Knowledge-Based Systems xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Automated playtesting in collectible card games using evolutionary algorithms: A case study in hearthstone

Pablo García-Sánchez[*,a], Alberto Tonda[b], Antonio M. Mora[c], Giovanni Squillero[d], Juan Julián Merelo[e]

[a] Dept. of Computer Engineering, University of Cádiz, Spain
[b] UMR 782 GMPA, INRA, Thiverval-Grignon, France
[c] Dept. Computer Sciences and Technology, Universidad Internacional de La Rioja, Spain
[d] Politecnico di Torino, Italy
[e] Dept. of Computer Architecture and Computer Technology, University of Granada, Spain

ARTICLE INFO

Keywords:
Genetic algorithm
Hearthstone
Collectible card games
Artificial intelligence

ABSTRACT

Collectible card games have been among the most popular and profitable products of the entertainment industry since the early days of *Magic: The Gathering*™ in the nineties. Digital versions have also appeared, with *HearthStone: Heroes of WarCraft*™ being one of the most popular. In Hearthstone, every player can play as a hero, from a set of nine, and build his/her deck before the game from a big pool of available cards, including both neutral and hero-specific cards. This kind of games offers several challenges for researchers in artificial intelligence since they involve hidden information, unpredictable behaviour, and a large and rugged search space. Besides, an important part of player engagement in such games is a periodical input of new cards in the system, which mainly opens the door to new strategies for the players. Playtesting is the method used to check the new card sets for possible design flaws, and it is usually performed manually or via exhaustive search; in the case of Hearthstone, such test plays must take into account the chosen hero, with its specific kind of cards. In this paper, we present a novel idea to improve and accelerate the playtesting process, systematically exploring the space of possible decks using an Evolutionary Algorithm (EA). This EA creates *HearthStone* decks which are then played by an AI versus established human-designed decks. Since the space of possible combinations that are play-tested is huge, search through the space of possible decks has been shortened via a new heuristic mutation operator, which is based on the behaviour of human players modifying their decks. Results show the viability of our method for exploring the space of possible decks and automating the play-testing phase of game design. The resulting decks, that have been examined for balancedness by an expert player, outperform human-made ones when played by the AI; the introduction of the new heuristic operator helps to improve the obtained solutions, and basing the study on the whole set of heroes shows its validity through the whole range of decks.

1. Introduction

Collectible card games (CCGs) are turn-based card games where players set up their decks in advance, carefully selecting cards in order to have the opportunity to exploit powerful combinations later, during an actual match. In this process, known as *deckbuilding*, players usually choose the cards from a large set (with hundreds or even thousands of possibilities), acquired in digital or physical sealed packs. Since every card has specific features, complex and rich gameplays usually emerge.

CCGs gained popularity in the 1990s, thanks to *Magic: The Gathering*™. In the following decades, the genre has been revitalized by the advent of digital versions (digital collectible card games, DCCGs), such as *Clash Royale*™ or *HearthStone: Heroes of WarCraft*™, both downloaded more than 50 million times [1].

Cards are clearly the main component of the game, and in order to increase players' engagement, available sets are updated regularly. On average, every two to six months a new set of cards is added to the game, and sometimes older ones are removed. Such *expansions* need to be thoroughly analyzed to avoid *breaking the game* [2]. However, testing these new cards is a difficult issue, because as the number of cards increases, also does the number of interactions or "combos", that can be potentially unbalanced and may affect players' enjoyment of the game.

Besides taking into account obvious constraints and checking for under- or over-powered cards, the end goal is the increase of

* Corresponding author.
E-mail addresses: pgarcia@atc.ugr.es, pablo.garciasanchez@uca.es (P. García-Sánchez).

https://doi.org/10.1016/j.knosys.2018.04.030
Received 8 December 2017; Received in revised form 22 March 2018; Accepted 24 April 2018

satisfaction, engagement, and ultimately *fun* for all players. However, the huge dimension of most games makes it very hard to fully test the impact on all facets of the game of new contents before including them in one of the released expansion packs. The situation calls for automatic testing tools to help designers to identify possible flaws or weaknesses, and to assess their impact in advance. This process is called *playtesting*, which is a way of searching the space of possible decks for superior ones.

Automatic deckbuilding methods may be useful to automatically test feasible and interesting decks, not only for developers, but also to players, as a decision-aid tool inside the game using their current collection as input. Moreover, they allow to add extra content, for example challenges where players have to defeat an Artificial Intelligence (AI) agent (or bot) with an original optimized deck.

In this paper, we propose a playtesting tool that can be used during the development of new content for a DCCG. More specifically, we focused on automatic deckbuilding in *HearthStone*, arguably the most successful DCCG nowadays. We propose to automatically build and test viable and competitive decks using an Evolutionary Algorithm (EA) [3].

