Inteligência Artificial em Magic: The Gathering
Proposta

Guilherme Souto Schützer - 8658544
Tomás Bezerra Marcondes Paim - 7157602

28 de abril de 2017

Capítulo 1

Introdução

1.1 Regras

Magic: The Gathering é um jogo de cartas criado em 1993 por Richard Garfield que introduziu o conceito moderno de *trading card game* (jogo de cartas colecionáveis). Com um acervo de mais de 15 mil cartas diferentes, os jogadores devem criar baralhos de 60 cartas (normalmente podendo haver repetição de até quatro cartas iguais) e competir, normalmente em jogos um contra um.

Durante o jogo, os jogadores têm que lidar com informações conhecidas (as cartas já jogadas previamente e as cartas em suas mãos) e informações desconhecidas (as cartas na mão de seu oponente e a ordem das cartas em seu baralho), o que faz com que seja praticamente impossível ter informação perfeita. Além disso, ambos os jogadores podem agir a praticamente qualquer momento dentro de um turno, o que adiciona uma camada a mais de complexidade.

Já existem inteligências artificiais feitas para jogar **Magic**, mas devido à complexidade do jogo, nenhuma inteligência artificial consegue ser realmente boa no jogo (por exemplo, elas não costumam levar em consideração as cartas que estão na sua mão, então é muito fácil fazer com que a IA sempre "morda sua isca").

Nosso objetivo com o trabalho será entender o problema e criar a nossa versão de uma inteligência artificial para uma versão simplificada do jogo (limitando o conjunto de cartas disponíveis, adicionando restrições à construção dos baralhos e simplificando algumas regras do jogo).

1.2 Objetivos

- Modelar com diferentes níveis de complexidade as decisões possíveis e representações do espaço de estados.
- Implementar diferentes algoritmos de Inteligência Artificial em um cliente-jogador e comparar os resultados.
- Implementação e treino de uma ou mais IAs baseadas em aprendizado de máquina.
- Usar jogos entre duas IAs (próprias ou existentes) para avaliação de desempenho.

1.3 Regras básicas

Um jogo usual de *Magic: the Gathering* conta com dois jogadores munidos de um baralho de 60 *cards*, ambos começando com 20 pontos de vida, com o objetivo de reduzir o total de pontos de vida do oponente a 0. Para tanto, é preciso usar os *cards* disponíveis na mão, que podem representar feitiços, criaturas ou terrenos (existem outros tipos de *cards*, mas para nossa implementação vamos focar nesses três). Feitiços são *cards* que têm um efeito que acontece no momento em que são jogados e então são colocados no cemitério (como é chamada a pilha de descarte no jogo).

Figura 1.1: Divinação (Feitiço)

Por exemplo, o feitiço Divinação tem um efeito simples: ”Compre dois cards”. O jogador que joga este card pega os dois cards do topo de seu baralho e os coloca em sua mão, aumentando o leque de possibilidades. Assim, **feitiços** podem alterar o estado do jogo de diversas maneiras (como fazer com que os jogadores comprem ou descartem cards, alterar o total de pontos de vida de um jogador ou destruir uma criatura) e são a principal forma de interagir com o oponente ou desenvolver o seu lado do campo de batalha (como é chamada a zona em que ficam as criaturas e terrenos).

Criaturas são cards permanentes (uma vez jogados eles permanecem no campo de batalha até que sejam destruídos por algum feitiço ou durante o combate) que possuem poder (quantidade de dano causado em combate), resistência (quantidade de dano necessária para ser destruída) e muitas vezes habilidades que afetam o andamento do combate ou que fazem algum efeito no momento em que são jogadas.

Por exemplo, Anjo da Misericórdia tem a habilidade de Voar (que limita as interações do oponente durante o combate) e concede a seu controlador 3 pontos de vida ao entrar no campo de batalha. Além disso, seu poder e resistência são 3 e 3, respectivamente, conforme indicado na caixa no canto inferior direito.

