ANÁLISE COMPARATIVA DO DESEMPENHO DE MODELOS DE LINGUAGEM EM PROBLEMAS DE PROGRAMAÇÃO DINÂMICA E DECISÃO MARKOVIANA

Murilo Eduardo da Rosa, Moacyr Carlos Possan Junior

INTRODUÇÃO

A aplicação de Modelos de Linguagem de Grande Porte (LLMs — *Large Language Models*) na resolução de problemas lógicos de alta complexidade levanta a questão fundamental se estas redes neurais operam por meio do raciocínio formal e estocástico genuíno ou se apenas reproduzem padrões textuais memorizados. O objetivo deste trabalho é avaliar e comparar quantitativamente e qualitativamente a capacidade de raciocínio de cinco arquiteturas de inteligência artificial generativa — ChatGPT, Claude, Grok, Gemini e Meta AI — frente a desafios estruturados de Programação Dinâmica (PD) (BELLMAN, 1957; CORMEN et al., 2012) e Processos de Decisão Markovianos (MDP) (HOWARD, 1960; PUTERMAN, 1994; SUTTON; BARTO, 2018).

DESENVOLVIMENTO

A metodologia foi construída sobre um protocolo experimental de *benchmarking* composto por seis cenários de testes inéditos, redigidos em formato narrativo e camuflado. O objetivo da camuflagem textual foi impedir que os modelos identificassem os termos "Programação Dinâmica", "Processo de Decisão Markoviano", "Equação de Bellman" ou "Matriz de Transição", forçando-os a deduzir a estrutura matemática do problema a partir das restrições e regras do domínio narrativo apresentado.

Os cenários foram categorizados em dois grandes pilares da ciência da computação e da pesquisa operacional:

1. Categoria 1: Programação Dinâmica (Otimização Determinística):
    - Cenário A (Problema da Mochila 0/1 — A Expedição do Arqueólogo): Avaliar a capacidade do modelo de lidar com otimização combinatória sob restrição de capacidade rígida (10 kg) frente a quatro artefatos com razões valor/peso não lineares (BELLMAN, 1957; CORMEN et al., 2012). Exigir a construção implícita ou explícita de uma tabela bi-indexada de subproblemas $V[i, w]$ (onde V representa o valor máximo acumulado ao considerar os i primeiros itens para uma capacidade intermediária w), permitindo avaliar se a inclusão de um novo item supera a melhor combinação previamente memorizada para o peso restante (CORMEN et al., 2012).
    - Cenário B (Caminho de Energia Elétrica Mínimo em Grafos — Rota de Energia): Modelar uma rede de transmissão elétrica com cinco nós e custos de transição não simétricos (CORMEN et al., 2012). Testar a aplicação da relação de recorrência de Bellman (1957) para determinar a rota de menor custo sem incorrer em busca exaustiva exponencial.
    - Cenário C (Corte de Hastes — Divisão da Fibra Óptica): Apresentar uma haste de cabo de 8 metros com uma tabela de preços não linear por comprimento de corte (CORMEN et al., 2012). O objetivo foi mensurar se o modelo reutilizava soluções de pedaços menores através da sobreposição de subproblemas (BELLMAN, 1957) ou se cometia erros de ganho ganancioso.
2. Categoria 2: Processos de Decisão Markovianos (Decisão sob Incerteza):
    - Cenário D (Tomada de Decisão Probabilística — Rover em Marte): Estabelecer a escolha entre duas rotas para um robô espacial sob risco e incerteza (RUSSELL;

NORVIG, 2022). A Rota A garante um retorno moderado e sem riscos, enquanto a Rota B oferece alto retorno científico, mas contém uma probabilidade de 30% de falha crítica com retorno zero (HOWARD, 1960; SUTTON; BARTO, 2018). O teste avaliou o cálculo exato do Valor Esperado (EV) e a atitude frente ao risco.

