

PROGRAMA DE PÓS-GRADUAÇÃO E PESQUISA

Processo Seletivo do Mestrado
2019

PROPOSTA PRELIMINAR DE PESQUISA

Thiago Filipe de Medeiros

Sistema de passes e marcação de futebol de robôs baseado em aprendizado de máquina profundo (*deep reinforcement learning*)

1. Área de Interesse

Planejamento de ações, estratégia e tomada de decisão em robótica móvel. A área de pesquisa é a inteligência artificial aplicada à Robótica. O programa de Pós-Graduação em Engenharia Eletrônica e Computação – PG/EEC, especificamente a divisão de Informática – EEC-I seria ideal para orientar o trabalho.

2. Motivação

2.1. Contextualização do Problema

A equipe ITAndroids foi recriada em 2012 e hoje participa de competições a nível internacional. Uma das categorias da ITAndroids é o Very Small Size Soccer (VSS), que consiste numa competição de futebol de pequenos robôs (com dimensões de cerca de 8 cm x 8 cm x 8 cm). O número de robôs por time varia com a competição: atualmente na América Latina é comum a versão 3 x 3, mas há a perspectiva a médio prazo para competições 5 x 5; internacionalmente, existem competições, como as organizadas pela FIRA (The Federation of International Robot-Soccer Association) na Coreia do Sul, que chegam à 11 x 11. Na **Fig. 1** é possível observar uma típica partida da categoria no formato 3 x 3.

Os robôs da categoria VSS possuem 2 rodas, chassi e um *hardware* para controle e comunicação com o computador central do time, além de detalhes mecânicos que facilitam domínio e condução da bola, os chamados *ball guide*. Na **Fig. 2** podem-se observar exemplares de jogadores da categoria.

O objetivo da partida é marcar gols no adversário e vencer por sistema de placar (o time com maior placar no final do jogo vence a partida). Antes da partida, algumas calibrações são realizadas para o sistema de visão computacional de cada equipe: as cores de interesse a serem avaliadas nas imagens; correção de distorções da lente da câmera e o mapeamento de coordenadas físicas reais *versus* coordenadas em *pixel*. No início da partida, cada equipe executa o programa principal, assiste a partida e aguarda alguma interrupção (seja pela marcação de uma falta, pedido de pausa por parte de alguma equipe, marcação de um gol ou encerramento do jogo), a partir da qual é permitido interagir novamente com os robôs. Os jogos normalmente são divididos em dois tempos de 5 minutos e o fim da partida se dá pelo término da contagem do tempo ou pela diferença de 10 gols no placar, o que ocorrer primeiro.

Durante a partida, o seguinte ciclo de processamento ocorre várias vezes por segundo: a câmera registra o estado atual do jogo; a imagem lida é confrontada com a calibração de cores realizada; os jogadores de ambos os times são identificados pelo padrão de cores sobre cada jogador; as posições, orientações e velocidades dos jogadores e da bola são calculados; os dados são observados pela estratégia e uma decisão em alto nível é tomada; a decisão é convertida em um comando de controle em baixo nível e enviada, via rádio, para os robôs aliados; aguarda-se a próxima leitura do estado do jogo.

Regras adicionais específicas sobre o jogo, requisitos do robô e tamanho do campo podem ser encontradas no site http://www.cbrobotica.org/wp-content/uploads/2014/03/VerySmall2009_ptbr.pdf.

2.2. Identificação do Problema de Pesquisa

Niveladas as outras componentes dos projetos dos times, como controle dos robôs, planejamentos de trajetória, processamento das imagens, qualidade do *hardware* e mecânica utilizados, a estratégia é o fator determinante para a vitória nos jogos da categoria. Porém, devido ao alto grau de abstração envolvido, a criação de uma boa estratégia para o jogo da categoria VSS é um tema complexo e tipicamente acompanhado da utilização de heurísticas.

