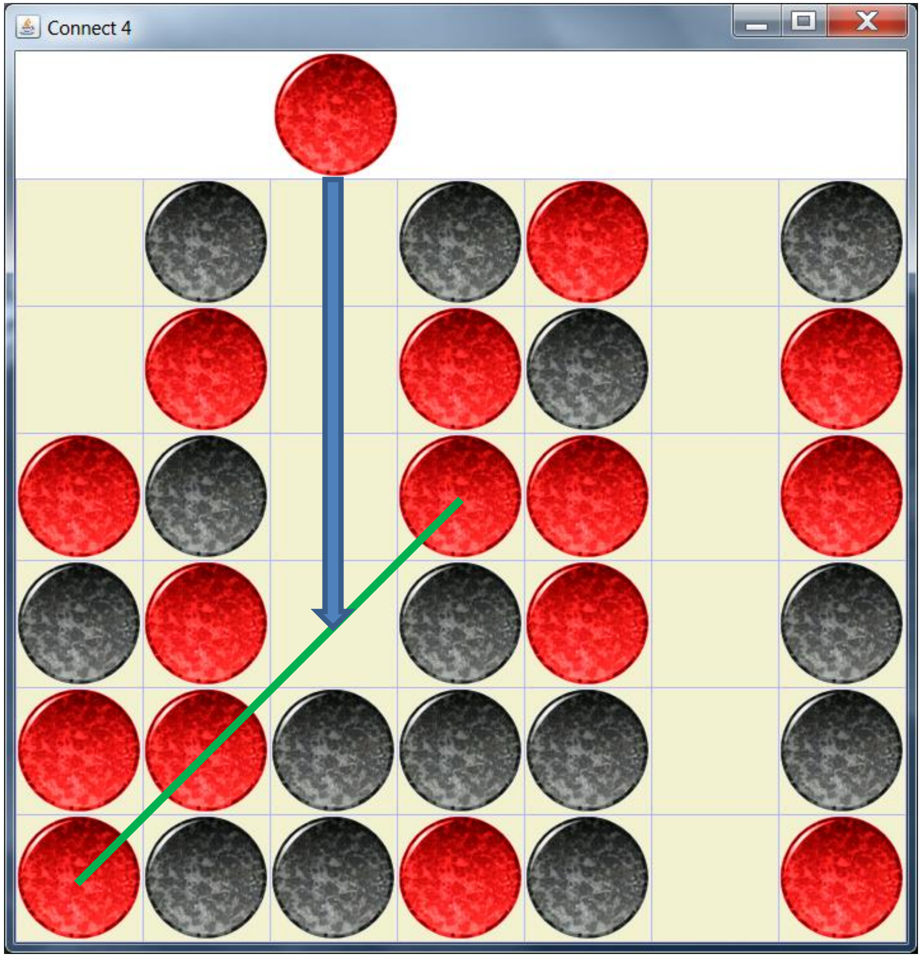
```
Player 1 discards 4H.
Player 1 has [[AC, 2C, 3C], [2S, 3S, 4S], [TC, TH, TD], [AH]] with 1 deadwood.
Player 1 melds [[AC, 2C, 3C], [2S, 3S, 4S], [TC, TH, TD]] with 1 deadwood from [AH].
Player 0 has 15 deadwood with [2H, 3H, 2D, 3D, 5D]
Player 0 melds [[4C, 5C, 6C, 7C, 8C]].
Player 1 scores the deadwood difference of 14.
Number of opponent cards known: 0
Number discarded: 12
Number of candidates: 30.0
>>>> est. 5.801766427069696E-4 unif. 1.693508780843028E-5 ratio 34.25885057520386
```

