Evolutionary Deckbuilding in HearthStone

Pablo García-Sánchez*, Alberto Tonda†, Giovanni Squillero‡ Antonio Mora§ and J.J. Merelo*
*Dept. of Computer Architecture and Computer Technology. University of Granada, Granada, Spain.
Email: pablogarcia@ugr.es,jjmerelo@geneura.ugr.es
†UMR 782 GMPA, INRA. Thiverval-Grignon, France. Email: alberto.tonda@grignon.inra.fr
‡Politecnico di Torino, Torino, Italy. Email: giovanni.squillero@polito.it
§Department of Software Engineering. University of Granada, Granada, Spain. Email: amorag@ugr.es

***Abstract*—One of the most notable features of collectible card games is deckbuilding, that is, defining a personalized deck before the real game. Deckbuilding is a challenge that involves a big and rugged search space, with different and unpredictable behaviour after simple card changes and even hidden information. In this paper, we explore the possibility of automated deckbuilding: a genetic algorithm is applied to the task, with the evaluation delegated to a game simulator that tests every potential deck against a varied and representative range of human-made decks. In these preliminary experiments, the approach has proven able to create quite effective decks, a promising result that proves that, even in this challenging environment, evolutionary algorithms can find good solutions.**

I. INTRODUCTION

Collectible Card Games (CCGs) have been part of the mainstream gaming culture since the 90s, when *Magic: the Gathering*™ first became popular. Such games have suffered a recent growth thanks to *HearthStone: Heroes of WarCraft*™ [1], a game that, thanks to a very effective Free-To-Play model, reached a record of 40 million registered accounts in 2016 [2].

The common objective in a wide set of turn-based card games is to beat the opponent by using on him different types of cards (such as *spells* or *minions*). In CCGs every player is asked to construct a specific deck before the actual match. As the cards include specific rules that deeply affect the interaction between players, building a deck promotes an interesting and rich game play.

These kind of games are an interesting test bed in AI research, as players need to deal with hidden information and randomness, with the combination of states, rules and cards that may imply complex or unpredicted reactions, such as *combos*, combination of card so explosive that probably have not been anticipated even by the creators of the game.

Several authors have applied diverse computational intelligence methods to a variety of problems related to this field. For example, Cowling et al. compared different Monte Carlo tree search methods to deal with the imperfect information of the *Magic: the Gathering* game, obtaining better results that an expert rule-based agent [3]. CCGs have also been used as an example of the application of a framework to automatically detect design issues of new games [4].

However, previous works dealt with the AI aspects of the game in terms of automatic playing and behavior design. Being the construction of the deck a very important part of this kind of games, where players may spend hundred of dollars in buying cards, it is quite surprising the lack of works in the literature proposing computational methods for automatic deck generation and analysis of their effectiveness.

These techniques can also be interesting for CCGs manufacturers or developers, as adding new sets of cards may unbalance the game. Balancing games is a complex task, as new cards can affect previous rules, as well as all the possible combinations of card effects [5].

This paper proposes a methodology to automatically create decks for CCGs using an *Evolutionary Algorithm* (EA), an optimization technique loosely inspired by natural evolution. In EAs, potential solutions are encoded in a suitable format, and an objective function called *fitness* is automatically optimized [6]. EAs have already been extensively used in AI generation for videogames [7], [8], [9].

EAs are commonly used in combinatorial problems, as they commonly produce very effective combinations of elements, yet quite different from what a human expert would do. In the current framework, it makes possible to obtain competitive decks from scratch, i.e. without adding human knowledge. The proposed approach encode the candidate decks as vectors. The fitness function used to drive the evolutionary process is based on a series of actual matches against properly selected opponents. The resulting statistics are then analyzed and parsed to obtain a numerical metric.

The rest of the paper is structured as follows: after some background in CCGs and Evolutionary Algorithms, the proposed approach is described in Section III. After the experimental setup (Section IV), the results are discussed in Section V. Finally, the Conclusions and future lines or work are addressed.

II. BACKGROUND

In this section we present some preliminary concepts that will help the reader to better understand the work.

A. Collectible card games

The field of CCGs, that exploded with *Magic: the Gathering* in 1993, over time developed a specific terminology. There is a set of shared concepts in this field: *deckbuilding* (decks can be prepared by the player, following certain rules); *competitive play* (the objective is to defeat the opponent); *card costs* (players have limited resources available every turn, and playing each card consumes some of these resources). While there are many variations on CCGs, ranging from cooperative play to games with no card costs, the popular ones — namely *Magic*, *HearthStone*, and *Yu-Gi-Oh!* — include all these concepts.

1) Deck types: Some of these terms are referred to the type of decks:

- *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 a player with an Aggro deck cannot end the game fast enough, he will eventually lose in the mid or late game.
- *Combo* is a deck where player's main objective is to survive until he manages 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 decks 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.

2) Metagame: The term *metagame* is used to describe conceptually difficult activities associated with game play, perceived by players as 'peripheral' to the game itself, but important to the whole game experience. Concretely, in the context of CCGs, *metagame* indicates the types of decks that a player entering a specific competitive event (or ladder) is expected to find, in largest numbers. Or in other words, 'what everyone else is playing' [10].

