

Hanabi is NP-complete, even for cheaters who look at their cards

Citation for published version (APA):

Baffier, J.-F., Chiu, M. K., Diez, Y., Korman, M., Mitsou, V., van Renssen, A., Roeloffzen, M., & Uno, Y. (2016). Hanabi is NP-complete, even for cheaters who look at their cards. In E. D. Demaine, & F. Grandoni (Eds.), *8th International Conference on Fun with Algorithms, FUN 2016* (Vol. 49). Article 4 (Leibniz International Proceedings in Informatics (LIPIcs); Vol. 49). Schloss Dagstuhl - Leibniz-Zentrum für Informatik. <https://doi.org/10.4230/LIPIcs.FUN.2016.4>

Document license:

CC BY

DOI:

[10.4230/LIPIcs.FUN.2016.4](https://doi.org/10.4230/LIPIcs.FUN.2016.4)

Document status and date:

Published: 01/06/2016

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Hanabi is NP-Complete, Even for Cheaters Who Look at Their Cards

Jean-Francois Baffier^{1,9}, Man-Kwun Chiu^{2,9}, Yago Diez³,
Matias Korman⁴, Valia Mitsou⁵, André van Renssen^{6,9},
Marcel Roeloffzen^{7,9}, and Yushi Uno⁸

- 1 National Institute of Informatics (NII), Tokyo, Japan
jf_baffier@nii.ac.jp
- 2 National Institute of Informatics (NII), Tokyo, Japan
chiu@nii.ac.jp
- 3 Tohoku University, Sendai, Japan
yago@dais.is.tohoku.ac.jp
- 4 Tohoku University, Sendai, Japan
mati@dais.is.tohoku.ac.jp
- 5 SZTAKI, Hungarian Academy of Sciences, Hungary
vmitsou@sztaki.hu
- 6 National Institute of Informatics (NII), Tokyo, Japan
andre@nii.ac.jp
- 7 National Institute of Informatics (NII), Tokyo, Japan
marcel@nii.ac.jp
- 8 Department of Mathematics and Information Sciences, Graduate School of
Science, Osaka Prefecture University, Japan
uno@mi.s.osakafu-u.ac.jp
- 9 JST, ERATO, Kawarabayashi Large Graph Project

Abstract

This paper studies a cooperative card game called Hanabi from an algorithmic combinatorial game theory viewpoint. The aim of the game is to play cards from 1 to n in increasing order (this has to be done independently in c different colors). Cards are drawn from a deck one by one. Drawn cards are either immediately played, discarded or stored for future use (overall each player can store up to h cards). The main feature of the game is that players know the cards their partners hold (but not theirs. This information must be shared through hints).

We introduce a simplified mathematical model of a single-player version of the game, and show several complexity results: the game is intractable in a general setting even if we forego with the hidden information aspect of the game. On the positive side, the game can be solved in linear time for some interesting restricted cases (i.e., for small values of h and c).

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems, G.2 Discrete Mathematics, F.1.2 Modes of Computation

Keywords and phrases algorithmic combinatorial game theory, sorting

Digital Object Identifier 10.4230/LIPIcs.FUN.2016.4

1 Introduction

When studying mathematical puzzles or games, mathematicians and computer scientists are often interested in winning strategies, designing computer programs that play as well (or even better than) humans. The computational complexity field studies the computational



© Jean Francois Baffier, Man-Kwun Chiu, Yago Diez, Matias Korman, Valia Mitsou,
André van Renssen, Marcel Roeloffzen, and Yushi Uno;
licensed under Creative Commons License CC-BY

8th International Conference on Fun with Algorithms (FUN 2016).

Editors: Erik D. Demaine and Fabrizio Grandoni; Article No. 4; pp. 4:1–4:17



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

complexity of the games. That is, how hard it is to obtain a solution to a puzzle or to decide the winner or loser of a game [6, 10, 11]. Another interest is to design algorithms to obtain solutions. Some games and puzzles of interest include for example Nim, Hex, Sudoku, Tetris, and Go. Recently, this field has been called ‘algorithmic combinatorial game theory’ [11] to distinguish it from games arising from other fields, especially classical economic game theory.

In this paper we study a cooperative card game called Hanabi. Designed by Antoine Bauza and published in 2010, the game has received several tabletop game awards (including the prestigious *Spiel des Jahres* in 2013 [8]). In the game the players simulate a fireworks show¹, playing cards of different colors in increasing order.

In this paper we study the game from the viewpoint of algorithmic combinatorial game theory. We first propose mathematical models of a single-player variant of Hanabi, and then analyze their computational complexities. As done previously for other multiplayer card games [7, 12], we show that even a single-player Hanabi is computationally intractable in general, while the problem becomes easy under very tight constraints.

1.1 Rules of the Game

Hanabi is a multi-player, imperfect-information cooperative game. This game is played with a deck of fifty cards, where each card has a number (from 1 to 5) and a suit (or a color) out of five colors (red, yellow, green, blue and white). There are ten cards of each suit. The values of the cards are 1, 1, 1, 2, 2, 3, 3, 4, 4, and 5, respectively. That is, there are two copies of each card except for the lowest and highest cards of each color (that appear three and one time, respectively). Players must cooperate to play the cards from 1 to 5 in increasing order for all suits independently.

One of the most distinctive features of the game is that players cannot see their own cards while playing: each player holds his cards so that they can be seen by other players (but not himself). A player can do one of the following actions in each turn: play a card, discard a card from his hand to draw a new one, or give a hint to another player on what type of cards he or she is holding in hand. A second characteristic of this game is that each player can hold only a small number of cards in hand (4 or 5 depending on the number of players) drawn at random from the deck. Whenever no card is playable, a player may discard a card and draw a new one from the deck. See Appendix A for the exact rules of the Hanabi game, or [1, 2] for more information on the game.

1.2 Related Work

There is an extensive amount of research that studies the complexity of tabletop and card games. In virtually all games, the total description complexity of the problem is bounded by a constant, thus they can be solved in constant time by an exhaustive search approach.

Thus, the literature focuses on the extensions of those games in which the complexity is not constant. For example, it is known that determining the winner in chess on an $n \times n$ board needs exponential time [9]. If the playing board can be any graph, *Pandemic* (a popular tabletop game in which players try to prevent a virus from spreading) is NP-complete [13]. A somehow more surprising result is that determining the winner in *Settlers of Catan* is also an NP-complete problem, even after the game has ended [5].

When considering card games, the complexity is often expressed as a function of the number of cards in the deck. The popular trading card game *Magic: The Gathering* is Turing

¹ The word *hanabi* means fireworks in Japanese.

complete [3]. That is, it can simulate a Turing machine (and in particular, it can simulate any other tabletop or card game).