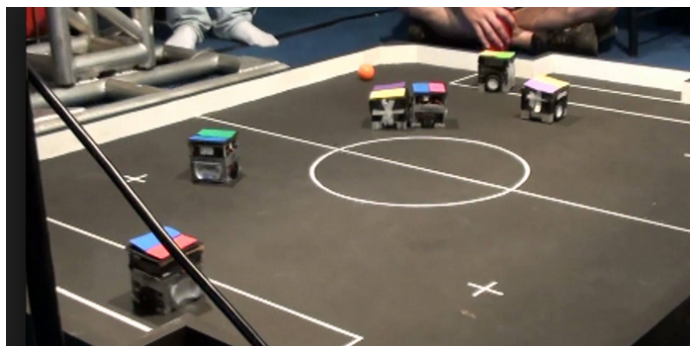


Figura 1. Exemplo de partida da categoria VSS no formato 3 x 3. As câmeras dos times ficam apoiadas em um suporte sobre o campo (não mostrado) capturando as informações do jogo. O ciclo de processamento é repetido várias vezes por segundo: aquisição e processamento de imagens; estimação de posições, orientações e velocidades; decisão estratégica em alto nível e comando de controle em baixo nível via rádio para os robôs.

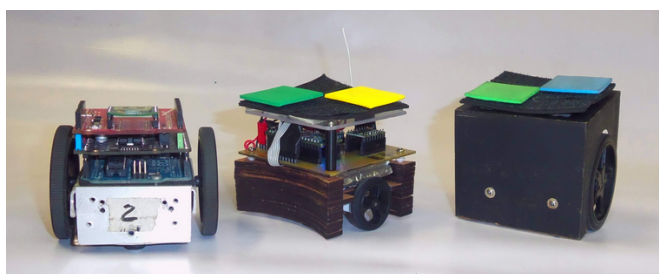


Figura 2. Exemplos de robô utilizado na competição da categoria VSS. À esquerda, um robô sem a cobertura com o circuito interno exposto. Ao centro, um robô com cobertura parcial e com o *ball guide*. À direita, um robô com a cobertura completa, mas sem o *ball guide*. Os padrões coloridos sobre os jogadores são utilizados para identificar os jogadores do time aliado e adversário, bem como estimar posição e orientação dos mesmos.

Duas formas interessantes de explorar o time adversário para conseguir vantagens estratégicas são: **O passe**, que consiste da passagem rápida da bola de um jogador para outro do mesmo time; e **A marcação**, que visa bloquear a movimentação de jogadores adversários impedindo que se posicionem adequadamente. Contudo, basta a simples consideração das inúmeras configurações de estado possíveis durante uma partida para que a ideia de construção de um código que avalie estados a partir do simples raciocínio heurístico do programador se torne inviável. Neste terreno, do reconhecimento de padrões complexos e da tomada de decisão, uma área da Inteligência Artificial (I.A.) recentemente vem ganhando força: o aprendizado por reforço ou aprendizado profundo (*deep reinforcement learning*).

Em 2015, o mundo pôde contemplar a utilização desta técnica para a criação de uma I.A., o AlphaGo, capaz de vencer de 5 x 0, no jogo chamado Go, o 3 vezes campeão europeu Fan Hui, um feito inédito e impressionante para a história da I.A. [1]. Para se ter noção da dimensão, basta entender que o jogo Go possui um número de configurações superior, por exemplo, ao xadrez. Dessa forma, o uso da técnica parece viável no cenário onde o número de estados possíveis é proibitivo para uma busca simples ou um algoritmo heurístico ingênuo, como é o caso do problema de passe e marcação.

De fato, a técnica já é usada no futebol de robôs, por exemplo, na categoria Small Size League (SSL) [2], que é semelhante ao VSS, porém com dimensões maiores e desafios de engenharia mais voltados à estratégia do sistema de multiagentes. Assim, o incremento da estratégia com sistemas de passe e marcação tem grande potencial para melhorar o desempenho do time em competições da categoria VSS.