B. HearthStone

Launched in 2013, *Hearthstone: Heroes of Warcraft* is an online CCG, developed by Blizzard Entertainment. Players compete against each other, trying to reduce the enemy health from 30 to 0 points, building their decks from a pool of cards that is constantly increasing, when either expansion packs or single-player adventures that reward the player with collectible cards upon completion are published. Currently, there are 743 unique collectible cards in the game, with more planned to be added in the future through additional content. Every card has an associated probability to be obtained when the player buys an envelope, being related with its power: common, rare, epic and legendary.

Cards in HearthStone fall into two main categories: *spells* and *minions*. Spells are played, create an effect on the battlefield, and then are discarded. Minions, on the other hand, stay in play, and can be used to attack the enemy Hero or other minions. Each card has a cost, that is paid when the card is played, using *crystals* (also called *mana*), a resource that grows every turn. On their first turn, players can use a total of one crystal, on the second turn they are allotted two crystals, and so on, to a maximum of ten crystals for turns ten and later. The cost is used for balance: powerful cards have a higher cost, cheaper cards are not as effective. A deck has to feature cards of all costs, in order to be able to play effectively in the early, mid and late game.

In *HearthStone*, deckbuilding is further constrained by the *Hero* the player chooses: each Hero features a special power that can be activated during the game, and exclusive cards that can only be used for that Hero. There are currently 9 different types of Hero: Druid, Hunter, Mage, Paladin, Priest, Rogue, Shaman, Warlock and Warrior. Even if it is theoretically possible to build an Aggro/Combo/Control deck using each Hero, in practice most Heroes are more suited to a single deck type. For example, the Priest's ability and exclusive cards make it a very powerful choice for Control (with several variations of Priest Control decks), but a poor one for Aggro.

C. Evolutionary algorithms

Evolutionary algorithms (EAs) [6], [11] are bio-inspired meta-heuristics that can be effectively used to find nearly optimal solutions for optimization problems. Usually an EA starts by generating a set of random solutions, called *population*, following a user-defined description. Then, it evaluates each candidate solution, called *individual*, assigning it a *fitness* value, that describes how good the individual is, with regards to the target problem. New solutions are then generated by the application of *operators* that either mutate a single existing solutions or recombine different existing solutions. After each iteration, called *generation*, the least fit individuals are removed, and the process continues until a user-defined stop condition is met.

What makes EAs particularly interesting is their ability to manipulate complex structures such as binary trees or graphs [12]; and their relying only upon the fitness values, that can be provided by black-box evaluation, with no need of assumptions of regularity or stochasticity of the search space. For all these reasons, EAs have been already successfully employed in game design, for example evolving the parameters of an agent that play RTS such as *Planet Wars* [7] or generating the strategy of a *StarCraft*™ bot [13], and even the automatic creation of card games [14].

III. PROPOSED APPROACH

In this work, we propose to use an EA to optimize the deck for a specific metagame. The EA initially generates random decks, and then mutates and combines the most promising ones to generate new solutions. Candidate decks are evaluated using an AI capable of playing *HearthStone*, against a set of representative human-designed decks that define the target metagame. Their fitness is tied to the total number of victories obtained. We will go into details of the different aspects of the approach next.

A. Candidate solutions

Solutions in our problem are decks: following *HearthStone*'s rules, a deck has to be composed of exactly 30 cards, with no more than two copies of each, or exactly one copy in the case of Legendary cards. Decks can include both Neutral cards and those reserved to a single specific Hero.

B. Fitness function

As the evolutionary algorithm can freely manipulate decks, swapping any card for any other, crossing two decks and so on, it is possible that it will obtain decks that violate the rules of the game: for example, by having more than 2 copies of the same card, or more than 1 copy of a Legendary one. Also, while the total number of victories obtained is important, at the same time we desire a deck with a fair chance to win against all decks in the metagame, and not one that mercilessly slaughters specific opponents and loses badly against other ones. Also, and due to the stochastic nature of the game, a single execution of a game against a deck would not be statistically significant [15], so for each opponent at least 15 games should be played. For these reasons, the fitness function is divided into three parts, evaluated following a lexicographical order:

1) **Correctness**: this metric takes into account the number of errors in the decklist (repeated cards). decks that have this fitness value bigger than 0 are not evaluated further, and all their remaining fitness values are set to the lowest possible amount. This fitness value is to be minimized.
2) **Victories**: straightforwardly, this is the total number of victories obtained by the decklist played 16 times against each of the decks in the target metagame. This fitness value is to be maximized.
3) **Standard deviation**: this value is computed by evaluating the number of victories obtained against each opponent, and computing the standard deviation with regards to the number of victories against other opponents. If the deck obtains the same number of victories against all opponents, its standard deviation will be optimal. This fitness value is to be minimized.

This type of lexicographical fitness, using different parry opponents has been successfully used in previous works [13], [7].

IV. EXPERIMENTAL EVALUATION

This section describes the algorithm used and the decisions taken into account to model the fitness function.