Even though deckbuilding is a cornerstone of CCGs [4], there is not a lot of work in literature dealing with this specific issue. Indeed the possibility of optimizing deckbuilding in *HearthStone* was firstly explored by the authors in [5]. However that paper just presented a proof-of-concept study considering a limited experimental setup. In fact, we addressed as future work the possibility to evolve decks for different Heroes, and to perform a play-by-play analysis of the decks, to try and assess the generality of our approach.

Therefore, we have extended the search space of possible decks by including all available Heroes in the game. Since this extension boosts the size of the search space and thus the time of search, a new heuristic deck mutation operator has been added to the evolutionary algorithm. Also, as this method is intended to be a playtesting methodology, the best way to evaluate it is to use an expert to validate the decks obtained automatically by our algorithm, which is what we have done in this paper.

The rest of the paper is structured as follows. Section 2 introduces a few key concepts related to our subject. The state of the art and related works are commented in Section 3. The proposed method is described in Section 4, followed by the specific experimental setup in Section 5. Obtained results and findings about the evolutionary deckbuilding and playtesting process are exposed, followed by an analysis on the used cards and their impact on the matches in Sections 6 and 7, respectively. Finally, reached conclusions and an outline of future directions are presented.

2. Background

In this section, we summarize the terminology used, then the particular CCG we are working with, Hearthstone, and finally introduce the method we are applying, an Evolutionary Algorithm.

2.1. Collectible card games: concepts and terminology

CCGs became known to the wide public with the release of *Magic: The Gathering* in 1993. The culture that evolved around them developed specific terms, defined in the following.

2.1.1. Deckbuilding

The deck that a player deploys in a CCG must be carefully chosen, starting from a common pool of available cards. Building a deck is one of the most important parts of the experience: while playing the cards optimally is obviously important, players can only hope to draw a card that already is in their deck. Deckbuilding is a complex activity that requires understanding the current state of the game as well as evaluating other players; not surprisingly, it monopolizes the vast majority of articles and discussions among players on the Internet.

2.1.2. Metagame

All important activities associated with the experience, yet perceived by players as *peripheral* to the game itself, are cumulatively termed *metagame*. In the context of CCGs, metagame describes the *type* of decks that one is expected to find in a specific *ladder* (i.e. a competitive, ranked system), that is, “what everyone else is playing” [6].

2.1.3. Mana curve

In most CCGs each card has a *cost*, indicating the number of resources needed to play it, and usually called “casting cost”. The card cost is used for balance, as resources increase over time, and cheaper cards tend to be weaker but can be played early in the game, while expensive cards are potentially game-changers. Such cost is generally called “mana” after the term used for resources in *Magic: The Gathering*™. The *mana curve*[1] is a histogram plot representing a deck by counting the number of cards per each casting cost. From such mana curves, it’s easy to understand the reference archetype. In any case, the mana curve should usually be somehow balanced, because completely lacking cards with a certain casting cost might leave the player unable to be effective in the early, mid or late game.

2.1.4. Deck archetypes

Since CCGs have a huge card pool, players can create decks with many different behaviours. A *deck archetype* is a category of deck formed by a specific subset of cards that allows a particular style of play. While each CCG features its own exclusive archetypes such as *Suicide Black* in Magic: The Gathering and *Malylock Warlock* in HearthStone [7], there are a few broader typologies that all decks can be roughly reduced to.

- *Aggro*, short for “aggression”, is a deck driven by a relatively simple strategy: the player attempts to finish the game in its early stages, quickly consuming lots of resources to inflict the maximum possible damage to the opponent. Typically, if players with *Aggro decks* cannot end the game fast enough, they will eventually lose in the mid or late game. This kind of decks will typically have a considerable number of low-cost cards, with a mana curve shifted to the left.
- *Combo* is a deck where the main objective of the players is to survive until they manage to draw all the necessary pieces of a combination. Combos usually include two or more synergistic cards that allow the player to unleash a considerable amount of damage (ideally lethal) over the span of a single turn, securing the game. Players with these deck archetypes may lose if the opponent is able to produce a significant attack before all the pieces of the combination are gathered, or if the opponent is prepared to somehow counter it.
- *Control* is a deck chosen to keep the opponent in check, neutralizing early-game threats to prolong the match until the late game, where they can finish off using high-cost, high-value cards. Players with Control decks risk losing if they cannot find good answers for the cheap, effective threats of Aggro decks, or if they fail to counter the lethal combinations of Combo decks. This kind of decks will typically feature a mana curve shifted to the right.

2.2. Hearthstone

HearthStone: Heroes of WarCraft is an online DCCG launched in 2013 by Blizzard Entertainment. Players build a deck of 30 cards from a huge card pool that can be expanded buying random packs, or converting owned cards to in-game currency to buy new ones. To win, players need to reduce the health of the other human opponent (or Hero) from 30 to 0, using the two types of cards available: *spells*, that affect the battleground and are then discarded, and *minions*, that stay in play and can

[1] http://www.hearthstone.gamepedia.com/Mana_curve.