FGA0221 - Inteligência Artificial

Portfólio 04

Janeiro, 2025

Tema do portfólio:

Agentes Lógicos.

Aluno: João Matheus de Oliveira Schmitz
Matrícula: 200058525
Turma: T01
Semestre: 2024.2

Janeiro, 2025

Sumário

Janeiro, 2025

1. Agente baseado em conhecimento

No campo da inteligência artificial, os **agentes baseados em conhecimento** são aqueles que recriam o processo de raciocínio humano baseado em representações internas de conhecimento. Um exemplo é a capacidade de uma pessoa de saber que, ao atravessar a rua quando um carro está passando, ela irá ser atropelada, portanto ela toma a decisão de olhar se um carro está vindo antes de atravessar.

Os agentes baseados em conhecimento fazem uso da **lógica** para realizar **inferências** sobre seu conhecimento atual do ambiente, derivando assim novas informações baseadas naquelas que ele já sabia e tomando suas ações de acordo. Essa característica é extremamente importante pois permite que os agentes aprendam novas informações, os tornando capazes de se **adaptar à mudanças no ambiente** e navegar ambientes parcialmente observáveis com maior facilidade. Além disso, também são capazes de aceitar novas tarefas sob a forma de metas descritas de modo explícito, permitindo uma maior flexibilidade em seu uso.

O componente central de um agente baseado em conhecimento é sua **base de conhecimento**, ou KB. Esta, por sua vez, se caracteriza por ser um conjunto de **sentenças**, onde cada uma representa uma afirmação sobre o ambiente. As sentenças são expressas em uma **linguagem de representação de conhecimento** e podem ser divididas em:

- **Axiomas**, quando a sentença for tomada como dada sem ser derivada de outras sentenças.
- **Derivadas**, quando são inferidas a partir de uma ou mais sentenças já conhecidas.

Para adicionar informações à base de conhecimento ou tomar uma ação com base no que se sabe, são usadas as operações **TELL** e **ASK**, respectivamente. Ambas podem envolver inferência e, segundo [2](Russell & Norvig, 2009): "A inferência deve obedecer ao requisito fundamental de que, quando se formula (com ASK) uma pergunta para a base de conhecimento, a resposta deve seguir do que foi informado (com TELL) anteriormente à base de conhecimento".

Existem duas abordagens diferentes na criação de agentes baseados em conhecimento, elas são:

- **Abordagem declarativa**: Com a utilização do TELL por parte de um programador, o agente é informado das informações que ele precisará para operar em seu ambiente.
- **Abordagem procedural**: O comportamento desejado do agente é codificado de maneira direta em seu código.

Ambas as abordagens têm suas vantagens, porém, ao invés de utilizar uma ou outra, os melhores agentes são aqueles que utilizam uma **mistura de ambas as abordagens**.

1.1. Mundo de Wumpus

O método mais comum de exemplificar o funcionamento de agentes baseados em conhecimento é através do micromundo conhecido como Mundo de Wumpus. Neste portfólio, irei usar como exemplo uma versão simplificada onde o agente é incapaz de atirar flechas.

As **características do mundo** são:

- É uma caverna composta por salas interligadas na vertical e na horizontal;
- O Wumpus devora aqueles que entram em sua sala;
- Salas podem conter poços que aprisionam aqueles que entram nela (fora o Wumpus);
- Existe um tesouro e o objetivo do agente é encontrá-lo e retornar ao início;

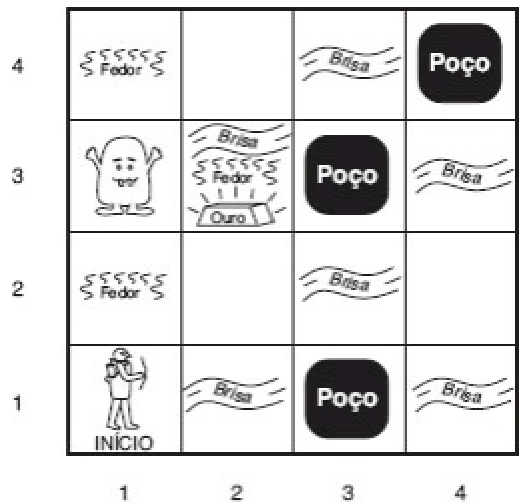

Figura 01 - Mundo de Wumpus

A **medida de performance** do agente é:

- +1000 se sair da caverna com o tesouro;
- -1000 se for devorado pelo Wumpus ou cair em um poço;
- -1 para cada ação tomada;
- O jogo termina quando o agente morre ou sai da caverna;

O **ambiente** é:

- Um grid 4x4;
- O agente começa na sala [1,1] virado para o leste;
- A posição do Wumpus e do tesouro são aleatórias;
- Uma sala tem 20% de chance de ser um poço;

Seus **atuadores** são:

- Mover para a frente;
- Virar 90º à direita;
- Virar 90º à esquerda;
- Pegar o tesouro (se estiver na mesma sala);
- Sair da caverna (na sala [1,1]);

Seus **sensores** são:

- Nas salas conectadas ao Wumpus, o agente percebe um fedor;
- Nas salas conectadas a um poço, o agente percebe uma brisa;
- Na sala onde está o tesouro, o agente percebe um brilho;
- Quando o agente tentar se mover em direção à uma parede, ele percebe um solavanco;
- As percepções serão dadas da forma [Fedor, Brisa, Brilho, Solavanco];

Para representar cada elemento, podemos utilizar:

- W para Wumpus;
- P para poços;
- T para tesouro;
- B para brisa;
- F para fedor;
- C para brilho;
- S para solavanco;
- L para o agente;

2. Lógica

Como dito anteriormente, a base de conhecimentos de um agente é representada por um conjunto de sentenças, as quais são expressas em uma linguagem de representação de conhecimento. Para ser mais exato, elas são expressas de acordo com a **sintaxe** da linguagem, a qual pode ser descrita como as regras que decidem se uma sentença está ou não de acordo com determinada linguagem.

Uma maneira simples de exemplificar sentenças verdadeiras e falsas é através de equações matemáticas. Qualquer pessoa que entenda as regras básicas da aritmética pode olhar para as seguintes equações e dizer qual está correta e qual está errada: "$x + y = 10$" e "$x7y5 =$ " , nesse caso a primeira é a correta e a segunda errada. Outra forma seria através da sintaxe de uma língua, como o português. Ao analisar as frases “Eu li um livro” e “Um eu livro li”, podemos facilmente identificar que a segunda não está de acordo com a sintaxe.

Além da sintaxe, também existe a **semântica**, que, por sua vez, define se **uma sentença é verdadeira ou falsa em cada modelo/mundo possível**. Ao tomar como exemplo as equações anteriores, podemos dizer que "$x + y = 10$" é verdadeiro quando "$x = 3$" e "$y = 7$", e falso quando "$x =- 4$" e "$y = 2$". Nesse caso, pode-se dizer que os modelos possíveis para este exemplo são todas as combinações de números reais para "x" e "y". Se uma sentença A for verdadeira no modelo M_1, podemos dizer que M_1 **satisfaz** A, ou que M_1 **é um modelo de** A. É utilizado M(A) para expressar o conjunto de todos os modelos que satisfazem A.

Existem alguns conceitos importantes para a lógica, eles são:

- **Consequência lógica**: uma sentença é consequência de outra sequência. Isso é escrito da seguinte forma: $A \vDash B$, ou seja A tem como consequência B e implica que a sentença B só é verdadeira **se e somente se** a sequência A for verdadeira. Para identificar se uma sequência A tem como consequência B, verificamos se $M(A) \subseteq M(B)$, somente se isso for verdadeiro podemos afirmar que $A \vDash B$.
- **Equivalência lógica**: duas sentenças são logicamente equivalentes se ambas são verdadeiras no mesmo conjunto de modelos. Escrevemos isso da seguinte forma: $A \equiv B$. Um exemplo simples é dizer que $(A \wedge B) \equiv (B \wedge A)$.

- **Validade**: uma sentença é válida se for verdadeira em todos os modelos. Essas sentenças também são chamadas de **tautologias** e, como exemplo, temos $A \lor \neg A$.