TABLE I
PARAMETERS USED BY THE EA. THE ACTIVATION PROBABILITIES OF THE OPERATORS ARE SELF-ADAPTED. FOR MORE INFORMATION ON THE PARAMETERS, SEE [16] OR VISIT HTTPS://SOURCEFORGE.NET/P/UGP3/WIKI/HOME/.

Parameter	Meaning	Value
μ	Population size	10
λ	Operators applied	10
α	Self-adapting inertia	0.9
σ	Initial mutation strength	0.9
τ	Size of the tournament selection	[2-4]
G	Number of generations	50
R	Replacement mechanism	Generational
e	Number of parry decks	8
t	Number of games per parry deck	16
Operators used	`singleParameterAlterationMutation` `onePointCrossover` `twoPointCrossover`	

A. Evolutionary algorithm

The EA used in the experience is μGP, a general-purpose evolutionary framework [16], designed to easily implement different optimization problems out-of-the-box, thanks to its flexible definition of individual structure and external evaluator. The project is available on SourceForge[1]. During all the experiments, μGP has been configured with the parameters reported in Table I. The evolutionary operators collectively allow the EA to replace a card with any other card and cross over two decks.

B. MetaStone

MetaStone is an open-source HearthStone simulator[2]. It allows the manual creation of decks using the cards available in HearthStone and simulate games between decks, obtaining several statistics, such as turns taken or the damage done. Different heuristics can be selected for the AI engine, based on a score given to the actions that are evaluated in each turn, taking into account a combination of weights of the type of minions/spells used.

- Play Random: each turn the actions (moves) to play are selected randomly.
- Greedy Optimize Move: in each turn the AI selects each move ordered by score.
- Greedy Optimize Turn: in each turn the AI selects the combination of all possible moves with the higher score.
- Flat MonteCarlo Tree: during a certain number of iterations the AI simulates random moves until possible ends of the match to calculate the score.

C. Opponents decks

For the experimental evaluation, we consider the metagame of Season 18 of *HearthStone* competitive play, featuring the base set, the adventures *Curse of Naxxramas* and *Blackrock Mountain*, and the expansions sets *Goblin vs Gnomes* and *The Grand Tournament*, that overall include 694 cards. We have chosen this set of cards because it is the one used in the

[1] http://ugp3.sourceforge.net/
[2] https://github.com/demilich1/metastone

last season before the metagame changed to current (and still changing) one: just before the newest expansion (*League of Explorers*) appeared. This season has a good representation of different deckbuilding strategies, and we selected 4 representative human-designed Aggro decks (Hunter, Mage, Paladin, Shaman), 3 Control (Priest, Warrior, Warlock) and 2 Combo (Druid, Rogue).

The considered decks have been taken from the website of Tempo Storm[3], an American e-sports professional video game team, and selected among the ones able to reach the highest rank in the competitive ladder during season 18.

1) MidRange Druid (Combo): This Combo deck aims at using a combination of 2 cards, *Force of Nature* and *Savage Roar*, that can inflict from 14 to 30 damage to the opponent, depending on board conditions. However, the combined cost of the two cards is 9 (6+3), thus the deck has to stall for time in the early game, and slowly build a *ramp* by using specific Druid cards that increase your resources faster than the opponent's.

2) MidRange Hunter (Aggro): This Hunter deck is slower than similar Aggro decks, trading cheap cards for cost-effective minions that are harder to remove, and thus more difficult to deal with for Control decks.

3) Mage Tempo (Aggro): In CCGs, *Tempo* is basically a measurement of the speed of a player's progression through the game. This Aggro Mage deck uses cards that are able to improve one's progression, while at the same time slowing down the opponent, making enemy minions unusable for one or more turns.

4) Aggro Paladin (Aggro): A fast, effective Aggro deck, that attempts to swarm the battlefield with a lot of weak but cheap minions. It includes a few ways to neutralize problematic answers from the opponent.

5) Shadow Madness Priest (Control): A classical Priest Control deck, that makes use of a few twists. The Priest's Hero power normally would cure minions or the player; but there are a few Priest cards (with the keyword *Shadow*) that change this ability into inflicting an equal amount of damage. This deck tries to switch between curing and dealing damage depending on board conditions, to keep the match under control until it can finish off the opponent using relatively powerful creatures.

6) Oil Rogue (Combo): Another Combo deck, it exploits the Rogue's ability to play multiple cards in the same turn, reducing their costs thanks to the aid of other cards. In the very first turns this Rogue deck will try to remove the opponent's threats, all the while slowly building a large hand of cards, to finally unleash lethal damage in one single turn.

7) Mech Shaman (Aggro): This Aggro deck exploits the synergy of some Shaman cards with a specific category of minions, the *Mechs*. The Mechs are not as cheap as the minions used in other Aggro decks, but they are harder for the opponent to deal with, and interact nicely with each other, as some Mechs provide bonuses to all other Mechs in play.