Example Output

Rank	A	2	3	4	5	6	7	8	9	T	J	Q	K
C	0.4894	0.5650	0.6587	FALSE	FALSE	FALSE	FALSE	FALSE	0.3366	0.4660	FALSE	0.1442	FALSE
H	0.4867	FALSE	FALSE	FALSE	0.2741	0.1352	FALSE	FALSE	0.2413	0.3678	FALSE	0.1727	FALSE
S	0.5438	0.5429	0.6530	0.3665	0.4975	0.2655	0.3018	0.0944	0.3211	0.4566	FALSE	FALSE	0.1257
D	0.4720	FALSE	FALSE	0.3248	FALSE	FALSE	0.1043	FALSE	0.1903	0.2938	0.1083	FALSE	FALSE

```
Player 1 discards 4H.
Player 1 has [[AC, 2C, 3C], [2S, 3S, 4S], [TC, TH, TD], [AH]] with 1 deadwood.
Player 1 melds [[AC, 2C, 3C], [2S, 3S, 4S], [TC, TH, TD]] with 1 deadwood from [AH].
Player 0 has 15 deadwood with [2H, 3H, 2D, 3D, 5D]
Player 0 melds [[4C, 5C, 6C, 7C, 8C]].
Player 1 scores the deadwood difference of 14.
Number of opponent cards known: 0
Number discarded: 12
Number of candidates: 30.0
>>>> est. 5.801766427069696E-4 unif. 1.693508780843028E-5 ratio 34.258850575203866
```

A Connect Four Example



Flat Monte Carlo

- Algorithm:
  - For each legal play action *a*:
    - For *n* samples:
      - Make play action *a*
      - Simulate a random game to completion
      - Accumulate the result (win = +1, loss = -1, draw = 0)
    - Compute the average result for *a*
  - Choose to play *a* with the maximum average result.
- Lesson: Randomly sampled actions can inform play.
- Why is sampling important and necessary for complex games?

https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.297.3086

DeepStack

- DeepStack and Libratus were the first two approaches to "solve" No-Limit Texas Hold-em Poker.
- Here, "solve" means that full-time play across a human lifetime could not find significant exploitable (suboptimal) play.

DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík[♠,♥,†], Martin Schmid[♠,♥,†], Neil Burch[♠], Viliam Lisý[♠,♣],
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling[♠,*]
[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada
[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic
[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

[†]These authors contributed equally to this work and are listed in alphabetical order.
[*]To whom correspondence should be addressed; E-mail: bowling@cs.ualberta.ca

Artificial intelligence has seen several breakthroughs in recent years, with games often serving as milestones. A common feature of these games is that players have perfect information. Poker is the quintessential game of imperfect information, and a longstanding challenge problem in artificial intelligence. We introduce DeepStack, an algorithm for imperfect information settings. It combines recursive reasoning to handle information asymmetry, decomposition to focus computation on the relevant decision, and a form of intuition that is automatically learned from self-play using deep learning. In a study involving 44,000 hands of poker, DeepStack defeated with statistical significance professional poker players in heads-up no-limit Texas hold'em. The approach is theoretically sound and is shown to produce more difficult to exploit strategies than prior approaches. [1]