- **Satisfatibilidade**: uma sentença é satisfatível se for verdadeira em algum modelo qualquer. Para exemplificar, podemos utilizar os Problemas de Satisfação de Condições abordados no portfólio anterior, pois eles estão, essencialmente, perguntando se as restrições são satisfatíveis pela atribuição de alguma combinação de valores às variáveis.

2.1. Lógica Proposicional

Para representar sentenças na base de conhecimento, podemos utilizar a linguagem de representação conhecida como **Lógica Proposicional**.

2.1.1. Sintaxe

A sintaxe define quais sentenças são permitidas, e essa podem ser divididas em:

- **Sentenças atômicas**: consistem em uma única **proposição**, a qual pode ser verdadeira ou falsa (e.g. $W_{2,2}$ para representar que o Wumpus se encontra na sala [2,2]);
- **Sentenças complexas**: são construídas a partir de sentenças mais simples em conjunto com o uso de parênteses e **conectivos lógicos**.

Os conectivos lógicos, em ordem de precedência, são:

- **Negação ($\neg$)**: Nega a proposição que acompanha, como na sentença $\neg W_{2,2}$ (o Wumpus **NÃO** está na sala [2,2]);
- **E ($\land$)**: Uma sentença baseada nesse conectivo é chamada **conjunção**, como a sentença $\neg W_{2,2} \land \neg W_{1,2}$ (o Wumpus não está na sala [2,2] **E** não está na sala [1,2]);
- **Ou ($\lor$)**: Uma sentença baseada nesse conectivo é chamada **disjunção**, como a sentença $W_{2,2} \lor W_{3,1}$ (o Wumpus está na sala [2,2] **OU** na sala [3,1]);
- **Implica ($\Rightarrow$)**: Compõe sentenças no formato '$premissa$' $\Rightarrow$ '$conclusão$', onde ambas premissa e conclusão são sentenças, como exemplo

$F_{1,2} \Rightarrow W_{2,2} \vee W_{3,1}$ (fedor na sala [1,2] **IMPLICA** no Wumpus estar na sala [2,2] **OU** na sala [3,1]);

- **Se e somente se (⇔)**: Está presente em sentenças **bicondicionais**, onde um lado será verdadeiro somente se o outro também for, como na sentença $W_{3,1} \Leftrightarrow \neg W_{2,2}$(o Wumpus estará na sala [3,1] **SE E SOMENTE SE** ele não estiver na sala [2,2]);

2.1.2. Semântica

A semântica é responsável por definir as regras que determinam se uma sentença é verdadeira ou falsa. No caso da lógica proposicional, um modelo deve atribuir o chamado **valor-verdade** para cada proposição, ou seja, se ela é verdadeira ou falsa.

Ao utilizar o exemplo da Figura 02, podemos representar as informações do modelo da base de conhecimento que se encontra no topo da imagem como: $M_1 = \{W_{1,2} = falso,\ W_{2,2} = verdadeiro,\ W_{3,1} = falso\}$. A quantidade de modelos possíveis é dada pela fórmula 2^n, onde 'n' é igual ao número de proposições, embora no exemplo anterior o Wumpus só possa estar em uma sala, o que diminui esse número de 8 para 3. Entretanto, se, ao invés do Wumpus, estivéssemos falando de poços, seriam gerados todos os 8 modelos e eles seriam representados através de uma **tabela-verdade**, que nada mais é do que uma tabela onde são atribuídos todas a combinações de valores possíveis para as proposições, assim como no exemplo a seguir:

-	$P_{1,2}$	$P_{2,2}$	$P_{3,1}$
M_1	Verdadeiro	Verdadeiro	Verdadeiro
M_2	Verdadeiro	Verdadeiro	Falso
M_3	Verdadeiro	Falso	Verdadeiro
M_4	Verdadeiro	Falso	Falso
M_5	Falso	Verdadeiro	Verdadeiro

M_6	Falso	Verdadeiro	Falso
M_7	Falso	Falso	Verdadeiro
M_8	Falso	Falso	Falso

Tabela 01 - Tabela verdade de poços no Mundo de Wumpus

A semântica da lógica proposicional define 5 regras distintas, uma para cada conectivo lógico. Elas são:

- **Negação**: $\neg P$ é verdadeiro se e somente se P for falso;

P	$\neg P$
Verdadeiro	Falso
Falso	Verdadeiro

Tabela 02 - Tabela verdade da regra de negação

- **Conjunção**: $P \land Q$ é verdadeiro se e somente se P for verdadeiro e Q for verdadeiro;

P	Q	$P \land Q$
Verdadeiro	Verdadeiro	Verdadeiro
Verdadeiro	Falso	Falso
Falso	Verdadeiro	Falso
Falso	Falso	Falso

Tabela 03 - Tabela verdade da regra de conjunção

- **Disjunção**: $P \lor Q$ é verdadeiro se e somente se P for verdadeiro ou Q for verdadeiro;

P	Q	$P \lor Q$

Verdadeiro	Verdadeiro	Verdadeiro
Verdadeiro	Falso	Verdadeiro
Falso	Verdadeiro	Verdadeiro
Falso	Falso	Falso

Tabela 04 - Tabela verdade da regra de disjunção

- **Condicional**: $P \Rightarrow Q$ é verdadeiro, a menos que P seja verdadeiro e Q seja falso;

P	Q	$P \Rightarrow Q$
Verdadeiro	Verdadeiro	Verdadeiro
Verdadeiro	Falso	Falso
Falso	Verdadeiro	Verdadeiro
Falso	Falso	Verdadeiro

Tabela 05 - Tabela verdade da regra de condicional

- **Bicondicional**: $P \Leftrightarrow Q$ é verdadeiro se e só se P e Q forem ambos verdadeiros ou falsos;

P	Q	$P \Rightarrow Q$
Verdadeiro	Verdadeiro	Verdadeiro
Verdadeiro	Falso	Falso
Falso	Verdadeiro	Falso
Falso	Falso	Verdadeiro

Tabela 06 - Tabela verdade da regra de bicondicional

3. Processo de inferência

Para exemplificar conceitos como **consequência lógica** e demonstrar como um agente **infere novos conhecimentos**, podemos utilizar o Mundo de Wumpus como base. Imagine que o agente começou na sala [1,1] e não percebeu nada, andou para a sala [2,1] e percebeu Fedor. Com base nos conhecimentos básicos do agente, ele irá **inferir** que o Wumpus se encontra nas salas [2,2] ou [3,1]. Podemos então analisar três conclusões:

- **A** -> O Wumpus não está na sala [1,2];
- **B** -> O Wumpus não está na sala [2,2];
- **C** -> O Wumpus não está na sala [3,1];