[3]https://tempostorm.com/articles/meta-snapshot-18-from-warrior-to-warrior

8) Warlock MalyLock (Control): The Warlock's default power allows the player to draw extra cards, in exchange for life points. Exploiting this feature, this deck tries to go through the deck, finally obtaining a single, expensive, powerful creature: the dragon Malygos. Malygos increases the amount of damage dealt by all of the player's spells by a large quantity, allowing the Warlock to quickly close the game the turn after Malygos enters the field.

9) Warrior Control (Control): The Warrior's Hero power allows it to cumulate *Armor*, a sort of shield that protects the life points: before damaging the player's hit points, the opponent has to destroy all the Armor. Interestingly, while there is a cap for the hit points, there is no maximum limit for Armor. The Warrior tries to use the Armor to survive the early game, removing the most pernicious threats, while waiting for powerful, expensive minions that will be extremely effective in the late game.

D. Opponent decks analysis

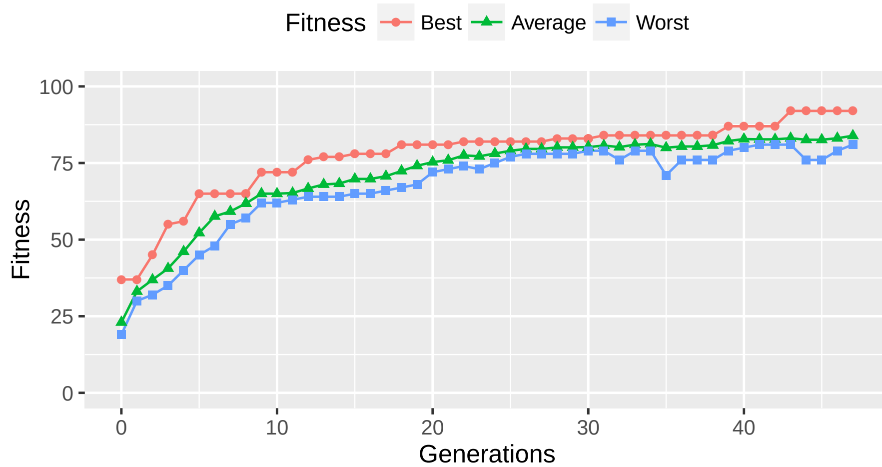
In order to get an estimate of how well MetaStone can play the human-designed decks, we run a first tournament, where each deck was paired against every other for 256 games, using all combinations of the 4 possible AIs . Thus, 11520 games were played. From that results we discovered that the AI GreedyOptimizeTurn obtained the best percentage of victories, winning 4320 games out of the 11520 (37.5%). Therefore, we set this AI as the one to bet during the rest of the experiments. Focusing on the deck behavior using this AI, Table II shows the win ratio of each one.

TABLE II
NUMBER OF GAMES WON BY THE GREEDY OPTIMIZE TURN AI. EACH DECK SHOWN IN THIS TABLE PLAYED A TOTAL OF 256 MATCHES.

Deck name	Games Won	Games Lost	Win/Lose ratio
Aggro Paladin	182	74	0.7109
Mage Tempo	177	79	0.6914
Shadow Madness Priest	152	104	0.5937
Midrange Hunter	143	113	0.5585
Mech Shaman	119	137	0.4648
Oil Rogue	106	150	0.4140
Control Warrior	104	152	0.4062
Midrange Druid	85	171	0.3320
Warlock MalyLock	83	173	0.3242

E. Experimental results

As the fitness evaluator requires a lot of computational time to simulate the large number of games for each individual, every execution of the algorithm requires several days. However, as this is a proof-of-concept, we have performed two preliminary experiments, limiting each run to a different set of cards. The first is aimed at evolving a Mage deck, normally played as a control deck, and the second is focused on a Hunter deck, usually played as aggro. In both experiments, each candidate decklist was played t times (16) against every human-designed deck in the metagame, with the exception of the deck featuring the same Hero (e=8), so each individual is tested 128 times in each evaluation.



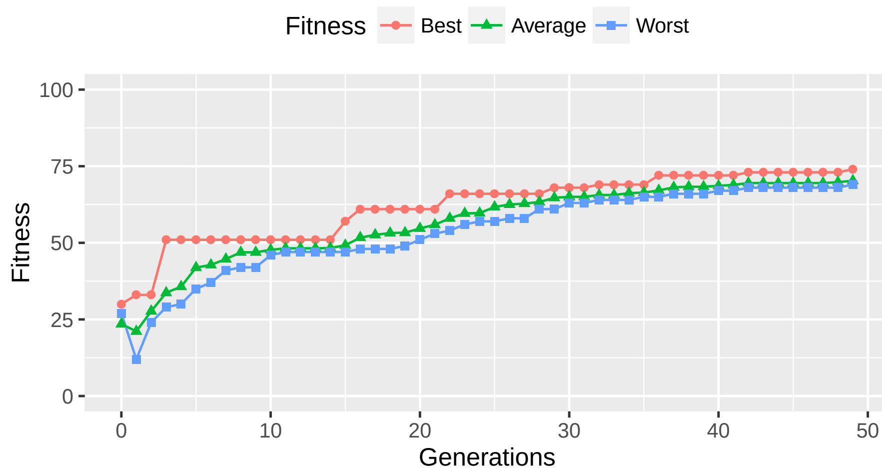


Fig. 1. Evolution of the number of victories in the population during the experiments for Mage (top) Hunter (bottom).