There is little research that studied algorithmic aspects of Hanabi. Most of the existing research [14, 4] propose different strategies so that players can share information and collectively play as many cards as possible. Several heuristics are introduced, and compared to either experienced human players or to optimal play sequences (assuming all information is known).

Our approach diverges from the aforementioned studies. We show that, even if we forego its hidden information trademark feature, the game is intractable, which means that there is an intrinsic difficulty in Hanabi beyond information exchange. In fact we show hardness for a simplified solitaire version of the game where the single player has complete information about which cards are being held in his hand as well as the exact order in which cards will be drawn from the deck.

1.3 Model and Definitions

We represent a card of Hanabi with an ordered pair (a_i, k_i) , where $a_i \in \{1, \dots, n\}$ and $k_i \in \{1, \dots, c\}$. The term a_i is referred to as the *value* of the card and k_i as its *color*. The whole deck of cards is then a sequence σ of N cards. That is, $\sigma = ((a_1, k_1), \dots, (a_N, k_N))$. The *hand size* h is the maximum number of cards that the player can hold in hand at any point during the game. The *multiplicity* r of cards in a card sequence σ is the maximum number of times that any card appears in σ .

In a game, the player scans the cards in the order fixed by σ in a streaming fashion. In each turn a player has three options: play a card from their hand, discard one to get a hint token, or give a hint. After his turn, he draws a new card to replace the played/discarded one. As our model drops completely the information sharing feature of the game, we replaced the hinting move with a move where the single player takes no action thus, ‘storing’ the card. Since we are focused in the single player case and turn order does not matter, we allow doing several actions before redrawing. The three available options are thus: *play*, *discard* or *store* the card. If a card is discarded, it is gone and can never be used afterwards. If instead we store the card, it is saved and can be accessed afterwards (remember that at any instant of time the maximum number of cards that can be stored in hand is h). Cards can be played only in increasing order for each color independently. That is, we can play card (a_i, k_i) if and only if the last card of color k_i that was played was $(a_i - 1, k_i)$ (or $a_i = 1$ and no card of color k_i has been played). After a card has been played we can also play any cards we may have stored in hand in the same manner. The objective of the game is to play all cards from 1 to n in all c colors. Whenever this happens we say that the sequence of play/discard/store is a *winning play sequence*.

Thus, a problem instance of the SOLITAIRE HANABI (or HANABI for short) consists of a hand size $h \in \mathbb{N}$ and a card sequence σ of N cards (where each card is an ordered pair of a value and a color out of n numbers and c colors, and no card appears more than r times). The aim is to determine whether or not there is a winning play sequence for σ that never stores more than h cards in hand.

1.4 Results and Organization

In this paper, we study computational complexity and algorithmic aspects of HANABI with respect to parameters N , n , c , r and h . Unfortunately the problem is NP-complete, even if

■ **Table 1** Summary of the different results presented in this paper, where N , n , c , r and h are the number of cards, the number of values, the number of colors, multiplicity, and the hand size, respectively.

Case Studied	Approach Used	Running Time	Observations
$r = 1$	Greedy	$O(N) = O(cn)$	Lemma 1 in Sec. 2
$c = 1$	Lazy	$O(N + n \log h)$	Theorem 4 in Sec. 3
General Case	Dynamic Programming	$O(Nhc^h n^{h+c-1})$	Theorem 6 in Sec. 4
$h = 2, r = 2$	NP-complete		Theorem 7 in Sec. 5

we fix some parameters to be small constants. Specifically, in Section 5 we show that the problem is NP-complete even if we restrict ourselves to the case in which $h = 2$ and $r = 2$.

Given the negative results, we focus on the design of algorithms for particular cases. For those cases, our aim is to design algorithms whose running time is linear in N (the total number of cards in the sequence), but we allow slightly larger running times as a function of n , k , and r (the total number of values, colors and multiplicity, respectively).

In Section 2 we give a straightforward $O(N)$ algorithm for the case in which $r = 1$ (that is, no card is repeated in σ). This approach is afterwards extended for $c = 1$ (and unbounded r) in Section 3. In Section 4 we give an algorithm for the general problem. Note that the algorithm runs in exponential time (expected for an NP-complete problem), but the running time reduces to $O(N)$ whenever, h , c and n are constants. The exact running times of all algorithms introduced in this paper are summarized in Table 1.

2 Unique Appearance

As a warm-up, we consider the case in which each card appears only once (i.e., $r = 1$). In this case we have exactly one card for each value and each color. Thus, $N = cn$ and the input sequence σ is a permutation of the values from 1 to n in the c colors.

Since each card appears only once, we cannot discard any card in the winning play sequence. In the following, we show that the natural greedy strategy is essentially the best we can do: play a card as soon as it is found (if possible). If not, store it in hand until it can be afterwards played.

The game rules state that we cannot play a card (a_i, k_i) until all the cards from 1 to $a_i - 1$ of color k_i have been played. Thus, we associate an interval to each card that indicates for how long that card must be held in hand. For any card (a_i, k_i) , let f_i be the largest index of the cards of color k_i whose value is at most a_i (i.e., $f_i = \max_{j \leq N} \{j : k_j = k_i, a_j \leq a_i\}$). Note that we could have $i = f_i$, but this only happens when all cards of value smaller than a_i appear before card (a_i, k_i) . Otherwise, we must have $f_i > i$, and card (a_i, k_i) cannot be played until we have reached card (a_{f_i}, k_{f_i}) .

We associate each index i to the interval $[i, f_i]$. Let \mathcal{I} be the collection of all nonempty such intervals. Let w be the maximum number of intervals that overlap (i.e., $w = \max_{j \leq N} |\{[i, f_i] \in \mathcal{I} : j \in [i, f_i]\}|$).

► **Lemma 1.** *There is a solution to any HANABI problem instance with $r = 1$ and hand size h if and only if $w \leq h$. Moreover, a play sequence can be found in $O(N)$ time.*

Proof. Intuitively speaking, any interval $[i, j] \in \mathcal{I}$ represents the need of storing card (a_i, k_i) until we have reached card (a_j, k_j) . Thus, if two (or more) intervals overlap, then the corresponding cards must be stored simultaneously. By definition of w , when processing the

input sequence at some point in time we must store more than h cards, which in particular implies that no winning play sequence exists.

In order to complete the proof we show that the greedy play strategy works whenever $w \leq h$. The key observation is that, for any index i we can play card (a_i, k_i) as soon as we have reached the f_i -th card. Indeed, by definition of f_i all cards of the same color whose value is a_i or less have already appeared (and have been either stored or played). Thus, we can simply play the remaining cards (including (a_i, k_i)) in increasing order.

Overall, each card is stored only within its interval. By hypothesis, we have $w \leq h$, thus we never have to store more than our allowed hand size. Furthermore, no card is discarded in the play sequence, which in particular implies that the greedy approach will give a winning play sequence with hand size h .

