**Reward is enough for social intelligence**

O artigo estudado trata de temas bastante abstratos e – de certa forma – até mesmo filosóficos. Dessa forma, para o entendimento completo dos conceitos, é necessário primeiro definir o que é uma agente inteligente. David Silver, o autor do artigo, começa sua escrita exatamente com essa discussão, porém, não focando somente em inteligência artificial, mas com uma definição ampla de inteligência.

A conclusão desenvolvida pelo autor é que um agente inteligente é aquele que usa diversas habilidades e áreas de conhecimento distintas para tomar uma ação. Dessa forma, é somente lógico que para entender o comportamento inteligente, é necessário entender o que motiva esse conjunto de ações. Para o autor, recompensas são suficientes para motivação de ações, que por fim desencadeiam a utilização de habilidades distintas de forma coordenada.

Ainda é argumentado que recompensas são uma das melhores formas de se criar comportamentos inteligentes. Isso se dá por 2 principais motivos: (i) Como já mencionado, o uso coordenado de habilidades diversas e que não seriam relacionadas caso a recompensa não estivesse motivando as ações; e (ii) o desenvolvimento de novas habilidades por tentativa e erro. Em um exemplo com agentes autônomos inteligentes, um jogador de Atari 2600 com objetivo de maximizar sua pontuação no jogo, desenvolve diversas habilidades. Dentre elas, o acompanhamento dinâmico de um objeto, posicionamento da peça em uma região favorável para a maximização dos pontos, cálculo de um ângulo de incidência e reflexão e, por fim, estimativa de onde o projétil irá atingir os obstáculos. Diversas habilidades são desenvolvidas com tentativa e erro, sempre balizando as melhores habilidades com a maximização de recompensa.

Esse *feedback* constante é justamente a relação entre o ambiente e o agente: O ambiente é influenciado pelas ações realizadas no passado pelo agente, devolvendo uma observação para que o agente possa tomar uma nova ação. Dessa forma, se a observação contribuir para a maximização de recompensas, uma ação no mesmo sentido da última será realizada, caso contrário, uma nova estratégia terá que ser desenvolvida por tentativa e erro.

Mesmo parecendo um conceito relativamente conhecido, esse contexto fica um pouco menos obvio quando aplicado à inteligência social. Diversas teorias e modelos matemáticos já foram desenvolvidos para tentar racionalizar interações interpessoais e contextos sociais, porém, nenhum deles funciona tão bem como a maximização de recompensas. Inicialmente, é necessário mencionar a Teoria dos Jogos e o Equilíbrio de Nash, em que decisões são medidas com base na ponderação das recompensas para cada um dos participantes. Dessa forma, intrinsecamente o modelo do equilíbrio de Nash funciona para situações de soma-zero, em que para um participante ganhar, o outro tem que perder. Porém, grande parte das interações sociais não se enquadram como situações soma-zero, criando-se uma necessidade para o desenvolvimento de um outro modelo de comportamento. É assim que entra a maximização de recompensas: Dado que o ambiente é influenciado pelas ações do agente, em um ecossistema com mais de um participante, as ações de cada um deles influenciariam as decisões dos próximos agentes, causando uma observação conjunta e permitindo que os agentes não aprendam somente com suas próprias ações, mas também com os comportamentos de seus pares.

**Referência bibliográficas**

[1] AXELROD, R. A evolução da cooperação. São Paulo: Leopardo Editora, 2010.

[2] SIGMUND, K. Games of life: explorations in ecology, evolution and behaviour.

Penguim Books, 1995.

[3] Os primeiros indícios da máxima “Olho por olho, dente por dente” foram encontrados no Código de Hamurabi, de 1780 a.C., no reino da Babilônia. Esse princípio impede que as pessoas façam justiça por elas mesmas e de modo desproporcional, ao punir crimes e delitos. A pena, chamada de retaliação, deveria ser exatamente igual ao crime.

[5] PONDSPONE, W. Prisoners’ dilemma: John Von Neumann, game theory and the puzzle of the bomb. Anchor Books, 1993.

[6] AXELROD, R. The evolution of cooperation. Basic Books, 2006., p.112.

[7] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, et al., Human-level control

through deep reinforcement learning, Nature 518 (7540) (2015) 529

[8] Y. Shoham, R. Powers, T. Grenager, If multi-agent learning is the answer, what is the question?, Artif. Intell. 171 (7) (2007) 365–377