V. Discussion

The proposed approach is proven able to discover decks with a satisfying win ratio against competitive human-designed decks in the target metagame. Figure 1 shows the evolution of victories of the best, the average and worst individuals in each generation, showing the fitness improvement during the evolution. In both cases (Mage and Hunter) the final win ratio outperformed the Mage Tempo and Midrange Hunter decks from season 18, respectively: the best evolved Mage wins 71.87% of the matches (vs Mage Tempo, 69.14%) and the best evolved Hunter wins 57.81% of the matches (vs Midrange Hunter, 55.85%).

As the fitness evaluation is dependent on the MetaStone AI, however, our methodology might incur in *overfitting* with regards to the AI capabilities and playing style. By looking at the results of the preliminary tournament among the human-designed decks in Table II, it is immediately evident that MetaStone can use some decks (and some playing styles) better than others: for example, the *Mid-Range Druid* and *Malylock* decks, while pretty effective in the hand of an experienced human player, have relatively low performances when played by MetaStone. The human strategy for both decks relies upon waiting for specific cards (*Force of Nature+Roar* for the Druid, *Malygos* for the Warlock) and play them at the right moment, which is something that MetaStone might not be capable of. For this reason, we deem it useful to perform an expert card-by-card analysis of the deck found by the evolutionary approach, to understand whether the deck contains cards (and cards combinations) that are considered powerful by humans, or rather that MetaStone could play more effectively. The expertise comes from one of the authors, an average competitive HearthStone player, able to reach rank 10 in the season ladder (ranks 1-10 contain more or less the top 10% of the registered users [17]), that has played over 7,000 matches since the Open Beta of the game.

A. Evolved Mage deck

Figure 2 contains the best Mage decklist obtained at the end of the process. Mana curve (a histogram of the number

MINIONS	SPELLS
Antique Healbot (C)	**Blizzard** (R)
Antique Healbot (C)	**Blizzard** (R)
Argent Commander (R)	**Dragon's Breath** (C)
Baron Geddon (L)	**Fireball** (C)
Clockwork Gnome (C)	Flame Lance (C)
Clockwork Gnome (C)	**Flamestrike** (B)
Clockwork Knight (C)	**Flamestrike** (B)
Coliseum Manager (R)	**Mirror Image** (B)
Dancing Swords (C)	Polymorph Boar (R)
Fallen Hero (R)	Polymorph Boar (R)
Flesheating Ghoul (C)	**Pyroblast** (E)
Gormok The Impaler (L)	
Imp Master (R)	
Leper Gnome (C)	
Leper Gnome (C)	
Razorfen Hunter (B)	
Razorfen Hunter (B)	
War Golem (B)	
Water Elemental (B)	

Fig. 2. The best decklist obtained through the evolutionary approach for Mage. Cards that are considered particularly powerful by a human expert are highlighted in **bold**. Cards that are considered sub-optimal are in *italics*. The rarity of each card is also marked as (in decreasing order of rarity) Legendary (L), Epic (E), Rare (R), Common (C), Basic (B).

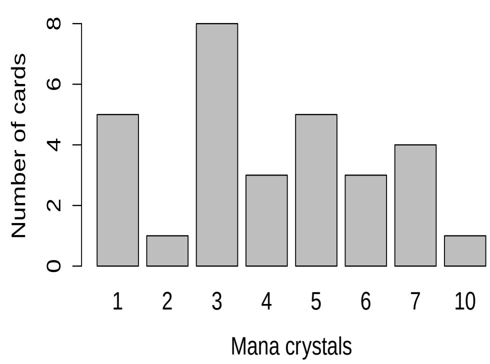


Fig. 3. Mana curve of the evolved Mage deck.

of cards grouped by cost, shown in Figure 3) shows that the deck is clearly *Aggro*, using several small, effective minions as early threats (*Clockwork Gnome, Fallen Hero, Leper Gnome, Razorfen Hunter*); blocking the opponent's minions in the mid-game through so-called *freeze* spells (*Blizzard, Frost Nova*), that prevent hit minions from acting during the next turn, or directly wiping the board with *Flamestrike*; and finally attempting to finish off the game through large minions and powerful spells (*Baron Geddon, Fireball, War Golem, Pyroblast*).

There are a few remarkable properties of the evolved deck, that we are going to describe in more detail. First of all, the majority of cards appear in two copies, the maximum number allowed, even if there is no explicit pressure to have this configuration in the fitness function. The evolutionary algorithm autonomously discovered that possessing a higher number of copies of some cards is better, since it makes the deck more reliable. Secondly, a considerable percentage of cards appearing in the deck have been often used in the competitive ladder, with the exception of *Coliseum Manager*, considered a sub-optimal minion, *Razorfen Hunter*, that has several strong competitors in the same niche, and *Gormok the Impaler*, which is sometimes used but considered very circumstantial by the players.