3. Objetivo da Pesquisa

A pesquisa propõe o desenvolvimento de um sistema de passes e de marcação utilizando a técnica chamada de aprendizado de máquina profundo (*deep reinforcement learning*). Para tanto, os seguintes passos serão cumpridos:

1. Definir métricas para avaliação da qualidade de passes e marcações;
2. Utilizar tais métricas na técnica de *deep reinforcement learning* em um ambiente simulado;

3. Analisar o desempenho dos agentes simulados com as métricas arbitradas após o treinamento em (2).
4. Se possível, analisar, também, o desempenho dos agentes no mundo real com as métricas arbitradas após o treinamento em (2).

4. Informações Complementares

Pretende-se utilizar trabalhos já feitos pelas categorias da ITAndroids em anos anteriores e trabalhos decorrentes da competição Mirosot (organizada pela FIRA) como literatura base, pois são voltados especificamente para o futebol de robôs. Adicionalmente, serão consultadas literaturas gerais de *deep reinforcement learning*, como os artigos relacionados ao DeepMind [3] (projeto da Google para pesquisa e desenvolvimento na área de máquinas e I.A., criador da I.A. AlphaGo) e o toolkit Open A.I. Gym [4].

Além disso, há a intenção de utilizar técnicas tidas como ‘estado-da-arte’ de *deep reinforcement learning*, como os algoritmos T.R.P.O. (*Trust Region Policy Optimization*) e P.P.O. (*Proximal Policy Optimization*).

Enfim, um cronograma preliminar das atividades a serem desenvolvidas:

Período	Atividades
2019 – 1º Semestre	(1) Estudo de <i>deep reinforcement learning</i> e formas de aplicação ao problema proposto; (2) Cumprimento dos créditos de disciplina necessários. Obs. Este semestre é livre para os alunos de Eng. de Computação (não há disciplinas obrigatórias)
2019 – 2º Semestre	(1) Implementação e testes dos algoritmos de interesse ao problema proposto; (2) Cumprimento de eventuais créditos de disciplina remanescentes.
2020 – 1º Semestre	(1) Confecção e defesa da tese; (2) Cumprimento de eventuais créditos de disciplina remanescentes.

5. Aplicação dos Conhecimentos

O problema envolve a investigação de alternativas para incrementar a estratégia de uma categoria de competição de robôs (VSS), de forma que os resultados obtidos podem ser aplicados com propósito geral, em estratégias de sistemas multiagentes onde o aprendizado por reforço capacita o surgimento de comportamentos complexos sem a necessidade de heurísticas adicionais ou da interferência ativa do programador, ou em contexto acadêmico específico, como é o caso das competições da categoria VSS.

Para a instituição ambos os cenários são interessantes, uma vez que os incrementos na estratégia adquiridos com este projeto podem ser aproveitados em problemas diversos de aprendizado de máquina que é um tópico em foco atualmente, além do desenvolvimento da equipe ITAndroids ao melhorar o desempenho da mesma nas competições.

Além disso, os resultados gerados neste projeto podem ser aplicados em projetos de pesquisa futuros, tal qual é a visão da ITAndroids em relação à geração e compartilhamento do conhecimento, na área de Inteligência Artificial.

6. Referências

- [1] Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D. (2016). *Mastering the game of Go with deep neural networks and tree search*. *Nature*, 529 (7587), pp.484-489.
- [2] Schwab, D., Zhu, Y., Veloso, M., *Learning Skills for Small Size League RoboCup*, in *Proceedings of the RoboCup Symposium*, Springer, Montreal, Canada, Junho, 2018.
- [3] Artigos relacionados ao projeto DeepMind, disponível em <https://deepmind.com/research/publications/>, acessado em 10/10/2019.
- [4] *Toolkit de desenvolvimento e comparação de algoritmos de aprendizado por reforço*, disponível em <https://gym.openai.com/>, acessado em 10/10/2019.

São José dos Campos, 10/10/2019.