

Inteligência Artificial

Primeira Lista de Exercícios

Arthur do Prado Labaki - 11821BCC017

30-09, 2022

GBC063

Resolução do item 1)

Considerando o sistema de IA do futebol de robôs (www.robocup.org), temos que:

- **Medida de desempenho:** Acertar a bola, proteger o gol aliado e marcar gols nos adversários;
- **Ambiente:** O campo de futebol e os outros robôs em campo (tanto adversários quanto aliados);
- **Atuadores:** Rodas ou pernas artificiais articuladas;
- **Sensores:** Câmeras e diversos tipos de sensores.

Resolução do item 2)

De acordo com o artigo *Inteligência Artificial e sociedade: avanços e riscos* (Sichman, 2021):

Resolução do item A)

A expressão *Inteligência Artificial* (IA) surgiu no verão de 1956 na Dartmouth College Conference.

Resolução do item B)

Inverno da IA é uma expressão que remete ao período de interesse reduzido no campo da inteligência artificial, que levou à falta de financiamento. Esse período durou desde os anos de 1975/1980 até 1987/1993.

Resolução do item C)

O atual otimismo sobre os resultados que a IA pode alcançar pode ser justificado por três fatores:

1. O custo tanto de processamento quanto de memória estão mais baratos do que nunca;
2. O surgimento de novos paradigmas (como rede neurais profundas) estão produzindo inegáveis avanços científicos;
3. A grande quantidade de dados disponíveis na Internet em razão do grande uso de recursos (como redes e mídias sociais).

Resolução do item D)

”O objetivo da IA é desenvolver sistemas para realizar tarefas que, no momento não possuem solução algorítmica viável pela computação convencional.” Existem problemas que não têm uma solução exata, ou uma única solução (como a definição do pacote de turismo, citado no artigo) e esses casos são muito complexos para o uso de algoritmos, sendo usualmente mais bem solucionados por seres humanos. Então, o objetivo da IA é conseguir realizar essas tarefas que são atualmente feitas obrigatoriamente por humanos, em que ela deve utilizar os mesmos conceitos que nos humanos utilizamos, como coleção de modelos, técnicas e tecnologias que, isoladamente ou agrupadas, resolvem problemas de tal natureza.

Resolução do item E)

Segundo o artigo, existem alguns paradigmas distintos importantes para uma IA, que são:

- **Simbólico:** Identificar o conhecimento do domínio, representá-lo utilizando uma linguagem formal de representação e implementar um mecanismo de inferência para utilizar esse conhecimento;
- **Conexionista:** A linguagem é uma rede de elementos simples, inspirada no funcionamento do cérebro, onde neurônios artificiais, conectados em rede, são capazes de aprender e de generalizar a partir de exemplos;
- **Evolutivo:** Utiliza um método probabilístico de busca de soluções de problemas, onde soluções são representadas como indivíduos, aos quais se aplicam técnicas inspiradas na teoria da evolução como hereditariedade, mutação, seleção natural e recombinação para selecionar para as gerações seguintes os indivíduos mais adaptados, até os que maximizam uma função objetivo;

- **Probabilístico:** Utiliza modelos para representar o conceito estatístico de independência condicional, a partir de relacionamentos causais no domínio.

Resolução do item F e G)

IA explicável é um conjunto de processos e métodos que permite aos usuários compreender e confiar nos resultados e na saída gerada. A IA explicável é usada para descrever um modelo de IA, seu impacto esperado e possíveis vieses. Ela ajuda a caracterizar a precisão, a imparcialidade, a transparência e os resultados na tomada de decisão desenvolvida com IA.

Contrapondo essa ideia, uma IA não explicável seria um conjunto de processos e métodos que não são confiáveis para os usuários, de forma que os resultados da sua solução não podem ser compreendidos por humanos.

Resolução do item H)

Existem algumas classes de riscos em IA, entre eles, as falhas (ou *bugs*), a sua segurança (*cybersecurity*), sua autonomia compartilhada (*shared autonomy*), sua capacidade de compreender e analisar (*sorcerer's apprentice*) e seus impactos socioeconômicos.

Atualmente é praticamente impossível de que um sistema baseado em IA consiga aniquilar a raça humana. Porém, se esse avanço da tecnologia continuar de forma explosiva (principalmente com a computação quântica), e não existir normas e regras (ética e moral) para a criação dessas novas IA, mesmo que seja difícil, essa possibilidade pode existir.

Referências

Sichman J. S. (2021). “Inteligência Artificial e sociedade: avanços e riscos”. *Estudos Avançados* 35, pp. 37–50.