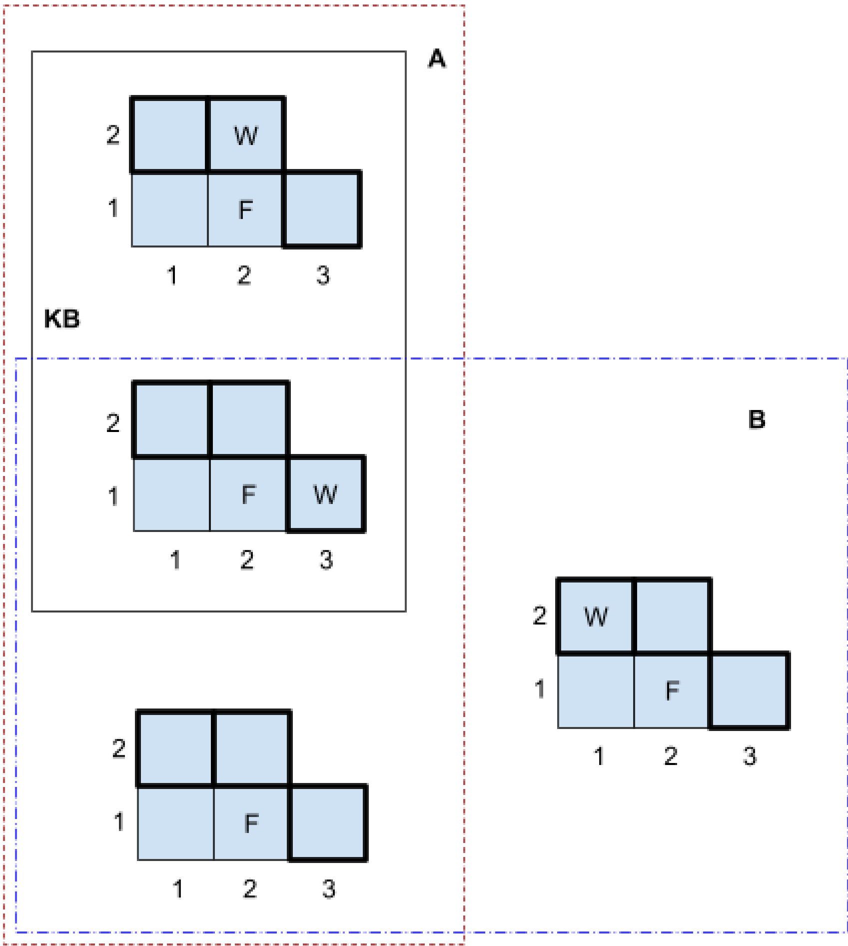

Figura 02 - Modelos possíveis para o Mundo de Wumpus

Na Figura 02, podemos ver os modelos possíveis para essa situação e analisar as conclusões A e B. Percebemos então que os modelos de KB (base de conhecimento) estão contidos nos modelos de A, ou seja $M(KB) \subseteq M(A)$. Portanto, $KB \vDash A$.

O agente então percebe que a conclusão A está correta, ou seja, **o Wumpus não está na sala [1,2]**. Além disso, os modelos de B (mostrado na figura) e de C (não mostrados na figura) não contém os modelos de KB, demonstrando que as duas conclusões podem estar incorretas e sugerindo que o **Wumpus pode estar na sala [2,2] ou na sala [3,1]**.

Segundo [2](Russell & Norvig, 2009), "... talvez ajude pensar no conjunto de todas as consequências de KB como um palheiro e de A como uma agulha. A consequência lógica é como a agulha estar no palheiro; a inferência é como encontrá-la. Essa distinção está corporificada em uma notação formal: se um algoritmo de inferência i pode derivar A de KB, escrevemos $KB \vdash_i A$, que lemos como 'A é derivável de KB por i' ou 'i deriva A de KB'".

Para realizar o processo de inferência, os agentes baseados em conhecimento utilizam os chamados **Algoritmos de inferência**, que, por sua vez, podem ser caracterizados de alguns modos distintos. Eles podem ser:

- **Corretos**, caso derivem apenas sentenças permitidas, ou seja, todo conhecimento que ele deriva tem como base o conhecimento atual, ele nunca inventa uma sentença do nada;
- **Completo**, caso consiga derivar qualquer consequência lógica;

Os melhores algoritmos de inferência são aqueles que apresentam completude e preservam a verdade de suas sentenças.

3.1. Regras de inferência

Com base no conceito de equivalência lógica, foram desenvolvidas uma série de regras para inferir sentenças com base em outras sentenças. A tabela a seguir demonstra algumas dessas regras:

Comutatividade	$(A \land B) \equiv (B \land A)$ $(A \lor B) \equiv (B \lor A)$
Associação	$((A \land B) \land C) \equiv (A \land (B \land C))$ $((A \lor B) \lor C) \equiv (A \lor (B \lor C))$