https://arxiv.org/pdf/1701.01724.pdf

Structure of Poker Play

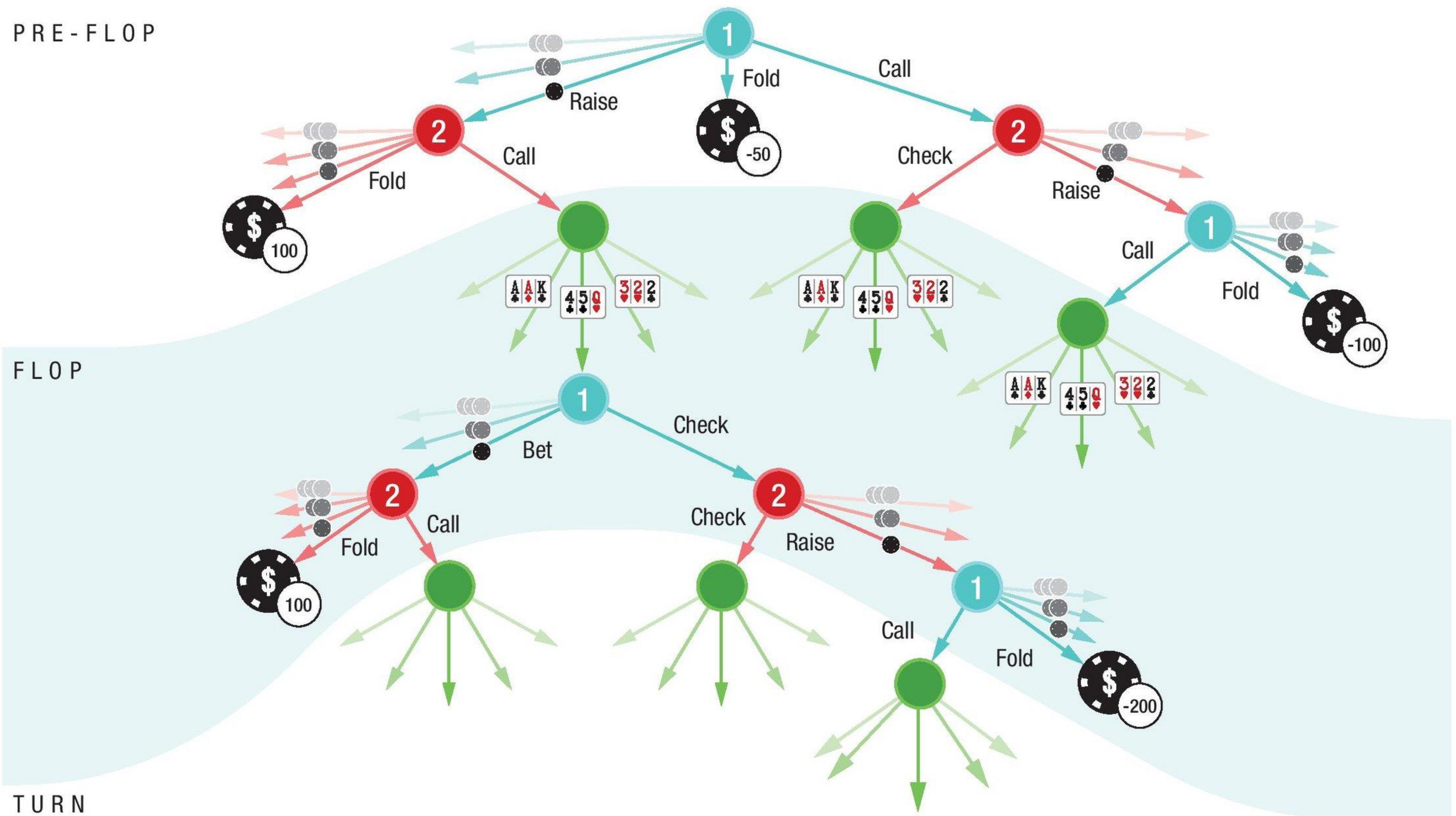


Image source: https://spencer-murray-zfht.squarespace.com/figures

Gin Rummy Game Tree Nodes

- Gin Rummy is also a game of imperfect information with choice and chance nodes.
- Types of choice nodes:
  - Take first up-card? (2 actions)
  - Draw face-up or face-down? (2)
  - Which card to discard? (11)
  - Whether or not to knock (when allowable)? (2)
  - (How to meld when knocking?) (≥1, often 1)
- Chance node:
  - Drawn face-down card (≤ 31)

