**NOME**

Edna Carvalho

Kaic de Oliveira Barros < [kaicbarros@gmail.com](mailto:kaicbarros@gmail.com)>

Marcos Neto Santos <marcos\_nto@hotmail.com>

Universidade Federal de Sergipe (UFS) - Curso de Sistemas de Informação – Campus Itabaiana Av. Vereador Olímpio Grande, S/N – Bairro Centro – CEP 495000-000 – Itabaiana - SE

Resumo

Abstract

**Title:**

1. Introdução

Aprendizado por Reforço é um ramo estudado pela ciência da computação. É um método de programação que consiste em um agente aprender como se comportar num ambiente dinâmico através de interações do tipo “tentativa e erro”. Nesse trabalho iremos demostrar os conceitos de Aprendizado por Reforço criando um simulador. Nele iremos utilizar a planta do Shopping Peixoto como local para o aprendizado de um robô, que por sua vez irá se locomover utilizando a base de aprendizado.

1. Aprendizado por Reforço

Em um ambiente de aprendizado por reforço, um agente é inserido em um ambiente e interage com ele através de percepções e ações. A cada passo o agente recebe como entrada uma indicação do estado atual do ambiente. O agente então escolhe uma ação a tomar, e gera a sua saída. A ação altera o estado do ambiente, e uma medida dessa mudança de estado é informada ao agente através de um valor de sinal de reforço. O comportamento do agente deve tomar ações que maximizem o valor final da soma dos reforços recebidos em um intervalo de tempo ou tentativas. Tal politica deve ser aprendida através de um processo de tentativa e erro.



Formalmente, o modelo é constituído por:

* Um conjunto discreto de estados que o ambiente pode assumir;
* Um conjunto discreto de ações que o agente pode tomar sobre o ambiente;

A tarefa que o agente deve desempenhar é encontrar uma política que mapeie os estados e ações, que maximizem a medida de reforço em longo prazo. Geralmente o sistema é não determinístico, isto é, uma mesma ação tomada em um mesmo estado pode levar a diferentes estados, com diferentes valores de retorno percebidos.

1. recursos utilizados

Para o desenvolvimento do simulador, que se trata de um vigia noturno, utilizamos uma planta de um ambiente (Shopping Peixoto) para aprendizado do mesmo. Nela foram inseridos pontos que representam os locais de acessos possíveis.