Figura 1.2: Anjo da Misericórdia (Criatura)

Terrenos são a fonte de *mana*, que é o recurso utilizado para pagar por criaturas e feitiços. O custo de *mana* dos *cards* que não são terreno está indicado no canto superior direito do card (por exemplo, para jogar Divinação é necessário usar terrenos, sendo que um deles deve ser necessariamente uma Ilha, como na 1.3). Terrenos são, portanto, um dos tipos de *cards* mais importantes, pois sem eles não há maneira (normalmente) de jogar seus outros *cards*. A mecânica associada ao uso de recursos é **virar** (em inglês, *tap*): para conjurar

Cada “cor” de *mana* tem um tipo de terreno associado: assim como azul (Ilha), temos o branco (Planície), verde (Floresta), vermelho (Montanha) e preto (Pântano). O jogo inclusive conta com uma espécie de “diagrama filosófico” chamado *Color Pie* (do inglês, “torta das cores”, pois se trata de um gráfico circular) associando cada cor a valores e traços de personalidade. A título de curiosidade, alguns dos principais aspectos de cada cor são: conhecimento, para o azul; ordem, em branco; instinto, no verde; impulso, para o vermelho e amoralidade, característica do preto. Os aspectos de cada cor são refletidos mecanicamente nas cartas e, consequentemente, acabam muitas vezes refletindo, em nível mais abstrato, nas estratégias associadas à montagem e uso dos *decks*. Por exemplo: comprar *cards* representa acumular conhecimento, portanto Divinação é um feitiço que representa bem sua cor. Entraremos mais a fundo no assunto na seção (INSERIR SEÇÂO), na apresentação dos decks usados para teste e jogo.

No começo do jogo é decidido aleatoriamente quem será o jogador inicial, e então os dois jogadores compram uma mão inicial de sete cartas. Antes do jogo propriamente dito começar, os jogadores tomam as decisões de *mulligan*, um processo onde cada jogador pode decidir se aceita a sua mão inicial ou se quer embaralhar de volta suas cartas no deck e comprar uma nova mão com uma carta a menos, repetir o processo até que decida manter a mão. Uma vez que os dois jogadores tiverem escolhido uma mão inicial para manter, cada jogador que realizou pelo menos um mulligan olha a carta do topo de seu deck e decide se quer colocá-la no fundo.

O jogo então começa, com os jogadores alternando entre turnos, onde o jogador que "controla o turno" é chamado de *jogador ativo*, com a seguinte estrutura, simplificada (iremos entrar em maiores detalhes nas próximas seções):

- **Início do turno**: Permanentes do jogador ativo são desviradas. Jogador ativo compra uma carta de seu *deck*.
- **Primeira Fase Principal**: Jogador ativo pode jogar as cartas da mão.
- **Combate**: Jogador ativo "declara atacantes" (escolhe quais de suas criaturas irão atacar seu oponente); em seguida, seu oponente "declara bloqueadores" (escolhe quais de suas criaturas irão bloquear as criaturas atacantes). Cada criatura não-bloqueada, então, causa dano igual ao seu poder ao oponente e todas as criaturas bloqueadas e bloqueadoras causam dano entre si.
- **Segunda Fase Principal**: Igual à primeira Fase Principal.

A estrutura acima se repete até o jogo terminar, o que acontece geralmente quando algum jogador chega a 0 pontos de vida, mas também pode acontecer de outras maneiras como, por exemplo, se o baralho de um jogador acabar.

Figura 1.3: Ilha (Terreno)

Capítulo 2

Modelagem

Neste capítulo abordaremos um jogo de *Magic* como um problema de Inteligência Artificial. A estrutura de um turno, como descrita no capítulo anterior, será repetida algumas vezes até o jogo acabar, porém é necessário determinar a mão inicial de cada jogador e, por isso, trataremos este problema em separado.

2.1 *Mulligan*

Baseando-se na experiência própria, a estrutura do problema do ***mulligan*** é notavelmente diferente do resto de um jogo de *Magic*. A principal diferença é que apesar de ambos os jogadores tomarem decisões alternadamente nessa etapa, as ações do oponente não têm nenhuma influência direta sobre as ações do agente, que deve se concentrar em obter uma **mão inicial** que possibilite jogadas nos primeiros turnos do jogo. A figura 2.1 representa as escolhas que o jogador poderá fazer para determiná-la, com cada conjunto de cartas representando um conjunto de estados.