Finally, the deck includes several interesting synergies. *Antique Healbot, Clockwork Gnome, Clockwork Knight*: these are all minions of type *Mech*, and Clockwork Knight is able to boost other Mechs. *Imp Master, Gormok the Impaler, Razorfen Hunter*: Gormok is a Legendary minion with a powerful ability that rarely activates, since it requires the presence of at least other 4 other minions on your side of the field; the other cards all spawn extra minions on the battlefield, making it easier to activate Gormok. *Flesheating Ghoul*, that increases its strength every time a creature on the board dies, works very nicely with spells able to wipe the board such as *Blizzard* and *Flamestrike*.

Table III presents results and play statistics of the evolved decklist against each one of the human-designed decks. The deck is able to win reliably against most of the opponents, being particularly effective against Warlock, Shaman and even Paladin (the highest ranking deck in the preliminary evaluation). On the other hand, the toughest match-ups seems to be Hunter and Warrior: the former is probably often able to out-run the Mage deck in a damage race; the latter is more of a control deck relying on large, dangerous minions that are hard to deal with for the Mage deck.

B. Evolved Hunter deck

Figure 4 showcases the best Hunter decklist obtained at the end of the evolutionary process. Again, the deck is clearly Aggro (see Figure 5), but this time it exploits a relatively large selection of creature-removal spells, that can probably be used to control the field in the mid-game. Interestingly, the deck exploits either minions with a low cost (*Gadgetzan Jouster, Jungle Panther, Lance Carrier*), or with a large cost (*Gazlowe, King Krush, Piloted Sky Golem, Sneed's Old Shredder*), while featuring lots of spells with intermediate cost (*Multi-shot, Cobra Shot, Deadly Shot, Powershot*).

Again, even without a specific pressure to do so, the algorithm found it useful to include double copies of several cards. And most of the cards have seen play in the competitive ladder: *Sylvanas Windrunner* and *Loatheb* are two Legendaries considered extremely powerful, *Animal Companion*, *Kill Command* and *Unleash the Hounds* are included in almost all Hunter decks, and the same can be said for one of the two weapons, *Glaivezooka*. *Annoy-o-Tron* and *Defender of Argus* are also pretty popular, albeit they are more used in decks featuring other Heros. *Fel Reaver* is perhaps the most surprising choice, being a large, cheap creature with a huge drawback: every time the opponent plays a card, Fel Reaver destroys the top three cards of the player's deck. The usual response to Fel Reaver is to play as many cards as possible, in

TABLE III
STATISTICS OF THE BEST INDIVIDUAL USING MAGE CARDS, AGAINST ALL THE HUMAN-DESIGNED DECKS (16 TIMES PER DECK). WIN RATES OF THE MAGE TEMPO SEASON 18 (MTS18) DECK IS ALSO SHOWN AS COMPARISON.

	Druid	Hunter	Paladin	Priest	Rogue	Shaman	Warlock	Warrior
% of wins of MTS18	90.625	59.375	34.375	53.125	78.125	78.125	87.5	71.875
% of wins of Evolved Mage	87.5	62.5	50	62.5	81.25	81.25	93.75	56.25
Damage Dealt	37.25	57.13	60.50	76.81	48.50	66.75	60.44	65.00
Healing Done	1.50	5.50	4.00	4.50	5.00	5.00	4.00	4.00
Mana Spent	22.94	31.31	25.69	46.06	37.56	39.06	41.69	44.69
Cards Played	9.06	11.63	11.31	16.69	12.94	15.44	16.69	16.38
Turns Taken	7.31	8.50	7.81	10.38	9.25	9.88	10.06	10.44
Cards Drawn	7.31	8.50	7.81	10.38	9.25	9.88	10.06	10.44
Minions Played	5.50	6.19	5.38	8.19	6.88	7.75	7.25	7.81
Spells Cast	2.06	2.94	2.81	5.19	3.63	4.00	5.19	4.69
Hero Power Used	1.50	2.50	3.13	3.31	2.44	3.69	4.25	3.88
Weapons Equipped	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

MINIONS	SPELLS
Annoy-o-tron (C)	**Animal Companion** (B)
Annoy-o-tron (C)	**Animal Companion** (B)
Blackwing Technician (C)	Arcane Shot (B)
Captain Greenskin (L)	*Bestial Wrath* (E)
Defender Of Argus (R)	Flare (R)
Defender Of Argus (R)	**Kill Command** (B)
Fel Reaver (E)	**Unleash The Hounds** (B)
Fel Reaver (E)	WEAPONS
Gilblin Stalker (C)	Gladiator's Longbow (E)
Goldshire Footman (B)	**Glaivezooka** (C)
Goldshire Footman (B)	
Hungry Crab (E)	
Kezan Mystic (R)	
Loatheb (L)	
Metaltooth Leaper (R)	
Piloted Sky Golem (E)	
Raging Worgen (C)	
Ship's Cannon (C)	
Sylvanas Windrunner (L)	
Timber Wolf (B)	
Twilight Guardian (E)	

Fig. 4. The best decklist obtained through the evolutionary approach for Hunter. Cards that are considered particularly powerful by a human expert are highlighted in **bold**. Cards that are considered sub-optimal are in *italics*. The rarity of each card is also marked as (in decreasing order of rarity) Legendary (L), Epic (E), Rare (R), Common (C), Basic (B).

order to remove a huge part of its controller's deck; but maybe the MetaStone AI is not able to assess correctly the drawback, and thus in this environment Fel Reaver might be even more effective. The deck also features some questionable choices: *Hungry Crab*, *Goldshire Footman* and *Ship's Cannon* simply have too many better competitors in their respective niches; while *Bestial Wrath*'s effect is considered too circumstantial to be useful in competitive play.