Regarding running time, it suffices to show that each element of σ can be treated in constant time. For the purpose, we need a data structure that allows insertions into \mathcal{H} and membership queries in constant time. The simplest data structure that allows this is a hash table. Since we have at most h elements (out of a universe of size cn) it is easy to have buckets whose expected size is constant.

The only drawback of hash tables is that the algorithm is randomized (and the bounds on the running time are expected). If we want a deterministic worst case algorithm, we can instead represent \mathcal{H} with a $c \times n$ bit matrix and an integer denoting the number of elements currently stored. With either data structure it is straightforward to see that insertions, removals, and membership queries take constant time, thus the algorithm takes $O(N) = O(cn)$ time as claimed. ◀

3 Lazy Strategy for One Color

We now study the case in which all cards belong to the same suit (i.e., $c = 1$). Note that we make no assumptions on the multiplicity or any other parameters. Unlike the last section in which we considered a greedy approach, here we describe a lazy approach that plays cards at the last possible moment.

We start with an observation that allows us to detect how important a card is. For any $i \leq N$, we say that the i -th card (whose value is a_i) is *useless* if there exist $w_1, \dots, w_{h+1} \in \mathbb{N}$ such that:

- (i) $a_i < w_1 < \dots < w_{h+1} \leq n$
- (ii) $\forall j \in \{i+1, \dots, N\}$ it holds that $a_j \notin \{w_1, \dots, w_{h+1}\}$

That is, there exist $h+1$ values that are higher than a_i none of which appears after the i -th card in σ . Observe that, for example, no card of value $n-h$ or higher can be useless (since the w_i values cannot exist) and that the last card is useless if and only if $a_N < n-h$.

► **Observation 2.** *Useless cards are never played in a winning play sequence.*

Proof. Assume, for the sake of contradiction, that there exists a winning play sequence that plays some useless card whose index is i . Since we play cards in increasing order, no card of value equal to or bigger than a_i can have been played at the time in which the i -th card is scanned. By definition of useless, the remaining sequence does not have more cards of values w_1, \dots, w_{h+1} . Thus, in order to complete the game to a winning sequence, these $h+1$ cards must all have been stored, but this is not possible with a hand size of h . ◀

Our algorithm starts with a filtering phase that removes all useless cards from σ . The main difficulty of this phase is that the removal of some useless cards from σ may create further useless cards, and so on. In order to avoid having to scan the input several times

we use two vectors and a max-heap as follows: for each index $i \leq N$, we store the index of the previous occurrence of the same card in a vector P (or $-\infty$ if none exists). That is, $P[i] = -\infty$ if and only if $a_j \neq a_i$ for all $j < i$. Otherwise, we have $P[i] = i'$ (for some $i' < i$), $a_i = a_{i'}$, and $a_j \neq a_i$ for all $j \in \{i' + 1, \dots, i - 1\}$. We also use a vector L such that each index $i \leq n$ stores the last non-useless card of value i (since initially no card has been detected as useless, the value $L[i]$ is initialized to the index of the last card with value i in σ). Finally, we use a max-heap \mathcal{HP} of $h + 1$ elements initialized with values $L[n - h], \dots, L[n]$.

Now, starting with $i = n - (h + 1)$ down to 1 we look for all useless copies of value i . The invariant of the algorithm is that for any $j > i$, all useless cards of value j have been removed from σ and that vector $L[j]$ stores the index of the last non-useless card of value j . The heap \mathcal{HP} contains the smallest $h + 1$ values among $L[j], \dots, L[n]$ (and since it is a max-heap we can access in constant time its largest value). These values will be the smallest possible candidate values for the witnesses w_1, \dots, w_{h+1} (properly speaking, \mathcal{HP} stores indices, but the values can be extracted in constant time). The invariants are satisfied for $i = n - (h + 1)$ directly by the way in which L and \mathcal{HP} is initialized.

Any card of value i whose index is higher than the top of the heap is useless and can be removed from σ (the indices in the heap \mathcal{HP} act as witnesses). Starting from $L[i]$, we remove all useless cards of value i from σ until we find a card of value i whose index is smaller than the top of the heap. If no card of value i remains we stop the whole process and return that the problem instance has no solution. Otherwise, we have found the last non-useless card of value i . We update the value of $L[i]$ since we have just found the last non-useless card of that value. Finally, we must update the heap \mathcal{HP} . As observed above, the value of $L[i]$ must be smaller than the largest value of \mathcal{HP} (otherwise it would be a useless card). Thus, we remove the highest element of the heap, and insert $L[i]$ instead. Once this process is done, we proceed to the next value of i . Let σ' be the result of filtering σ with the above algorithm.

► **Lemma 3.** *The filtering phase removes only useless cards from σ . Moreover, this process runs in $O(N + n \log h)$ time, and σ' contains no useless cards.*

Proof. Each time we remove a card from σ , the associated $h + 1$ witnesses w_1, \dots, w_{h+1} are present in \mathcal{HP} , thus the first claim follows. The fact that no more useless cards remain follows from the fact that we always store the smallest possible witness values.

Now we bound the running time. The heap is initialized with $h + 1$ elements, and during the whole depurating phase $O(n)$ elements are pushed. Hence, this part takes $O(n \log h)$ time. Vector P and L can be initialized by scanning σ once. During the iterative phase we can access the last occurrence of any value by using vector L . Once a card is removed, we can update the last occurrence using P . Thus, we spend constant time per card that is removed from σ (hence, overall $O(N)$ time). ◀

Now we describe the algorithm for our lazy strategy. The play sequence is very simple: we ignore all cards except when a card is the last one of that value present in σ' . For those cards, we play them (if possible) or store them (otherwise). Whenever we play a card, we play as many cards as possible (out of the ones we had stored).

Essentially, there are two possible outcomes after the filtering phase. It may happen that all cards of some value were detected as useless. In this case, none of those cards may be played and thus the HANABI problem instance has no solution. Otherwise, we claim that our lazy strategy will yield a winning play sequence.

► **Theorem 4.** *We can solve a HANABI problem instance for the case in which all cards have the same color (i.e., $c = 1$) in $O(N + n \log h)$ time.*

Proof. It suffices to show that our lazy strategy will always give a winning play sequence, assuming that the filtered sequence contains at least a card of each value. Our algorithm considers exactly one card of each value from 1 to n . The card will be immediately played (if possible) or stored until we can play it afterwards. Thus, the only problem we might encounter would be the need to store more than h cards at some instant of time.