DeepStack Operation

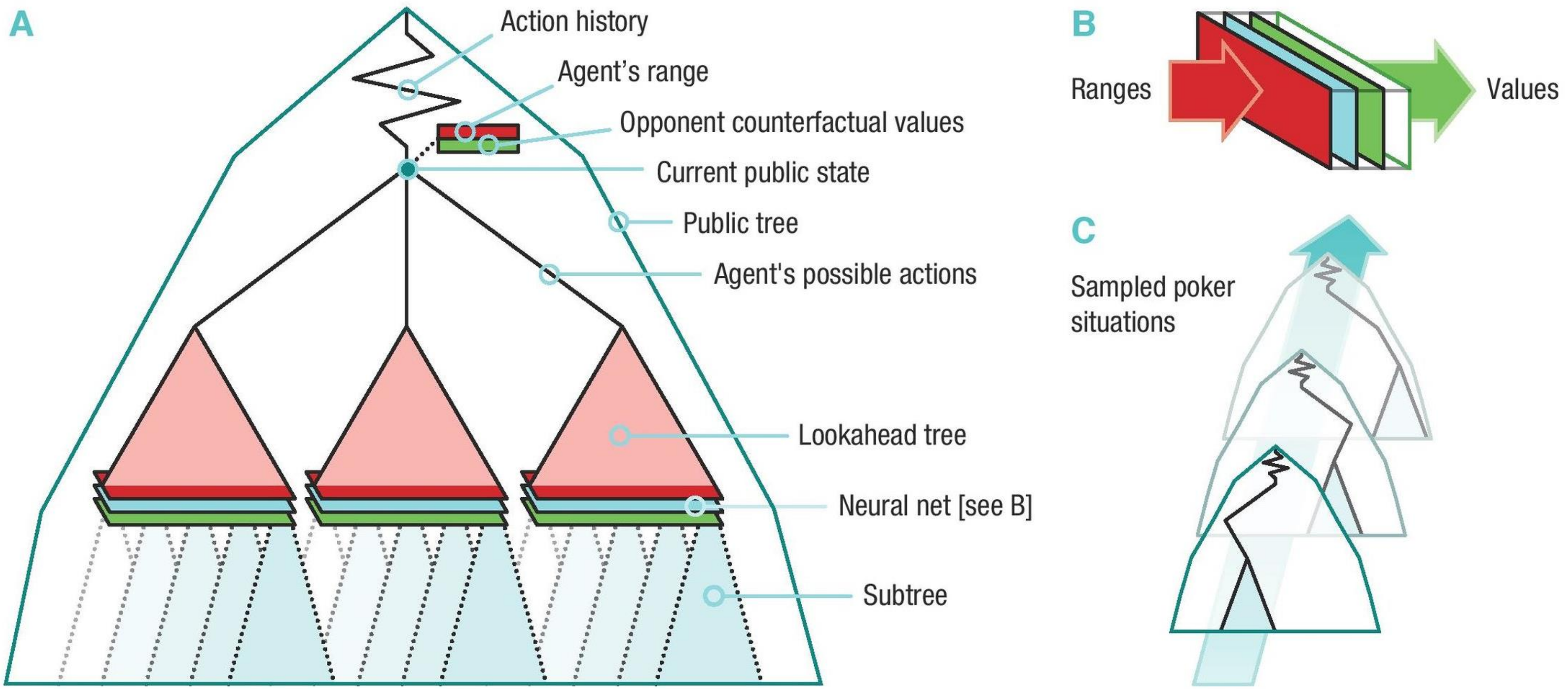


Image source: https://spencer-murray-zfht.squarespace.com/figures

DeepStack Operation

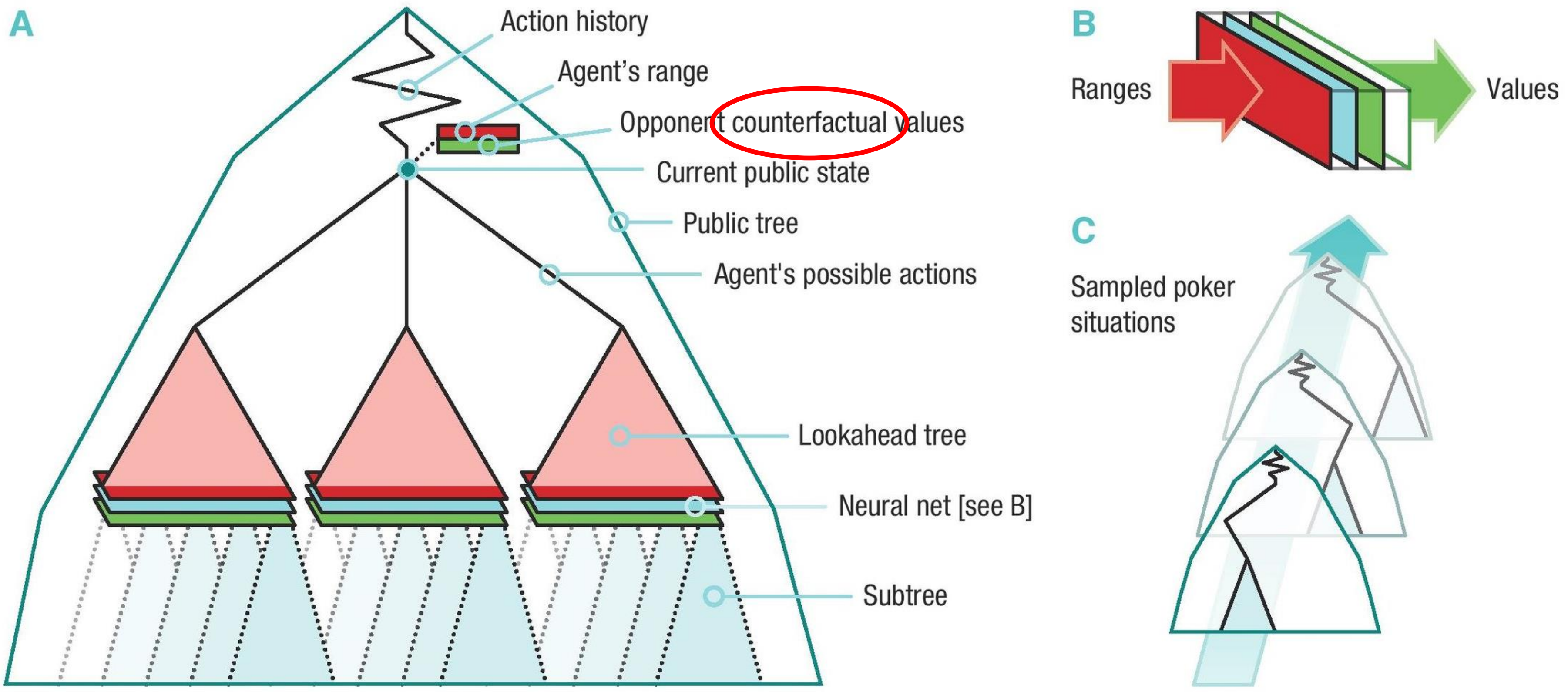


Image source: https://spencer-murray-zfht.squarespace.com/figures

Counterfactual Regret Minimization

- Predictability → exploitability
- Example: How often should one draw a face-up card that works towards meld(s)? Never? Always if low? Sometimes in some situations?
- Rock-Paper-Scissors (RPS)
  - 2 players, 3 possible simultaneous actions: rock (R), paper (P), scissors (S)
  - R, P, S beats S, R, P, respectively. Equal actions tie.
  - Win, tie, loss score +1, 0, -1, respectively

Regret

- Suppose you choose rock and your opponent chooses paper. Relative to your choice, how much do you regret not having chosen
  - paper?
  - scissors?
- Regret is the difference in utility between an action and your chosen action.
  - Initially **no regrets**, so picked from actions with equal probabilities.
  - New regrets: R→0 P→1 S→2

Regret Matching

- Choose an action with **probability proportional to positive regrets**.
- Regrets (0, 1, 2) normalized to probabilities: (0, 1/3, 2/3)
- Suppose we now choose S while our opponent chooses R.
  - Regrets: (1, 2, 0)
  - Cumulative regrets: (1, 3, 2)
  - Normalized cumulative regrets: (1/6, 3/6, 2/6)

Regret Minimization

- Regret Matching alone will not minimize regrets in the long run.
- However, the average strategy used over all iterations converges to a *correlated equilibrium.*
- In this example, average the strategies (1/3, 1/3, 1/3), (0, 1/3, 2/3), (1/6, 3/6, 2/6), etc.

