## 1. Introduction

## 2. Literature Review Of Suitable Techniques

### 2.1 Incomplete Information And Its Effect On The Game

When participating in a game of Hanabi an agent is able to access every other player’s hands, but not its own. The agent has **incomplete information** when assessing what action will maximise total utility for the game.

Humans are able to alter their perception of their (inaccessible) hand and note which cards are seemingly worthless or of value based on the behaviour of others [Solving Hanabi]. An agent is limited in this respect.

An agent needs to acquire information from each action performed by every other agent in the game.

When a player gives a hint, it is important to recognise that this encounters a cost of one hint token. The restriction to interact with other players is further limited. If receiving a hint, the information gained should be considered carefully. A hint gives two pieces of information. Those cards that *are*, and those cards that are *not*.

A discard can be interpreted that the player does not have enough information to play confidently or enough information has been given to discard a card.

If Hanabi were played such that there was **complete information** then a player would only give a hint if they wanted to skip a turn [Solving Hanabi]. An agent would not be required to interpret other agent’s actions.

### 2 Rule Based Approach

When playing Hanabi, intuitively a number of rules become apparent when attempting to maximise the end score of the game.

When all fireworks are empty, any card with the value of “one” can be played

When all fireworks stacks are above a certain value, any cards below/equal to this value can be discarded

There is only a single five for each colour so they must not be discarded when in a player’s hand

2s, 3s and 4s all have duplicates – however if any one of these is discarded the remaining card must be treated appropriately.

These rules can be hard-code quite easily into a simple reflex agent. However, issues arise when introducing the concept of hints. Which hint will maximise the final payoff? How is an agent meant to interpret a hint? An agent is limited by the number of hint tokens available too.

The approach considered and implemented was that of a conservative/cautious agent which would follow a hierarchy of actions.

Play a card that is guaranteed to work

Give a hint to another players card which should maximise payout

Discard a card

Hints were prioritised by the amount of information that can be obtained from them. It was important that discarding was limited as it was a risk and could limit the maximum score obtained.

For every state, a utility was assigned to each card based on the information known. Therefore, the current agent was able to identify (based on the above rules) which of their cards and other players cards were important, and which were not.

### 3 Monte Carlo Tree Search

### 4 Dynamic Programming

## Rationale Of Selected Technique

## Implementation Description

## Validation

Solving Hanabi: Estimating Hands By Opponent’s Actions In Cooperative Game With Incomplete Information