However, this cannot happen: assume, for the sake of contradiction, that at some instant of time we need to store a card (whose index is j) and we already have stored cards of values a_{i_1}, \dots, a_{i_h} . By construction of the strategy, there cannot be more copies of cards with value a_{i_1}, \dots, a_{i_h} or a_j in the remaining portion of σ' . Let p be the number of cards that we have played at that instant of time. Remember that we never store a card that is playable, thus $p + 1 \notin \{a_j, a_{i_1}, \dots, a_{i_h}\}$. In particular, the last card of value $p + 1$ must be present in the remaining portion of σ' . However, that card is useless (the values $\{a_j, a_{i_1}, \dots, a_{i_h}\}$ act as witnesses), which gives a contradiction.

Thus, we conclude that the lazy strategy will never need to store more than h cards at any instant of time, and it will yield a winning play sequence. Finally, observe that the sequence itself can be reconstructed, since vector L stores the last non-useless occurrence of each value. ◀

4 General Case Algorithm

In this section we study the general problem setting. Recall that this problem is NP-complete, even if the hand size is small (see details in Section 5), hence we cannot expect an algorithm that runs in polynomial-time. In the following, we give an algorithm that runs in polynomial time provided that both h and c are fixed constants (or exponential otherwise).

We solve the problem using a dynamic programming approach. Specifically, we build a table $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}]$ indexed by the following $c + 1$ parameters:

- s ($\leq N$) represents the number of cards from the sequence σ that we allow to scan.
- \mathcal{H} is the set of cards that we require to store in hand after card (a_s, k_s) has been processed. We might have no requirements on what needs to be in hand, in which case we simply set $\mathcal{H} = \emptyset$.
- p_1, \dots, p_{c-1} ($\leq n$) encode how many cards we require to play in the first $c - 1$ colors, respectively.

The entry of the table $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}]$ is a positive number equal to the maximum number of cards of the c -th color that we can play among all play sequences that preserve the above constraints. Whenever such a sequence is not feasible (i.e., we cannot play the required cards in some color or store the cards of \mathcal{H}), we simply set that position of the table to $-\infty$.

For example, if $c = 3$, the entry of table $DP[42, \{(15, 1), (10, 2)\}, 10, 4] = 6$ should be interpreted as *There is a play sequence that, after scanning through the 42 cards of σ has played exactly 10 cards of the first color, 4 of the second, 6 of the third, and has stored cards (15, 1) and (10, 2) in hand. Moreover, there is no play sequence that, after scanning the first 42 cards, plays 10, 4, and 7 cards of the three colors (respectively) and ends up with cards (15, 1) and (10, 2) in hand.*

When s is a small number we can find the solution of an entry by brute force (try all possibilities of discarding, storing or playing the first s cards). This takes constant time since the problem has constant description complexity. Similarly, we have $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = -\infty$ whenever $|\mathcal{H}| > h$ (because we need to store more than h cards in hand). In the following

we show how to compute the table DP for the remaining cases. For this purpose, we define three auxiliary values \mathcal{D} , \mathcal{S} , and \mathcal{P} (which stand for *Discard*, *Store*, and *Play*) as follows:²

$$\mathcal{D} = DP[s - 1, \mathcal{H}, p_1, \dots, p_{c-1}]$$

$$\mathcal{S} = \begin{cases} -\infty & \text{if } |\mathcal{H}| \geq h \\ DP[s - 1, \mathcal{H} \setminus \{(a_s, k_s)\}, p_1, \dots, p_{c-1}] & \text{otherwise} \end{cases}$$

and

$$\mathcal{P} = \begin{cases} -\infty & \text{if } k_s < c, a_s > p_{k_s} \\ DP[s - 1, \mathcal{H} \cup \{(a_s + 1, k_s), \dots, (p_{k_s}, k_s)\}, \\ \quad p_1, \dots, p_{k_s-1}, a_s - 1, p_{k_s+1}, \dots, p_{c-1}] & \text{if } k_s < c, a_s \leq p_{k_s} \\ \max_{t \in \{0, \dots, h\}} \{a_s + t: DP[s - 1, \mathcal{H} \cup \{(a_s + 1, c), \dots, (a_s + t, c)\}, \\ \quad p_1, \dots, p_{c-1}] \geq a_s - 1\} & \text{if } k_s = c \end{cases}$$

These auxiliary values allow us to compute an entry of the table efficiently.

► **Lemma 5.** $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = \max\{\mathcal{P}, \mathcal{S}, \mathcal{D}\}$.

Proof. Assume that $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}]$ is a positive number (i.e., it is feasible to satisfy all constraints). Consider any play sequence that realizes it, we distinguish three cases depending on what the play sequence does with card (a_s, k_s) :

(a_s, k_s) is discarded. When the last card is discarded, the entry of the table is the same as if we only allow the scanning of $s - 1$ cards. Thus, $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = DP[s - 1, \mathcal{H}, p_1, \dots, p_{c-1}] = \mathcal{D}$.

(a_s, k_s) is stored. Since this card is the last one that we are allowed to scan for the entry of the table that we are computing, storing (a_s, k_s) only makes sense if $(a_s, k_s) \notin \mathcal{H}$. Moreover, this operation is only possible if we do not exceed the hand size limit. That is, we have $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = DP[s - 1, \mathcal{H} \setminus \{(a_s, k_s)\}, p_1, \dots, p_{c-1}]$ if $|\mathcal{H}| < h$ (otherwise it should be $-\infty$). This coincides with the definition of \mathcal{S} .

(a_s, k_s) is played. In this case we claim that $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = \mathcal{P}$. In order to prove this we give an intuitive definition of \mathcal{P} . We consider three subcases depending on the color and value of card (a_s, k_s) .

$k_s < c$ and $a_s > p_{k_s}$. Recall we only need to play up to card p_{k_s} in color k_s (and this card is of higher value). In particular, the card need not be played in the play sequence. We set $\mathcal{P} = -\infty$ to make sure that this case is not considered by our algorithm.

$k_s < c$ and $a_s \leq p_{k_s}$. Consider only color k_s : we are required to play p_{k_s} cards and in order to do that we must specifically use card (a_s, k_s) . In order to do so, we must have played the first $a_s - 1$ cards of this color in advance, and must have the cards from $a_s + 1$ to p_{k_s} in hand. All constraints for other colors are unaffected. Thus, we have $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = DP[s - 1, \mathcal{H} \cup \{(a_s + 1, k_s), \dots, (p_{k_s}, k_s)\}, p_1, \dots, p_{k_s-1}, a_s - 1, p_{k_s+1}, \dots, p_{c-1}] = \mathcal{P}$ as claimed.

² Note that, strictly speaking, these terms depend on the parameters $s, \mathcal{H}, p_1, \dots, p_{c-1}$. Thus, a better notation would be $\mathcal{P}_{s, \mathcal{H}, p_1, \dots, p_{c-1}}$, $\mathcal{S}_{s, \mathcal{H}, p_1, \dots, p_{c-1}}$, and $\mathcal{D}_{s, \mathcal{H}, p_1, \dots, p_{c-1}}$. Since the subindices are clear from the context, we remove them for ease of reading.

