



OVERVIEW

The game of **HANABI** is a cooperative card game ... and so on.

Research question

How good is Monte Carlo in playing Hanabi?

PLAYABILITY

We consider the simplified situation where, as everybody probably knows, the following formula represents the number of states:

$$\sum_{\substack{y \prec x \\ |y| \leq k-2}} \frac{(-1)^{|y|} (x_1 + x_2 + \dots + x_k)! (k - |y| - 1)^{x_1 + x_2 + \dots + x_k - \sum_{i=1}^k (y_i \ominus 1)}}{(y_1 \ominus 1)! (y_2 \ominus 1)! \dots (y_k \ominus 1)! (x_1 + x_2 + \dots + x_k - \sum_{i=1}^k (y_i \ominus 1))!}$$

Here we denote $t \ominus 1 = \max(t - 1, 0)$; and $y \prec x$ if the ordered sequence $y = (y_1, y_2, \dots, y_k)$ satisfies $y_i \leq x_i$ for all i ; and $|y|$ is the number of non-zero elements in y . And k is the maximal card value, of course.

RULES OF THE GAME

Cards

The "classic" game of HANABI is played with a stack of $N = 50$ cards of $C = 5$ types, each having $k = 10$ cards.



Antoine Bauza (2011); published by R & R Games

Game play

Every turn, a player chooses one action:

- Do this.
- Do that.
- Do nothing.
- Cheat.
- Do more.
- Don't do.
- Do it.
- They do run run.
- Do the locomotion.
- Do something else.
- Do.
- Dodo.

Goal

The goal of the game is to win, or to lose. The game ends if either:

- No error tokens remain: score = 0.
- All C piles are complete: score = $C \cdot k$.
- The stack is empty. All players get one more turn; then sum the highest card number in each pile for the score.

RULE-BASED STRATEGY

For the **rule-based strategy**, every player acts according to the following preset rules:

1. If there is a card in my hand of which I am "certain enough" that it can be played, I play it.
2. Otherwise, if there is a card in my hand of which I am "certain enough" that it is useless, I discard it.
3. Otherwise, if there is a hint token available, I give a hint.
4. Otherwise, I discard a card.

Plaatjes, grafieken, ...

We show resulting scores for different parameter settings; $P = 3$, $R = 5$. The highest average score obtained is 15.4.

MONTE CARLO STRATEGY

The **Monte Carlo player** (who does the move with the best average score during random play-outs) uses the following refinements:

- When playing a card in the Monte Carlo phase, the hand of the current player is shuffled through the deck and a new hand is dealt which is consistent with all hint information obtained so far.
- During play-outs, the game does not end after 3 errors.
- The random player does "reasonable" moves.
- Only the score of the next D turns is taken into account to value a play-out.

The highest average score obtained is 14.5.

SAMPLE GAME STATE

North: ♠1 ♥3 ♦1 ♠4 ♣3

Table: ♥1 ♥2

West: ♠5 ♣3 ♥1 ♣4 ♣1 East: ♦1 ♥4 ♠2 ♠2 ♠4

South (me): ? ? ?2 ?2 ?≠2

South may give North a hint about his/her 1s.

Waarom is dit leeg?

FURTHER RESEARCH

Playability: Try to generalize the given formula for $R > 1$, and find an intuitive proof using the principle of in- and exclusion.

Strategies: Improve the Monte Carlo player using MCTS or learning methods. Moreover, one can delve deeper in the information truly contained in a given hint (cf. conventions).

Is er nog ergens plek voor Related work?

Voldoende plaatjes, grafieken, ...?

Deze poster heeft wel erg veel tekst!

Meer kleur?

[lastig] En wellicht een conclusie?

Nog meer plaatjes.

Welk block is geschikt om voor de "pitch" te gebruiken?