Nevertheless, we can observe again some interesting synergies: *Kill Command* and *Bestial Wrath* are enhanced by minions of type Beast, and the deck has 4 of them (counting the spell *Animal Companion*, that puts a random Beast in play); *Metaltooth Leaper* boosts minions of type Mech, which the deck plays 6 of; *Flare* and *Kezan Mystic* are two cards that are particularly effective against spells used by Mage and Paladin decks, the top two of our preliminary evaluation; finally, even if not particularly cost-effective, *Captain Greenskin* can

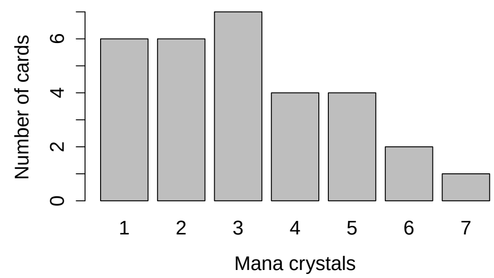


Fig. 5. Mana curve of the evolved Hunter deck.

enhance both *Gladiator's Longbow* and *Glaivezooka*.

From the results in Table IV, it is noticeable how the deck's performance is particularly good against Druid and Warlock, that are probably too slow to deal with the Hunter's aggression; while the worst match-ups are versus Priest, which is effective against low-strength minions, and Mage, another Hero whose cards are able to wipe the board, resulting in a big disadvantage for aggressive decks.

C. Remarks

While the presented proof-of-concept seems promising, there are a few weak points that are worth discussing. The most evident issue lies in the fitness function: MetaStone is a good AI, but so far it cannot attain human-comparable levels of play, especially with the settings we used for the experiments, based on a greedy choice; thus, it is hard to tell whether the optimization process is discovering generally good decks, or good decks just *for this specific AI*. This issue is hard to solve, but the fact that the evolutionary process created a deck with cards considered good by human players is at least encouraging. In future works, we plan to perform a play-by-play analysis of selected games using the evolved decks, in order to better study the problem.

Another possible issue lies in our definition of the search space. Currently, the evolutionary algorithm is free to replace a card with any other card in the set, with a low chance of obtaining an improvement. It would probably be more sensible

TABLE IV
STATISTICS OF THE BEST INDIVIDUAL USING HUNTER CARDS, AGAINST ALL OF THE HUMAN-DESIGNED DECKS (16 TIMES PER DECK). WIN RATE OF THE MIDRANGE HUNTER SEASON 18 (MHS18) DECK IS ALSO SHOWN AS COMPARISON.

	Druid	**Mage**	**Paladin**	**Priest**	**Rogue**	**Shaman**	**Warlock**	**Warrior**
% of wins of MHS18	81.25	40.625	40.625	25	65.625	62.5	65.625	65.625
% of wins of Evolved Hunter	100	37.5	43.75	37.5	62.5	62.5	68.75	50
Damage Dealt	34.81	34.06	45.69	54.00	33.94	37.69	41.50	42.56
Healing Done	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mana Spent	23.63	23.25	25.19	28.00	26.69	23.81	27.81	31.75
Cards Played	11.69	10.56	11.63	11.88	11.94	11.19	13.06	13.63
Turns Taken	7.81	7.44	7.56	7.94	7.81	7.56	8.25	8.75
Cards Drawn	7.94	7.69	7.75	7.94	7.69	7.69	8.69	9.13
Minions Played	6.00	6.00	6.13	6.94	6.44	6.00	6.75	7.31
Spells Cast	2.75	2.25	3.13	2.56	3.13	3.06	3.25	3.13
Hero Power Used	2.63	2.00	2.06	2.13	2.00	1.94	2.63	2.75
Weapons Equipped	0.31	0.31	0.31	0.25	0.38	0.19	0.44	0.44

to include mutations able to transform a card into other cards with the same cost, or similar characteristics, as human players often do when considering modifications to a decklist. The presence of such mutations could potentially help smoothen the fitness landscape, driving the algorithm towards interesting areas more effectively.

VI. CONCLUSIONS

In this paper we have presented a methodology for the automatic evolution of decks for collectible card games using *HearthStone* as a case study. An evolutionary algorithm is applied to the task. This EA uses as the structure of an individual a list of 30 cards, taken from the almost 700 available. The fitness function is the number of victories of the candidate deck against popular human-made competitive decks, performed through MetaStone, an AI able to play HearthStone. Two experiments have been conducted, and the proposed approach proved able to create a competitive Mage and Hunter deck for a specific real-world metagame, taken from Season 18 (the last one before the current, and still changing, metagame).