Dupla Negação	$(\neg\neg A) \equiv (A)$
Conjunção	$(A), (B) \equiv (A \land B)$
Simplificação	$(A \land B) \equiv (A)$ $(A \land B) \equiv (B)$
Adição	$(A) \equiv (A \lor B)$
Distribuição	$(A \land (B \lor C)) \equiv ((A \land B) \lor (A \land C))$ $(A \lor (B \land C)) \equiv ((A \lor B) \land (A \lor C))$
Contraposição	$(A \Rightarrow B) \equiv (\neg B \Rightarrow \neg A)$
Eliminação de Condicional	$(A \Rightarrow B) \equiv (\neg A \lor B)$
Eliminação de Bicondicional	$(A \Leftrightarrow B) \equiv ((A \Rightarrow B) \land (B \Rightarrow A))$
Modus Ponens	$(A \Rightarrow B), (A) \equiv (B)$
Modus Tollens	$(A \Rightarrow B), (\neg B) \equiv (\neg A)$
De Morgan	$\neg(A \land B) \equiv (\neg A \lor \neg B)$ $\neg(A \lor B) \equiv (\neg A \land \neg B)$
Silogismo Hipotético	$(A \Rightarrow B), (B \Rightarrow C) \equiv (A \Rightarrow C)$

Tabela 07 - Regras de inferência

4. Agente baseado em lógica proposicional

Com o conhecimento das seções anteriores, podemos começar a desenvolver um agente baseado em conhecimento que utiliza a lógica proposicional como linguagem de representação de conhecimento e que tem como ambiente o Mundo de Wumpus.

Para isso, precisamos primeiro criar uma base de conhecimento para que depois o agente possa começar a realizar ações em busca de seu objetivo. Essa base de conhecimento está ilustrada na tabela a seguir:

A sala inicial não tem o Wumpus, poços nem o tesouro.	$\neg W_{1,1}$ $\neg P_{1,1}$ $\neg T_{1,1}$
Uma brisa ocorre se e somente se existe um poço em salas adjacentes.	$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$ $B_{1,2} \Leftrightarrow (P_{1,1} \vee P_{1,3} \vee P_{2,2})$ $B_{1,3} \Leftrightarrow (P_{1,2} \vee P_{1,4} \vee P_{2,3})$ ... $B_{3,2} \Leftrightarrow (P_{3,1} \vee P_{3,3} \vee P_{2,2} \vee P_{4,2})$ ... $B_{4,4} \Leftrightarrow (P_{4,3} \vee P_{3,4})$
Um fedor ocorre se e somente se o Wumpus está em salas adjacentes.	$F_{1,1} \Leftrightarrow (W_{1,2} \vee W_{2,1})$ $F_{1,2} \Leftrightarrow (W_{1,1} \vee W_{1,3} \vee W_{2,2})$ $F_{1,3} \Leftrightarrow (W_{1,2} \vee W_{1,4} \vee W_{2,3})$ ... $F_{4,4} \Leftrightarrow (W_{4,3} \vee W_{3,4})$
Há pelo menos um Wumpus.	$W_{1,1} \vee W_{1,2} \vee W_{1,3} \vee W_{1,4} \vee \ldots \vee W_{4,4}$
Há no máximo um Wumpus.	$\neg W_{1,1} \vee \neg W_{1,2}$ $\neg W_{1,1} \vee \neg W_{1,3}$ $\neg W_{1,1} \vee \neg W_{1,4}$ ... $\neg W_{4,3} \vee \neg W_{4,4}$

Há pelo menos um tesouro.	$T_{1,1} \vee T_{1,2} \vee T_{1,3} \vee T_{1,4} \vee \ldots \vee T_{4,4}$
Há no máximo um tesouro.	$\neg T_{1,1} \vee \neg T_{1,2}$ $\neg T_{1,1} \vee \neg T_{1,3}$ $\neg T_{1,1} \vee \neg T_{1,4}$ $\neg T_{1,2} \vee \neg T_{1,3}$ ... $\neg T_{4,3} \vee \neg T_{4,4}$
Um brilho ocorre se e somente se o tesouro está naquela sala.	$C_{1,1} \Leftrightarrow T_{1,1}$ $C_{1,2} \Leftrightarrow T_{1,2}$ $C_{1,3} \Leftrightarrow T_{1,3}$ ... $C_{4,4} \Leftrightarrow T_{4,4}$

Tabela 08 - Base inicial de conhecimento de um agente do Mundo do Wumpus

Para que a base de conhecimento continue consistente mesmo após o agente começar a realizar ações, é necessária a inclusão de um novo eixo: o **tempo**. Isso ocorre pois o agente pode dizer em momento que não sente um fedor na sua posição atual, mas após se mover para outra sala, ele irá sentir; nesse caso, ocorre uma contradição. Ao levar isso em conta, podemos impedir que essas contradições ocorram. Chamamos os aspectos do mundo que variam ao longo do tempo de **fluentes**. Aqueles que não sofrem mudanças ao longo do tempo podem ser chamados de **variáveis atemporais**.