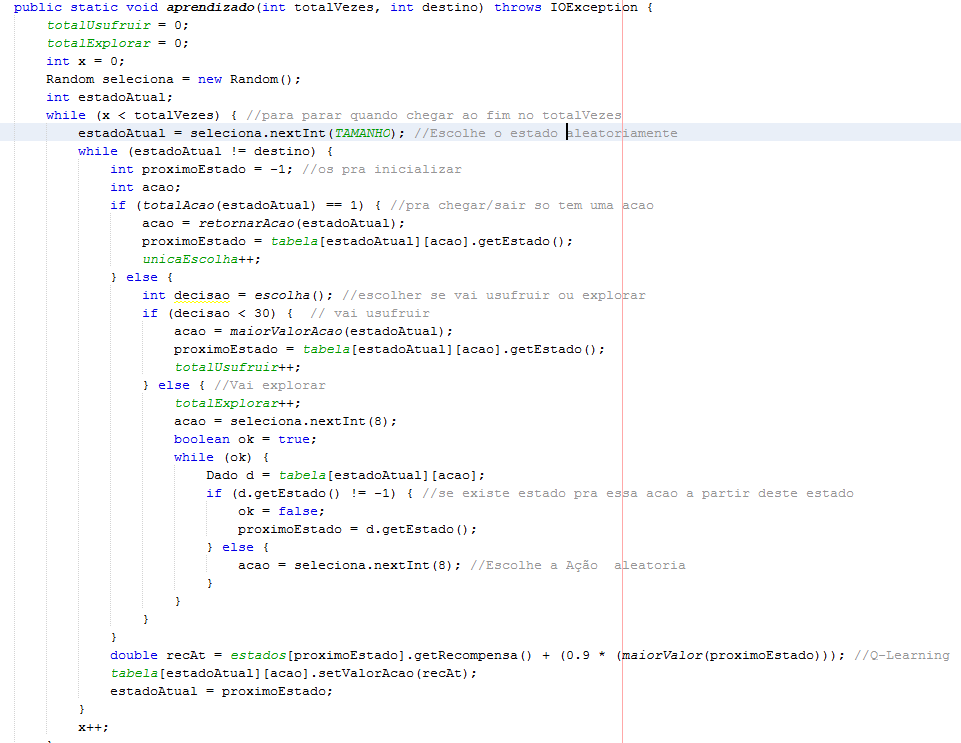
1. Projeto

O objetivo do projeto é simular o aprendizado de máquina em um ambiente, depois utilizar a base de dados aprendida para percorrer os caminhos do ambiente, de forma que simule um vigia.

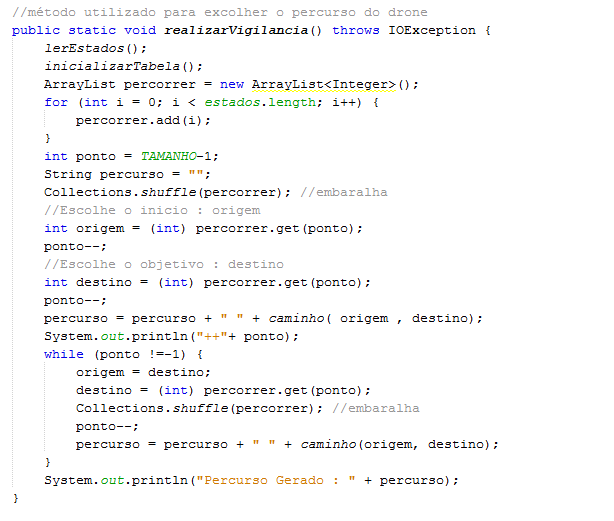
Para a construção do simulador foi necessário criar quatro classes:

* Estado, com os atributos, nome, valor e recompensa, um construtor e os métodos get e set.
* Dado, com os atributos, estado e valor da ação, construtor e métodos get e set.
* MetodosPrincipais, onde é construído todo o algoritmo de aprendizado por reforço.
* AprendizadoPorReforco, é a classe main do projeto.

No método aprendizado(int totalVezes, int destino) que recebe por parâmetro o total de vezes que ele vai aprender, e o destino. É sorteado um estado para iniciar o aprendizado, e enquanto o estado atual que foi sorteado for diferente do destino, verifica se o mesmo tem somente uma ação, se sim vai ao próximo estado com a ação, se não verifica se ele vai usufruir ou explorar, para tal escolha temos um array com cem posições embaralhadas em que pegamos o valor da ultima posição do array, se esse valor for menor que 30 iremos usufruir escolhendo a ação com maior valor do estado, se não vamos explorar sorteando uma ação possível que gere outro estado. A ação escolhida irá receber uma recompensa baseada da seguinte forma, iremos pegar o valor de recompensa do próximo estado mais a multiplicação entre 0.9 e a ação do próximo estado com maior valor. Esse passo será realizado n vezes (valor da variável totalVezes). A figura abaixo representa esse método.



O método realizarVigilancia() irá percorrer os caminhos do ambiente de acordo com o que foi aprendido no método aprendizado. A figura abaixo representa esse método.



1. conclusão

Concluímos que para a criação do simulador foi necessário a utilização dos conceitos aprendidos em sala de aula. O método de aprendizado por reforço pode ser utilizado para várias aplicações em que seja permitido o teste e erro, mas é necessário um longo prazo para aprendizado, sendo um ponto fraco da técnica.

1. Referências

https://www.maxwell.vrac.puc-rio.br/19637/19637\_4.PDF