$k_s = c$. This case is similar to the previous one. In this case we focus on color c ($= k_s$): as before, we need card (a_s, k_s) to be playable when we reach this card. This constraint is realized by restricting to entries of the table that allow to play at least $a_s - 1$ cards of color c (i.e., $DS[\cdot] \geq a_s - 1$).

Recall that our aim is to play as many cards of color c as possible. Thus, if we want to play t additional cards after (a_s, c) but are not allowed to scan more cards of the input, those t cards must have been previously stored. Thus, we are interested in the largest value of t so that cards of values $a_s + 1, \dots, a_s + t$ of color c can be stored in hand while making sure that all constraints are satisfied.

Note that t can be as small as zero (i.e., when we do not store additional cards) and at most h (since we cannot store more than h cards at any instant of time), giving $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] = \mathcal{P}$ as claimed.

Thus, when $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}]$ is a positive number, we have equality from the fact that we query feasible moves (thus $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}] \geq \max\{\mathcal{P}, \mathcal{S}, \mathcal{D}\}$) and exhaustiveness (since we try all options, the largest of them must satisfy $\max\{\mathcal{P}, \mathcal{S}, \mathcal{D}\} \geq DP[s, \mathcal{H}, p_1, \dots, p_{c-1}]$). Similarly, if an entry of DP is unbounded, its associated values \mathcal{P} , \mathcal{S} and \mathcal{D} will also be unbounded (since a bounded number would be a witness of a winning play sequence). ◀

► **Theorem 6.** *We can solve a HANABI problem instance in $O(Nhc^h n^{h+c-1})$ time using $O(c^h n^{h+c-1})$ space.*

Proof. By definition, there is a solution to the HANABI problem instance if and only if its associate table satisfies $DP[N, \emptyset, n \dots, n] = n$. Each entry of the table is solved by querying entries that have a smaller value in the first parameter, so we can compute the whole table in increasing order.

Recall that, entries of the table for which the associated set \mathcal{H} has more than h elements the answer is trivially $-\infty$ (since we cannot store that many cards). Thus, table DP will have $N \times \sum_{i \leq h} \binom{nc}{i} \times n^{c-1} \in O(Nc^h n^{h+c-1})$ nontrivial entries. Note that when computing them, each entry of the table queries for entries whose value of s is one smaller. Thus, after the values of some value of s have been computed, the smaller ones need not be stored. This way we never need to store more than $O(c^h n^{h+c-1})$ entries at the same time.

We now bound the time needed to compute a single entry of the table (say, $DP[s, \mathcal{H}, p_1, \dots, p_{c-1}]$). First notice that we can compute \mathcal{P} and \mathcal{S} with a constant number of queries to the table. Each query makes at most one insertion or removal into \mathcal{H} . Such insertions can be handled in constant time (see the proof of Lemma 1), thus overall they are computed in constant time.

In order to compute \mathcal{P} , we may have to do $O(h)$ queries onto the DP table and insert $O(h)$ elements into \mathcal{H} . Since each of these operations take constant time, we need $O(h)$ time to compute a single entry (and $O(Nc^h n^{h+c-1}h)$ for the whole table as claimed). ◀

► **Remark.** In principle, the DP table only returns whether or not the instance is feasible. We note that, we can also find a winning play sequence with the usual backtracking techniques.

5 NP-Hardness (Multiple Colors, Multiple Appearances)

In this section we prove hardness of the general HANABI problem. As mentioned in the introduction, the problem is NP-complete even if h and r are small constants.

► **Theorem 7.** *The HANABI problem is NP-complete for any $r \geq 2$ and $h \geq 2$.*

We prove the statement for $r = 2$, $h = 2$ and then show how to generalize it for larger values of r and h . Our reduction is from 3-SAT. Given a 3-SAT problem instance with v variables x_1, \dots, x_v and m clauses W_1, \dots, W_m , we construct a HANABI sequence σ with $2v + 1$ colors, $n = 6m + 2$, $r = 2$, $h = 2$ (and thus $N \leq 2(2v + 1)(6m + 2)$).

Before discussing the proof, we provide a birds-eye view of the reduction. The generated sequence will have a variable gadget V_i for each variable x_i and a clause gadget C_j for each clause W_j .

In the first phase of the game, the *variable assigning phase*, the player scans through the variable gadgets $V_i, i \leq v$. The variable gadget V_i associated to the i -th variable will have cards of colors $2i - 1$ and $2i$. After we have scanned through gadget V_i , the best we can do is play at most 5 cards of one color, and 1 from the other one. We assign a truth value to a variable depending on which of the two colors we played five cards. By repeating this in all variable gadgets we obtain a truth assignment.

In the next phase, the *clause satisfaction phase*, the player scans through the clause gadgets $C_j, j \leq m$. The clause gadget C_j corresponding to clause W_j is constructed in a way that when we scan through it we can play five additional cards (of all colors) if and only if the truth assignment satisfies the clause W_j . Thus, only when all clauses are satisfied the HANABI instance will have a solution.

As will be shown afterwards, it will be useful to temporarily reduce the amount of cards that can be stored in hand (or even make it zero). This can be enforced with an additional dummy color $2v + 1$. Indeed, assume that when scanning a portion λ of the input we want to make sure that hand size is one (for simplicity, we also assume that no card of the dummy color appears in the whole sequence). Then, it suffices to add cards $(2, 2v + 1)$ and $(1, 2v + 1)$ before and after λ , respectively. Since card $(2, 2v + 1)$ appears exactly once, then its unique appearance must be stored until card $(1, 2v + 1)$ is found. Thus, only one additional card can be stored while scanning λ . We call this gadget the *hand reduction gadget*.

Similarly, we can enforce independence between two gadgets by adding cards $(4, 2v + 1)$, $(5, 2v + 1)$, and $(3, 2v + 1)$ between them. If cards $(4, 2v + 1)$ and $(5, 2v + 1)$ appear exactly once in the sequence, they must be stored until the card $(3, 2v + 1)$ is scanned (at which point all three cards can be played). Since our hand size is two, this essentially makes sure that no card can be stored between gadgets. Note that this trick can be done arbitrarily many times provided that each time we use higher numbers of the dummy color (and each card appears only once in the whole sequence). We call this operation the *hand dump gadget*. Note that the gadgets are used to simulate a reduction in the hand size, but h remains constant. The temporary reduction is created by forcing some cards to be stored during a portion of the play sequence.

