



Projeto de Formatura – Turmas 2019 – Press Release PCS - Departamento de Engenharia de

Tema:

Impacto da cognição na cooperação de agentes de aprendizagem por reforço no dilema do prisioneiro para mais de dois jogadores.

Cognição e aprendizagem são aspectos chave para o aumento da cooperação em grupos de indivíduos da mesma espécie

Do ponto de vista biológico, cooperação é uma característica vantajosa. Espécies com maior cooperação conseguem vencer outras que cooperam menos na competição por recursos. Entretanto, há dois motivos para não cooperar: medo de ser explorado por outros indivíduos e interesse em se aproveitar da cooperação dos outros. A questão que se coloca é o que leva algumas espécies a desenvolverem alta cooperação entre indivíduos e outras espécies, um comportamento solitário?

Neste estudo, modelamos o problema como o *N-Person Prisoners Dilemma (NPD)*, que considera uma situação em que cada um dos N participantes tem uma escolha entre duas ações: cooperar entre si para o "bem comum" ou desertar, seguindo seus interesses egoístas de curto prazo. Para simular a evolução das reações dos indivíduos neste dilema foram utilizados algoritmos de Aprendizado por Reforço, que ganham conhecimento por tentativa e erro. Variando as características desses indivíduos, conseguimos inferir quais são os fatores que melhoram a cooperação.

Três parâmetros inerentes ao processo de aprendizado foram analisados: (i) o *learning step*, que regula a velocidade de aprendizado; (ii) o *discounting factor*, que ajusta a importância dada para ganhos futuros em relação a ganhos imediatos; e (iii) o *exploration factor*, que regula o quanto a agente explora novas e diferentes alternativas durante o aprendizado. Observamos que a cooperação aumenta para valores altos do *discounting factor* e pequenos de *learning step* e *exploration factor*, indicando que, para aumentar a cooperação, o indivíduo deve aprender devagar para acumular conhecimento por muito tempo, deve valorizar ganhos de longo prazo à ganhos imediatos e deve mudar pouco de comportamento para que a sociedade possa adaptar-se às suas ações.

Os outros dois fatores inerentes à cognição do indivíduo foram analisados: (i) a dimensão e o tipo de informação que o agente consegue perceber a respeito de sua situação; e (ii) a política de escolha de ações do indivíduo. Ao fixarmos a política de atuação ou a percepção do indivíduo e variarmos o outro fator, não observamos melhora significativa na cooperação. Entretanto, a combinação apropriada da percepção com uma política de escolha de ações gera grupos de indivíduos com níveis de até 80% de cooperação. Estes indivíduos aprenderam um mecanismo de recuperação abrupto de um estado de baixíssima cooperação para um de alta cooperação, o qual funciona da seguinte maneira: (i) se o indivíduo não cooperou no último turno, ele continua não cooperando a não ser que ninguém esteja cooperando – neste caso extremo de ninguém estar cooperando, o indivíduo então coopera com 60% de chance, o que move o grupo automaticamente de um estado de 0% para 60% de cooperação; e (ii) se o indivíduo cooperou no último turno, ele coopera muito se muitos estão cooperando e pouco se poucos estão cooperando. Estes indivíduos resolveram o problema do medo de ser explorado porque, com baixa cooperação, eles não cooperam, evitando serem explorados por muitos indivíduos não cooperadores; entretanto, estes indivíduos mostram certa leniência em relação a um baixo número de indivíduos não cooperadores, e, devido a isso, o pico de cooperação estabiliza nos 80%.

Integrante: João Vitor de Oliveira Barbosa

Professor Orientador: Professora Anna Helena Reali Costa