In future works, we plan to evolve decks for other Heroes, improve the evolutionary algorithm by adding context-aware mutations, and perform a play-by-play analysis of the decks, to try and assess the generality of our approach.

ACKNOWLEDGMENT

The authors would like to thank github user @demilich1 for creating MetaStone.

REFERENCES

[1] "Hearthstone: Heroes of Warcraft," http://eu.battle.net/hearthstone/, accessed: 24 April 2016.

[2] "Toucharcade: 'Hearthstone: Heroes of Warcraft' has more than 40 million registered players," http://toucharcade.com/2016/02/11/hearthstone-heroes-of-warcraft-has-more-than-40-million-registered-players/, accessed: 24 April 2016.

[3] P. I. Cowling, C. D. Ward, and E. J. Powley, "Ensemble determinization in monte carlo tree search for the imperfect information card game magic: The gathering," *IEEE Trans. Comput. Intellig. and AI in Games*, vol. 4, no. 4, pp. 241–257, 2012.

[4] J. C. Osborn, A. Grow, and M. Mateas, "Modular computational critics for games," in *Proceedings of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AIIDE-13, Boston, Massachusetts, USA, October 14-18, 2013*, G. Sukthankar and I. Horswill, Eds. AAAI, 2013.

[5] E. Ham, "Rarity and power: balance in collectible object games," *The International Journal of Computer Game Research*, vol. 10, no. 1, 2010.

[6] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing*, ser. Natural Computing Series. Springer, 2015. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-44874-8

[7] A. M. Mora, A. Fernández-Ares, J. J. Merelo-Guervós, P. García-Sánchez, and C. M. Fernandes, "Effect of noisy fitness in real-time strategy games player behaviour optimisation using evolutionary algorithms," *J. Comput. Sci. Technol.*, vol. 27, no. 5, pp. 1007–1023, 2012.

[8] N. Cole, S. J. Louis, and C. Miles, "Using a genetic algorithm to tune first-person shooter bots," in *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2004, 19-23 June 2004, Portland, OR, USA*. IEEE, 2004.

[9] A. M. Mora, R. Montoya, J. J. Merelo-Guervós, P. García-Sánchez, P. A. Castillo, J. L. J. Laredo, A. I. M. García, and A. Esparcia-Alcázar, "Evolving bot AI in unreal$^{\text{tm}}$," in *Applications of Evolutionary Computation, EvoApplicatons 2010: EvoCOMPLEX, EvoGAMES, EvoIASP, EvoINTELLIGENCE, EvoNUM, and EvoSTOC, Istanbul, Turkey, April 7-9, 2010, Proceedings, Part I*, ser. Lecture Notes in Computer Science, C. D. Chio, S. Cagnoni *et al.*, Eds., vol. 6024. Springer, 2010, pp. 171–180.

[10] M. Carter, M. R. Gibbs, and M. Harrop, "Metagames, paragames and orthogames: a new vocabulary," in *International Conference on the Foundations of Digital Games, FDG '12, Raleigh, NC, USA, May 29 - June 01, 2012*, M. S. El-Nasr, M. Consalvo, and S. K. Feiner, Eds. ACM, 2012, pp. 11–17.

[11] K. A. De Jong, *Evolutionary computation: a unified approach*. MIT press, 2006.

[12] J. R. Koza, *Genetic programming: on the programming of computers by means of natural selection*. MIT press, 1992, vol. 1.

[13] P. Garcia-Sanchez, A. Tonda, A. M. Mora, G. Squillero, and J. J. Merelo, "Towards automatic starcraft strategy generation using genetic programming," in *Computational Intelligence and Games (CIG), 2015 IEEE Conference on*, Aug 2015, pp. 284–291.

[14] J. M. Font, T. Mahlmann, D. Manrique, and J. Togelius, "Towards the automatic generation of card games through grammar-guided genetic programming," in *International Conference on the Foundations of Digital Games, Chania, Crete, Greece, May 14-17, 2013.*, G. N. Yannakakis, E. Aarseth, K. Jørgensen, and J. C. Lester, Eds. Society for the Advancement of the Science of Digital Games, 2013, pp. 360–363.

[15] J. J. Merelo, F. Liberatore, A. Fernández-Ares, R. H. García-Ortega, Z. Chelly, C. Cotta, N. Rico, A. M. Mora, and P. García-Sánchez, "There is noisy lunch: A study of noise in evolutionary optimization problems," in *Proceedings of the 7th International Joint Conference on Computational Intelligence (IJCCI 2015) - Volume 1: ECTA, Lisbon, Portugal, November 12-14, 2015.*, A. C. Rosa, J. J. Merelo-Guervós *et al.*, Eds. SciTePress, 2015, pp. 261–268.

[16] E. Sanchez, M. Schillaci, and G. Squillero, *Evolutionary Optimization: the μGP toolkit*. Springer Science & Business Media, 2011.

[17] "Battle.net: Hearthside chat - you're better than you think!" http://us.battle.net/hearthstone/en/blog/15955974/hearthside-chat-youre-better-than-you-think-9-18-2014, accessed: 24 April 2016.