Definimos vagamente uma mão jogável se esta contém tanto cartas que impactam o estado de jogo e recursos necessários para jogá-las. Para este efeito, separaremos as cartas do deck em duas categorias: terrenos (recursos) e não-terrenos (impactam o jogo). Outro fator importante é o tamanho da mão, pois cada carta perdida representa um turno de atraso em relação a uma mão com sete cartas (uma vez que cada jogador compra uma carta por turno). Dessa maneira, em uma mão com h cartas, temos $h+1$ possibilidades, cada uma com probabilidade

$$P_h(i) = \frac{\binom{t}{i}\binom{60-t}{h-i}}{\binom{60}{h}}, \quad i = 0, \ldots, h, \tag{2.1}$$

onde t é o número de terrenos no deck e i o número de terrenos na mão. Assim, a cada nível, como é mostrado na figura 2.1, o agente deve decidir entre realizar um *mulligan* (denotado por M), que resultará em uma mão com $h-1$ cartas, sendo i terrenos com probabilidade

$$P_{h-1}(i) = \frac{\binom{t}{i}\binom{60-t}{h-1-i}}{\binom{60}{h-1}}, \quad i = 0, \ldots, h-1$$

ou manter a mão (denotado por K), o que termina o problema, seguido (para $h < 7$) de uma ação *scry*, que permite uma pequena "filtragrem" do deck para os jogadores que já acumularam alguma desvantagem (em relação ao número de cartas na mão) no processo. Por fim, a decisão também é influenciada pela informação de quem é o jogador inicial, pois ele tem virtualmente uma carta a menos, dado que no primeiro turno do jogo não se compra uma carta.

2.1.1 Aprendizado por Reforço

Dada a estrutura bem conhecida do problema, o Aprendizado por Reforço é uma técnica que pode ser aplicada para que o agente determine um comportamento (política) para alcançar uma mão "jogável" para