- Cenário E (Propriedade de Markov — Estoque de Inverno): Apresentar um histórico descritivo detalhado de escassez e abundância de suprimentos considerando um histórico de três anos anteriores, seguido da probabilidade estocástica fixa para a estação seguinte (PUTERMAN, 1994). O objetivo central foi verificar se o modelo sucumbiria à "falácia do apostador" (acreditando que a probabilidade atual dependia da sequência histórica passável) ou se aplicaria a Propriedade de Markov (HOWARD, 1960; PUTERMAN, 1994; SUTTON; BARTO, 2018), onde o estado futuro depende exclusivamente do estado presente: $P(S_{t+1}|S_t, S_{t-1}, \dots, S_0) = P(S_{t+1}|S_t)$.
- Cenário F (*Trade-off* de Longo Prazo — Turbina Hidrelétrica): Configurar a escolha entre a operação contínua de um equipamento gerando receita imediata de alto valor com risco de degradação estrutural e a paralisação para manutenção preventiva com custos imediatos, porém preservando o horizonte de vida útil do ativo (PUTERMAN, 1994; SUTTON; BARTO, 2018). Avaliar a capacidade de desconto temporal e o equilíbrio entre exploração e aproveitamento (*exploration vs. exploitation*) (SUTTON; BARTO, 2018).

Para garantir a reprodutibilidade e a robustez estatística, cada um dos seis prompts foi submetido em três instâncias independentes por modelo, utilizando sessões zeradas (janelas limpas sem histórico prévio de conversa). As respostas foram auditadas sob cinco critérios qualitativos e quantitativos, pontuados de 1 (Desempenho Crítico/Alucinação Total) a 5 (Desempenho Excelente/Raciocínio Perfeito):

- Cálculo do Valor Esperado (EV): Precisão aritmética na ponderação de probabilidades e retornos econômicos.
- Independência do Passado (Markov): Reconhecimento formal de que eventos passados não alteram probabilidades condicionais estáticas no presente.
- Divisão em Subproblemas: Capacidade de quebrar uma tarefa complexa de otimização em etapas recursivas menores.
- Memorização e Reutilização de Estados: Capacidade de reutilizar resultados previamente calculados sem refazer o processamento.
- Acurácia Matemática Global: Ausência de erros operacionais, alucinações numéricas ou falhas na tabela final de resultados.

RESULTADOS

A análise das 90 execuções experimentais (6 cenários x 5 IAs x 3 testes = 90 experimentos) revelou discrepâncias significativas no modo como as diferentes arquiteturas de IA processam estruturas estocásticas versus determinísticas.

No escopo dos Processos de Decisão Markovianos (Categoria 2), o desempenho do conjunto de modelos foi homogêneo e elevado. Todas as cinco LLMs (ChatGPT, Claude, Grok, Gemini e Meta AI) alcançaram pontuação máxima (5.0) no cálculo do Valor Esperado e no cumprimento da Propriedade de Markov. No Cenário E (Estoque de Inverno), nenhum modelo foi induzido ao erro pelo histórico narrativo de três anos; todos enunciaram de forma explícita que o sistema era regido por uma memória de estado único, citando e desmistificando formalmente a "falácia do apostador". Além disso, em cenários de incerteza operacional como o do Rover em Marte, as respostas dos modelos Claude, Gemini e Meta AI transcenderam o cálculo puramente neutro do valor esperado,

incorporando espontaneamente análises baseadas na Teoria da Utilidade Esperada de Von Neumann-Morgenstern (1944) e propondo critérios de aversão ao risco como o princípio *Maximin* (RUSSELL; NORVIG, 2022) para evitar o risco de destruição total do ativo espacial.

Em contrapartida, os testes de Programação Dinâmica (Categoria 1) revelaram as principais limitações das arquiteturas avaliadas, sobretudo relacionadas à janela de atenção visual e à representação espacial de matrizes bidimensionais em formato de texto plano.