Let us first consider the variable assigning phase. We first describe the variable gadget. For any $i \leq v$, variable gadget V_i is defined as the sequence $V_i = 2, \bar{2}, 1, 3, 4, 5, \bar{1}, \bar{3}, \bar{4}, \bar{5}$, where overlined values are cards of color $2i$, whereas the other cards have color $2i - 1$. The first part of the HANABI problem instance σ simply consists of the concatenation of all gadgets V_1, \dots, V_v , adding card $(2, 2v + 1)$ in the very beginning and card $(1, 2v + 1)$ in the very end of the sequence, as to form a hand reduction gadget (see Figure 1). We call this sequence σ_1 .

► **Lemma 8.** *There is no valid play sequence of σ_1 that can play cards of value 2 of colors $2i - 1$, $2i$ and the dummy color $2v + 1$. This statement holds for all $i \leq v$.*

Proof. Assume, for the sake of contradiction, that there exists some $i \leq v$ and a sequence of plays for which we can play the three cards. In order to play card $(2, 2v + 1)$ we need to store it in the very beginning of the game enforcing the hand reduction gadget for the duration of the variable assigning phase, thus temporarily reducing the hand-size to one.

2	2 2 1 3 4 5 1 3 4 5	2 2 1 3 4 5 1 3 4 5	2 2 1 3 4 5 1 3 4 5	1
d	1 2 1 1 1 1 2 2 2 2	3 4 3 3 3 3 4 4 4 4	5 6 5 5 5 5 6 6 6 6	d

■ **Figure 1** Sequence σ_1 for a SAT instance with three variables. The upper row represents the numbers of the cards whereas the lower one represents the color of each card. Note that the dummy cards to reduce hand size are also added (color “d” stands for dummy color).

Further notice that each card appears exactly once in σ_1 (that is, the multiplicity of this part is equal to 1), and that the cards of color $2i - 1$ and $2i$ only appear in gadget V_i . More importantly, the value 2 in both colors appears before the value 1 in the respective colors. In particular, both must be stored before they are played. However, this is impossible, since we have decreased the hand size through the hand reduction gadget. ◀

From now on, for simplicity in the description, we only consider play sequences that play all the cards in the dummy color (recall that these cards appear exactly once. If any of them is not played the resulting sequence of moves cannot be completed). Similarly, we assume cards are played as soon as possible. In particular, if the card that is currently being scanned is playable, then it will be immediately played. We can make this assumption because holding it in hand is never beneficial. We call these two conditions the *smart play* assumption.

Thus, the best we can do after scanning through all variable gadgets is to play five cards of either color $2i - 1$ or $2i$ (and only one card of the other color). This choice is independent for all $i \leq v$, hence we associate a truth assignment to a play sequence as follows: we say that variable x_i is set to *true* if, after σ_1 has been scanned, the card $(5, 2i - 1)$ has been played, false if $(5, 2i)$ has been played. For well-definement purposes, if neither $(5, 2i - 1)$ or $(5, 2i)$ we simply consider the variable as *unassigned* (and say that an unassigned variable never satisfies a clause). This definition is just used for definement purposes since, as we will see later, no variable will be unassigned in a play sequence that plays all cards.

Let us now move on to the clause satisfaction phase by first describing the clause gadget C_j for clause W_j . We associate three colors to a clause. Specifically, we associate W_j with color $2i - 1$ if x_i appears positive in W_j . If x_i appears in negated form, we associate W_j with color $2i$ instead. Since each clause contains three literals, it will be associated to three distinct colors.

Let $o_j = 5(j - 1)$. Intuitively speaking, o_j indicates how many cards of each color can be played (we call this the *offset*). Our invariant is that for all $i \leq v$ and $j \leq m$, before scanning through the clause gadget associated to W_j , there is a play sequence that plays up to $o_j + 1$ cards in color $2i - 1$ and $o_j + 5$ in color $2i$ (or the reverse) and no play sequence can exceed those values in any color. Observe that the invariant is satisfied for $j = 1$ by Lemma 8.

The clause gadget C_j is defined as follows: we first add the sequence $o_j + 6, o_j + 7, o_j + 8, o_j + 9, o_j + 10$ for the three colors associated to W_j . Then we append the sequence $o_j + 5, o_j + 6, o_j + 7, o_j + 8, o_j + 9, o_j + 10, o_j + 2, o_j + 3, o_j + 4$ in all other colors (except the dummy color). After this we add three cards of the dummy color forming a hand dump gadget. Finally, we add the sequence $o_j + 3, \overline{o_j + 3}, \overline{\overline{o_j + 3}}, o_j + 2, o_j + 4, o_j + 5, o_j + 6, \overline{o_j + 2}, \overline{\overline{o_j + 4}}, \overline{\overline{\overline{o_j + 5}}}, \overline{\overline{\overline{\overline{o_j + 6}}}}$ in the three colors associated to W_j (as before, the single and double overline on the numbers is used to distinguish between the three colors). See Figure 2.

Let σ_2 be the result of concatenating all clause gadgets in order, where before each C_j we add three cards of the dummy color forming a hand dump gadget so as to make sure that no card from one gadget can be saved to the next one (see Figure 3). Further let $\sigma' = \sigma_1 \circ \sigma_2$.

We must show that, when scanning a clause gadget that is satisfied, we can play five cards of all colors. We start by showing that this is possible for the easy colors (i.e., colors for which we played five cards in σ_1 or those that are not associated to W_j).

	C_1							
4 5 3	6 7 8 9 10	5 6 7 8 9 10 2 3 4	7 8 6	3 3 3	2 4 5 6	2 4 5 6	2 4 5 6	
d d d	1,4, and 5	2, 3, and 6	d d d	1 4 5	1 1 1 1	4 4 4 4	5 5 5 5	
	C_2							
10 11 9	11 12 13 14 15	10 11 12 13 14 15 7 8 9	13 14 12	8 8 8	7 9 10 11	7 9 10 11	7 9 10 11	
d d d	1,3, and 6	2, 4, and 5	d d d	1 3 6	1 1 1 1	3 3 3 3	6 6 6 6	

■ **Figure 2** Sequence σ_2 for a SAT instance with three variables x_1, x_2, x_3 and two clauses $W_1 = (x_1 \vee \neg x_2 \vee x_3), W_2 = (x_1 \vee x_2 \vee \neg x_3)$. Colors 1, 4, 5 are associated to W_1 and colors 1, 3, 6 are associated to W_2 . The upper row represents the numbers of the cards whereas the lower one the color of each card. Note that the dummy cards to obtain independence between/inside gadgets are also added (color “d” stands for dummy color).

$$\sigma_1 \quad 4 \quad 5 \quad 3 \quad C_1 \quad 10 \quad 11 \quad 9 \quad C_2 \quad \dots \quad 6m - 2 \quad 6m - 1 \quad 6m - 3 \quad C_m$$

■ **Figure 3** Overall picture of the reduction. All cards depicted have dummy color (and are only used to obtain independence between gadgets).