DeepStack Operation

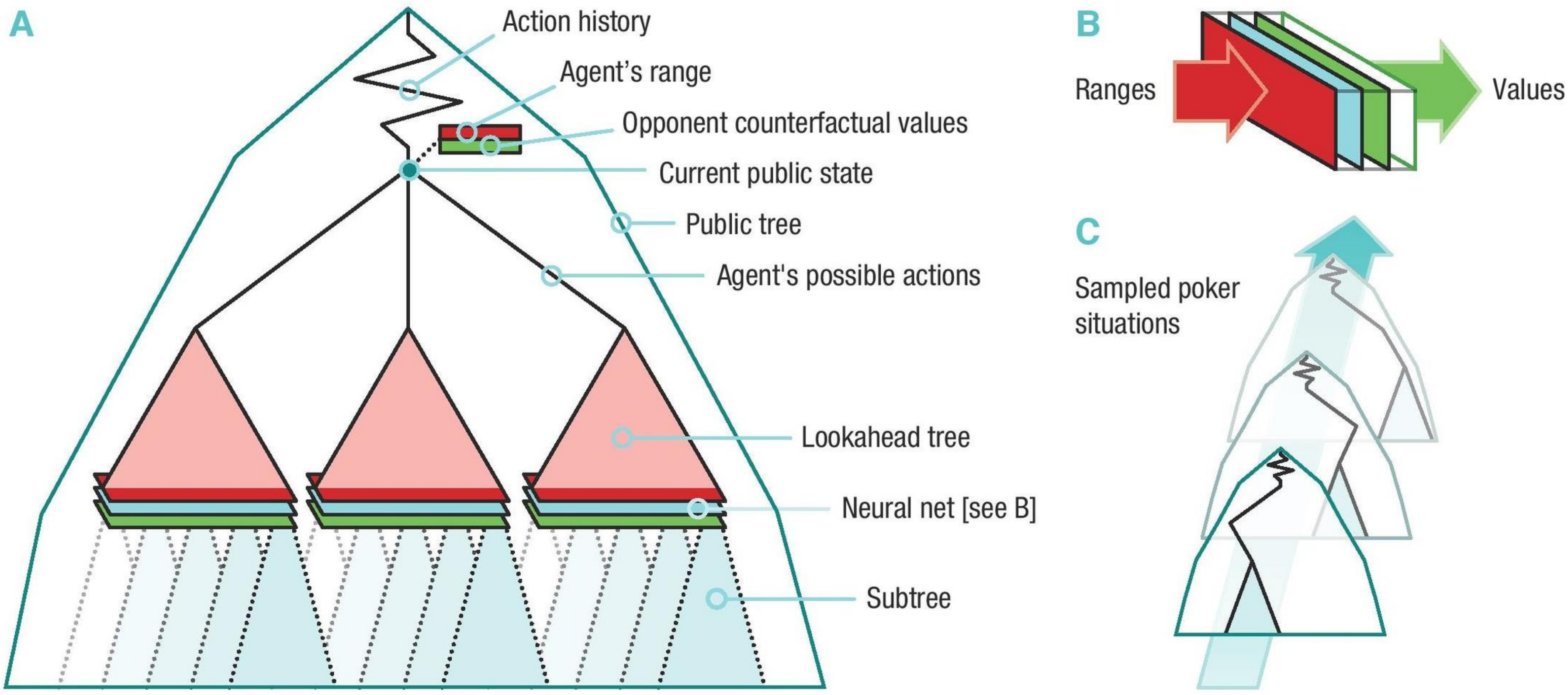


Image source: https://spencer-murray-zfht.squarespace.com/figures

Gin Rummy AI Research Topics

- Opponent Hand Estimation
  - Bayes' Rule Estimation
  - Markov Chain Monte Carlo
  - Particle Filtering
- Learning Game Values
  - Neural Networks
  - Gradient Boosted Decision Trees
- Sampling Techniques

The Gin Rummy AI Challenge

- This coming summer of 2020, I hope to get X-SIG funds to mentor students in Gin Rummy AI research.
- We will build an AI bot to compete against AI bots developed by other students at other institution.
- Goal: Successful bot and published work presented at EAAI, collocated at AAAI, the main general AI conference of this hemisphere.
- Stay tuned!

Questions?