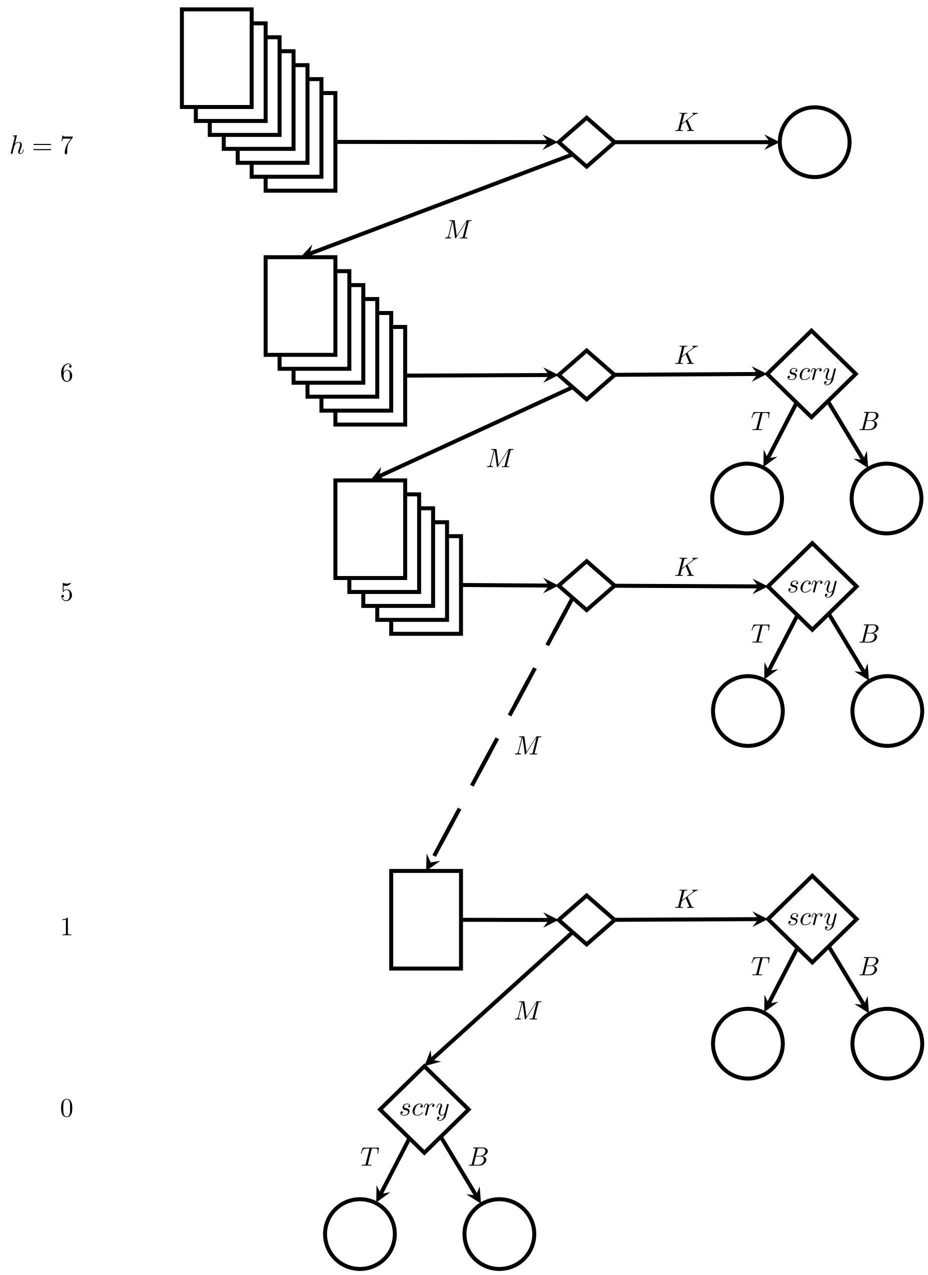

Figura 2.1: Espaço de estados de um problema de *mulligan*. Cada nível h representa o número de cartas na mão inicial do jogador.

uma partida de *Magic*.

O problema pode ser representado como uma série de estados, cada um representando uma configuração de quantidade de cartas na mão e quantidade de terrenos presentes entre essas cartas. Em cada mão, o agente pode decidir ou manter a mão e "resgatar" os pontos que aquela configuração representa, ou realizar um *mulligan*, que resultará em uma mão com uma carta a menos e distribuição de terrenos descrita na subseção anterior. Assim, depois de várias iterações, o agente irá encontrar uma política π que maximize a utilidade total $U(\pi) = \sum U^\pi(s)$, onde $U^\pi(s)$ é a função-utilidade do estado s. Usando a equação de Bellman, temos

$$U^\pi(s) = R(s) + \gamma \sum_{s'} P(s'|s, \pi(s))U^\pi(s'),$$

ou seja, a utilidade de cada estado é a recompensa $R(s)$ do estado somada à soma da utilidade de cada estado sucessor s' vezes a probabilidade de chegar em s' a partir de s e da política adotada em s (esta soma é multiplicada por um fator de desconto γ). Assim, temos que cada estado pode ser representado pelo par (h, i) e a função de Bellman, reescrita:

$$U^\pi(h, i) = R(h, i) + \gamma \sum_{i'=0}^{h-1} P(h-1, i'|h, i)U^\pi(h-1, i'). \tag{2.2}$$

A probabilidade de cada estado sucessor $(h-1, i')$ é dada pela equação 2.1:

$$P(h-1, i'|h, i) = P_{h-1}(i') = \frac{\binom{t}{i'}\binom{60-t}{h-1-i'}}{\binom{60}{h-1}}, \quad i = 0, \ldots, h-1.$$

Por fim, ainda precisamos determinar as recompensas $R(h, i)$ e o fator de desconto γ. Uma boa relação $R(h, i)$ de recompensas, por experiências pessoais, deve seguir as seguintes restrições:

- Com $h = h_0$, as recompensas (se existem) devem seguir a seguinte relação:

  $$R(h_0, 3) > R(h_0, 2) > R(h_0, 4) > R(h_0, 5) > R(h_0, 1) > R(h_0, 6) > R(h_0, 7) \gg R(h_0, 0), \tag{2.3}$$

  pois na média, uma mão ideal tem 3 terrenos, em segundo lugar 2 terrenos, etc. Uma mão sem terrenos é considerada **muito** ruim.

- $R(h, i) > R(h-1, i), \forall (h, i)$: uma mão com o mesmo número de terrenos e uma carta a menos deve sempre ser considerada pior do que a alternativa.

- $R(h, i) > R(h-1, i-1), \forall (h, i)$: uma mão com o mesmo número de não-terrenos e uma carta a menos deve sempre ser considerada pior do que a alternativa.

Dessa maneira, chegamos a uma fórmula geral, $R(h, i) = r(i) + \alpha h$, onde $r(i)$ assume valores pré-determinados de acordo com (2.3). Para $\alpha = 3$ e $r(i)$ determinado pela tabela 2.1, podemos atribuir as recompensas de acordo com a tabela 2.2

i	$r(i)$
0	-1000
1	-4
2	0
3	1
4	-1
5	-3
6	-5
7	-7

Tabela 2.1: recompensas-base para $i = 0, \ldots, 7$

	0	1	2	3	4	5	6	7
7	-979	17	21	22	20	18	16	14
6	-982	14	18	19	17	15	13	
5	-985	11	15	16	14	12		
4	-988	8	12	13	11			
3	-991	5	9	10				
2	-994	2	6					
1	-997	-1						
0	-1000							

Tabela 2.2: recompensas de cada estado calculadas com $R(i, h) = r(i) + 3h$