► **Lemma 9.** *Let $k \leq 2v$ be a color for which before processing clause gadget C_j we have played up to $o_j + 5$ cards of color k (for some $j \leq m$) or a color not associated to W_j for which we have played up to $o_j + 1$ cards of color k . Then, we can play up to five more cards of color k when processing the clause gadget C_j . Moreover no play sequence can play more than five cards of that color.*

Proof. Recall that there are hand dumps between different gadgets. Thus, any cards that is played while processing C_j must appear in C_j .

The case in which we played up to the value $o_j + 5$ of a color is easy, since C_j contains the sequence $o_j + 6, o_j + 7, o_j + 8, o_j + 9, o_j + 10$ in consecutive fashion in all colors. Thus, the five cards can be played without having to store anything in hand. Also note that a sixth card cannot be played since $o_j + 11$ is not present in any color in C_j .

The case in which we played up to $o_j + 1$ of a color not associated to W_j is similar. In this case, the cards of color k appear in the following order: $o_j + 5, o_j + 6, o_j + 7, o_j + 8, o_j + 9, o_j + 10, o_j + 2, o_j + 3, o_j + 4$. It is straightforward to verify that if we are only allowed to store two cards we can play at most five cards. ◀

The remaining case is that a color k is associated to W_j and only $o_j + 1$ cards have been played. Recall that, by the way in which we associated variable assignments and play sequences, this corresponds to the case that the assignment of variable $x_{\lceil k/2 \rceil}$ does not satisfy the clause W_j . We now show that five cards of color k will be playable if and only if at least one of the other two variables satisfies the clause.

► **Lemma 10.** *Let C_j be the clause gadget associated to W_j (for some $j \leq m$). We can play five cards in each of the three colors associated to W_j if and only if we have played card of value $o_j + 2$ in at least one of the three associated colors before W_j is processed. Moreover, we can never play more than five cards in the three colors associated to W_j .*

Proof. The proof is similar to the Lemma 9. By construction of our gadget, we first find the sequence $o_j + 6, o_j + 7, o_j + 8, o_j + 9, o_j + 10$ in all three colors associated to W_j . These cards are unplayable if we have only played up to $o_j + 1$, so the best we can do is to store them. However, before we find the smaller numbers of the same color, there is a hand dump

gadget. Thus none of these cards will be playable for colors in which, before C_j is processed, we have only played cards of value at most $o_j + 1$.

The only other cards of the associated color that are present in the gadget form the sequence $\overline{o_j + 3}, \overline{o_j + 3}, \overline{o_j + 3}, o_j + 2, o_j + 4, o_j + 5, o_j + 6, \overline{o_j + 2}, \overline{o_j + 4}, \overline{o_j + 5}, \overline{o_j + 6}, \overline{o_j + 2}, \overline{o_j + 4}, \overline{o_j + 5}, \overline{o_j + 6}$. Again, because of the hand dump gadget before and after this sequence, no other cards can be played.

Consider the case in which we have played only up to $o_j + 1$ of the three colors before processing C_j (or equivalently, the variable assignment does not satisfy clause W_j). The first three cards we find have the number $o_j + 3$ in the three colors. None is currently playable, thus ideally we would like to store them. Due to the limitations on our hand size, we can only store two of the three cards. In particular, from the color whose card $o_j + 3$ was discarded we will only be able to play one card. Note that for this situation to happen it is crucial that in none of the three colors we have played up to card $o_j + 5$. When this condition is not satisfied in at least one color we can make sure that five cards in all colors are played. ◀

From the above results we know that by the time we scan through σ' we can play at least up to value $o_m + 6$ (in half of the colors we can play up to value $o_m + 10$) if and only if the variable assignment created during the variable assignment phase satisfied all clauses. For the dummy color, we used one hand size reduction gadget and two hand dump gadgets per clause, thus $6m + 2$ cards will have been played. We pad σ' with values $o_m + 6$ to $6m + 2$ in increasing order in all colors (except the dummy color). Let σ be the resulting sequence.

► **Theorem 11.** *There is a valid solution of HANABI for σ and $h = 2$ if and only if the associated problem instance of 3-SAT is satisfiable.*

Proof. If the associated problem instance of 3-SAT is satisfiable, there exists a truth assignment satisfying all clauses, by Lemmas 8, 9 and 10, we can play all colors up to the card $6m + 2$ from σ . If the associated problem instance of 3-SAT is not satisfiable, for any truth assignment, there exists one or more clauses that are not satisfied. Let j be the index of the first clause that is not satisfied by the truth assignment. By Lemma 10, we will not be able to play card $o_j + 3$ in one of the three colors associated to C_j . Since the smallest number of the next gadgets is $o_j + 7$, no more cards can be played in that color. In particular, there cannot be a solution for this HANABI problem instance. ◀

The above reduction can be easily constructed in polynomial time. Further note that the reduction works for $r = 2$ and $h = 2$. If we want to have exactly r copies, it suffices to place next to the first appearance of each card $r - 1$ or $r - 2$ cards identical to it, so that we end up having a number of consecutive copies of each card of which only one will be useful. On the other hand, if $h > 2$ we can use a hand reduction gadget to reduce the hand size to exactly 2 for the interval in which σ is processed. This completes the proof of Theorem 7.

6 Conclusions

In this paper, we studied the complexity of a single player version of Hanabi, but the hardness also extends for the case in which we have p players. With more than one player the *give a hint* action allows a player to pass (i.e., neither draw nor discard). Thus, if we have enough hints (say, at least pN), the game with p players each with a hand size of h is equivalent to a game with a single player and hand size ph .

Even though our model is a bit far from the original game, it uncovers the importance of hand management. This was an aspect of the game that had been overlooked in previous

studies of the game. We hope that the ideas presented in this paper will help towards the design of strategies that are useful for this interesting game and that perform better than the currently existing ones.

Several questions regarding the complexity of the game remain unanswered. For example, is the game still NP-complete if we bound the number of colors instead of the hand-size? Also, can we obtain linear-time algorithms for the case where h and c are small constants (for example when $h = 1$ and/or $c = 2$)? Furthermore, as the problem is very rich in parameters, it would be fruitful to study it from a parameterized complexity point of view.

