Aspects of the cooperative card game Hanabi

Steven Student (supervized by Bert E. Geleider)

Leiden Institute of Advanced Computer Science

s.student@liacs.leidenuniv.nl

Overview

The game of Hanabi is a cooperative card game in which players together try to obtain the highest possible score. The goal of the game is simple: play out several sequences of cards in the right order. The catch, however, is that the **players can only see the cards in other player's hands and not their own**; information has to be gathered by a **system of hints** that reveal partial information.

Rules of the game

Materials and setup

The "classic" game of Hanabi is played with a stack of $N = 50$ cards. Every card has one out of $C = 5$ colors and a value between 1 and $k = 5$. There are $H = 8$ hint tokens and $E = 3$ error tokens.
At the start of the game, the stack is shuffled and a hand of $R = 4$ or $R = 5$ cards is dealt to each of the P players, $P \in \{2, 3, 4, 5\}$. Now, every player picks up his/her cards in such a way that the other players can see them, but they themselves cannot. The rest of the cards forms the face-down stack. All hint and error tokens are initially available.

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

Game play

Every turn, a player chooses one action:

- Give a hint: expend a hint token to point out all cards of a certain color or value in the hand of one other player.
- Discard a card: move a card from the hand to the discard pile, regain a hint token and draw a new card from the stack.
- Play a card: add a card from the hand to its pile on the table (if it fits) or spend an error token; then draw a new card.

Goal and game end

The goal of the game is to create C piles of cards going from 1 through k on the table, one of each color. 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.

Playability

We consider the simplified situation where the players can also see their own cards. We give a result for the one-player version, with $R = 1$: a player must immediately play or discard the newly received card. We consider only one color, but the number of cards of each value can be arbitrary. Hence, the initial stack is now a random sequence of integers in $\{1, \ldots, k\}$.

The maximum score can be obtained if and only if a subsequence 1–2–...–k exists. The number of ordered sequences with x_1 occurrences of the integer 1, x_2 occurrences of the integer 2, ..., x_k occurrences of the integer k, *without* a subsequence 1–2–...–k, equals

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

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

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
♣1

West:
♠5 ♣3 ♡1 ♣4 ♣1

East:
♢1 ♡4 ♠2 ♠2 ♠4

South (me): ? ? $?_2$ $?_2$ $?_{\neq 2}$

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?

Waar staat de onderzoeksvraag?

Voldoende plaatjes, grafieken, ...?

Deze poster heeft wel erg veel tekst!

Meer kleur?

[lastig] En wellicht een conclusie?