O uso de etapas de tempo se prova imprescindível na representação das tomadas de ação dos agentes. Isso ocorre pois o **modelo de transição** desses agentes é escrito como um conjunto de sequência lógicas. Por exemplo, se o agente estiver na sala [1,1], virado para o leste e andar para frente, ele irá se encontrar na sala [2,1] no instante de tempo seguinte. Representamos isso através do seguinte **axioma de efeito**:

$$(L^0_{1,1} \wedge VendoOLeste^0 \wedge AndarFrente^0) \Rightarrow (L^1_{2,1} \wedge \neg L^1_{1,1})$$

Entretanto, ao declarar axiomas de efeito, ocorre um problema onde não sabemos mais se os aspectos fluentes que não fazem parte do axioma permanecem ou não inalterados. Esse problema é conhecido como **Problema de Persistência**, e uma de suas soluções é a declaração de **axiomas de**

persistência, que são responsáveis por afirmar quais proposições fluentes continuam as mesmas. Por exemplo, ao seguir em frente o agente deve continuar virado para a mesma direção, então se ele estava olhando para o leste no instante 't' e andou para frente, então ele continuará olhando para o leste no instante 't+1'. Isso é escrito da seguinte forma:

$$(AndarFrente^{t}) \Rightarrow (OlhandoOLeste^{t} \Leftrightarrow OlhandoOLeste^{t+1})$$

Existem também os **axiomas de estado sucessor**, que definem o estado de um fluente em um instante futuro dependendo se uma ação que o afeta foi ou não utilizada no instante atual. Eles são utilizados no lugar de outros axiomas devido a existência de **Problemas de Persistência Inferencial**. Um exemplo é: se o agente está na sala [2,1] e se mover para a frente, estando virado para o oeste, ele chegará na sala [1,1] no instante seguinte; e será representado por:

$$(L_{1,1}^{t+1}) \Rightarrow (L_{2,1}^{t} \land (VendoOOeste^{t} \land AndarFrente^{t}))$$

A capacidade de um agente baseado em lógica proposicional de navegar seu ambiente e alcançar suas metas é incrível, mas isso só é possível quando a base de conhecimento do agente contém todos as informações básicas sobre o mundo e seus aspectos fluentes possuem todos os axiomas necessários para que não ocorra contradições.

5. Impressões sobre o conteúdo

A utilização de lógica proposicional como linguagem de representação de conhecimento, nos agentes discutidos neste portfólio, demonstra claramente a importância que outros campos de conhecimento podem ter no avanço da área de inteligência artificial.

Inicialmente, quando vi as regras do Mundo de Wumpus, pensei que a criação de um agente que conseguisse navegar esse micromundo fosse algo relativamente simples, pensamento este que foi reforçado quando vi que utilizava lógica proposicional, algo que eu já tenho conhecimento sobre.

O ponto de virada na minha percepção foi quando enxerguei a quantidade imensa de sentenças necessárias para popular por completo a base de conhecimento de um agente do Mundo de Wumpus, um micromundo relativamente simples.

Como conclusão, posso dizer que durante meu estudo deste conteúdo percebi que embora os blocos fundamentais usados na construção desses agentes sejam simples, eles são utilizados para criar algo bem mais complexo do que inicialmente acreditei.

6. Referências

[1] SOARES, Fabiano Araujo. Slides da aula 10 à 11. Apresentação do PowerPoint.

[2] Russell, S. & Norvig, P. **Artificial Intelligence - A Modern Approach**. 3ª ed. Pearson Education, Inc. 2009.

[3] GOUVEIA, Rosimar. Lógica Matemática. **Toda Matéria**, [s.d.]. Disponível em: https://www.todamateria.com.br/logica-matematica/. Acesso em: 15 jan. 2025.