References

- 1 Antoine Bauza. Hanabi. <http://www.antoinebauza.fr/?tag=hanabi>.
- 2 BoardGameGeek. <https://boardgamegeek.com/boardgame/98778/hanabi>.
- 3 Alex Churchill. Magic: The gathering is Turing complete. Unpublished manuscript available at <http://www.toothycat.net/~hologram/Turing/index.html>.
- 4 Christopher Cox, Jessica De Silva, Philip Deorsey, Franklin H. J. Kenter, Troy Retter, and Josh Tobin. How to make the perfect fireworks display: Two strategies for Hanabi. *Mathematics Magazine*, 88(5):323–336, 2015.
- 5 Erik D. Demaine. Personal communication.
- 6 Erik D. Demaine. Playing games with algorithms: Algorithmic combinatorial game theory. *CoRR*, cs.CC/0106019v2, 2008.
- 7 Erik D. Demaine, Martin L. Demaine, Nicholas J. A. Harvey, Ryuhei Uehara, Takeaki Uno, and Yushi Uno. UNO is hard, even for a single player. *Theor. Comp. Sci.*, 521:51–61, 2014.
- 8 Spiel des Jahres award. <http://www.spieldesjahres.de/en/hanabi>.
- 9 Aviezri S. Fraenkel and David Lichtenstein. Computing a perfect strategy for $n \times n$ chess requires time exponential in n . *J. Comb. Theory, Ser. A*, 31(2):199–214, 1981.
- 10 Martin Gardner. *Mathematical Games: The Entire Collection of His Scientific American Columns*. The Mathematical Association of America, 2005.
- 11 Robert Hearn and Erik D. Demaine. *Games, Puzzles, and Computation*. A. K. Peters, 2009.
- 12 Michael Lampis and Valia Mitsou. The computational complexity of the game of set and its theoretical applications. In *11th Latin American Symposium*, pages 24–34. Springer, 2014.
- 13 Kenichiro Nakai and Yasuhiko Takenaga. NP-completeness of pandemic. *JIP*, 20(3):723–726, 2012.
- 14 Hirotaka Osawa. Solving Hanabi: Estimating hands by opponent’s actions in cooperative game with incomplete information. In *AAAI workshop: Computer Poker and Imperfect Information*, pages 37–43, 2015.

A The Rules of Hanabi

In this appendix we introduce the official rules of Hanabi [1].

Game Material

50 fireworks cards in five colors (red, yellow, green, blue, white):

- 10 cards per color with the values 1, 1, 1, 2, 2, 3, 3, 4, 4, 5,
- 5 colorful fireworks cards with values of 1, 2, 3, 4, 5,
- 8 Clock (Note) tokens (+ 1 spare),
- 3 Storm (Fuse) tokens.

Aim of the Game

Hanabi is a cooperative game, meaning all players play together as a team. The players have to play the fireworks cards sorted by colors and numbers. However, they cannot see their own hand cards, and so everyone needs the advice of his fellow players. The more cards the players play correctly, the more points they receive when the game ends.

The Game

The oldest player is appointed first player and sets the tokens in the play area. The eight Clock tokens are placed white-side-up. The three Storm tokens are placed lightning-side-down.

Now the fireworks cards are shuffled. Depending on the number of players involved, each player receives the following hand:

- With 2 or 3 players: 5 cards in hand,
- With 4 or 5 players: 4 cards in hand.

Important: For the basic game, the colorful fireworks cards and the spare Clock token(s) are not needed. They only come in to use for the advanced game.

Important: Unlike other card games, players may not see their own hand! The players take their hand cards so that the back is facing the player. The fronts can only be seen by the other players. The remaining cards are placed face down in the draw pile in the middle of the table. The first player starts.

Game Play

Play proceeds clockwise. On a player's turn, he must perform exactly one of the following:

- A. Give a hint or
- B. Discard a card or
- C. Play a card.

The player has to choose an action. A player may not pass!

Important: Players are not allowed to give hints or suggestions on other players' turns!

A. Give a hint

To give a hint one Clock token must be flipped from its white side to its black side. If there are no Clock tokens white-side-up then a player may not choose the Give a hint action. Now the player gives a teammate a hint. He has one of two options:

1. **Color Hint.** The player chooses a color and indicates to his/her teammate which of their hand cards match the chosen color by pointing at the cards. *Important:* The player must indicate all cards of that color in their teammate's hand! Example: "You have two yellow cards, here and here." Indicating that a player has no cards of a particular color is allowed! Example: "You have no blue cards."
2. **Value Hint.** The player chooses a number value and gives a teammate a hint in the exact same fashion as a Color Hint. Example: "You have a 5, here." Example: "You have no Twos."

B. Discard a card

To discard a card one Clock token must be flipped from its black side to its white side. If there are no Clock tokens black-side-up then a player may not choose the Discard a card

action. Now the player discards one card from their hand (without looking at the fronts of their hand cards) and discards it face-up in the discard pile near the draw deck. The player then draws another card into their hand in the same fashion as their original card hands, never looking at the front.

C. Play a card

By playing out cards the fireworks are created in the middle of the table. The player takes one card from his hand and places it face up in the middle of the table. Two things can happen:

1. **The card can be played correctly.** The player places the card face up so that it extends a current firework or starts a new firework.
2. **The card cannot be played correctly.** The gods are angry with this error and send a flash from the sky. The player turns a Storm tile lightning-side-up. The incorrect card is discarded to the discard pile near the draw deck.

In either case, the player then draws another card into their hand in the same fashion as their original card hands, never looking at the front.

The Fireworks

The fireworks will be in the middle of the table and are designed in five different colors. For each color an ascending series with numerical values from 1 to 5 is formed. A firework must start with the number 1 and each card played to a firework must increment the previously played card by one. A firework may not contain more than one card of each value.

Bonus

When a player completes a firework by correctly playing a 5 card then the players receive a bonus. One Clock token is turned from black side to white side up. If all tokens are already white-side-up then no bonus is received. Play then passes to the next player (clockwise).

Ending the Game

The game can end in three ways:

1. The third Storm token is turned lightning-side-up. The gods deliver their wrath in the form of a storm that puts an end to the fireworks. The game ends immediately, and the players earn zero points.
2. The players complete all five fireworks correctly. The game ends immediately, and the players celebrate their spectacular victory with the maximum score of 25 points.
3. If a player draws the last card from the draw deck, the game is almost over. Each player—including the player who drew the last card—gets one last turn. Note: Cards cannot be drawn in this last round.

Finally, the fireworks will be counted. For this, each firework earns the players a score equal to the highest value card in its color series. The quality of the fireworks display according to the rating scale of the “International Association of Pyrotechnics” is:

- 0–5: Oh dear! The crowd booed.
- 6–10: Poor! Hardly applauded.
- 11–15: OK! The viewers have seen better.
- 16–20: Good! The audience is pleased.
- 21–24: Very good! The audience is enthusiastic!
- 25: Legendary! The audience will never forget this show!

Important Notes and Tips

- Players may rearrange their hand cards and change their orientation to help themselves remember the information they received. Players may not ever look at the front of their own cards until they play them.
- The discard pile may always be searched for information.
- Hanabi is based on communication – and non-communication – between the Players. If one interprets the rules strictly then players may not, except for the announcements of the current player, talk to each other. Ultimately, each group should decide by its own measure what communication is permitted communication. Play so that you have fun!