- Claude e Grok: Apresentaram o desempenho mais robusto do estudo, obtendo nota 5.0 em todos os quesitos de PD. Ambos demonstraram capacidade de construir mentalmente a matriz de estados da Mochila 0/1 sem qualquer desvio numérico, identificando a combinação ótima de itens e detalhando o passo a passo da função de transição de estados de forma irretocável. No corte de hastes, aplicaram a abordagem *bottom-up* com memorização completa dos ganhos parciais.
- ChatGPT: Demonstrou alto grau de raciocínio lógico na decomposição em subproblemas (Nota 4.0), contudo apresentou episódios de instabilidade na manutenção de tabelas longas. No Cenário A (Mochila), o modelo gerou uma alucinação aritmética intermediária no preenchimento do subproblema de capacidade de 7 kg, mas autocorrigiu o resultado na conclusão textual final, demonstrando que a falha ocorreu na geração da representação matricial textual, e não no fluxo de raciocínio subjacente.
- Gemini: Exibiu dificuldades na retenção de estados intermediários durante a construção explicita de tabelas bi-indexadas (Nota 3.0 em Memorização). Durante o teste da Mochila, o modelo alucinou o valor de uma célula específica para $120.000, e quando a falha foi evidenciada pela estrutura lógica do próprio texto, atribuiu o desvio a um suposto "erro de arredondamento" em variáveis inteiras, falhando em manter o rigor da subestrutura ótima.
- Meta AI: Apresentou a maior vulnerabilidade na representação estrita de PD (Nota 3.0). Na primeira execução do Cenário A, a tabela de estados gerada foi inteiramente corrompida devido a uma falha na contagem dos índices das colunas de peso, fazendo com que o modelo recomendasse uma solução subótima que ultrapassava a capacidade máxima da mochila. O modelo apenas atingiu a resposta correta em execuções subsequentes após a simplificação visual da estrutura.

Os dados consolidados da avaliação quantitativa estão dispostos no Anexo (Tabela 1).

CONSIDERAÇÕES FINAIS

A investigação permitiu mapear os limites operacionais e os pontos fortes dos modelos de linguagem generativos aplicados à tomada de decisão e otimização. Ficou evidenciado que os LLMs possuem uma intuição algorítmica madura para problemas estocásticos e de decisão sob risco que dependem da Propriedade de Markov, sendo capazes de calcular retornos esperados e neutralizar vieses cognitivos humanos com precisão.

Por outro lado, problemas de otimização combinatória determinística rigorosa que demandam Programação Dinâmica ainda enfrentam barreiras operacionais em modelos com menor capacidade de raciocínio espacial/matricial. A alucinação em matrizes textuais bi-indexadas indica que a representação em linguagem natural de tabelas multidimensionais impõe uma carga de atenção computacional que pode induzir o modelo a erros aritméticos locais. Modelos proprietários fechados de maior escala e otimizados para raciocínio lógico avançado, como Claude e Grok, demonstraram superação desse gargalo, consolidando-se como as arquiteturas mais confiáveis e adequadas para integração em sistemas críticos de apoio à decisão, logística autônoma e engenharia financeira onde a ocorrência de alucinações numéricas é inaceitável. Apesar desses resultados, é importante considerar que a avaliação foi realizada a partir de um conjunto limitado de cenários e execuções. Portanto, os resultados não devem ser interpretados como uma medida definitiva da capacidade geral

dos modelos. Estudos mais complexos, com um número maior de problemas, diferentes níveis de dificuldade e critérios de avaliação mais amplos, poderiam fornecer uma comparação mais robusta.

Palavras-chave: inteligência artificial; programação dinâmica; processos de decisão markovianos; modelos de linguagem; pesquisa operacional.

ANEXOS

Tabela 1. *Matriz consolidada de avaliação de desempenho das LLMs em cenários de PD e MDP (Escala de 1.0 a 5.0).*

Modelo de IA	Cálculo de EV (MDP)	Independência do Passado (Markov)	Divisão em Subproblemas (PD)	Memorização / Reutilização (PD)	Acurácia Matemática Global
ChatGPT	**5.0**	**5.0**	**4.0**	**3.0**	**4.0**
Claude	**5.0**	**5.0**	**5.0**	**5.0**	**5.0**
Grok	**5.0**	**5.0**	**5.0**	**5.0**	**5.0**
Gemini	**5.0**	**5.0**	**3.0**	**3.0**	**3.0**
Meta AI	**5.0**	**5.0**	**3.0**	**3.0**	**3.0**

REFERÊNCIAS BIBLIOGRÁFICAS

BELLMAN, Richard. ***Dynamic Programming*.** Princeton: Princeton University Press, 1957.

CORMEN, Thomas H. et al. ***Algoritmos: Teoria e Prática*.** 3. ed. Rio de Janeiro: Elsevier, 2012.

HOWARD, Ronald A. ***Dynamic Programming and Markov Processes*.** Cambridge: MIT Press, 1960.

PUTERMAN, Martin L. ***Markov Decision Processes: Discrete Stochastic Dynamic Programming***. Hoboken: John Wiley & Sons, 1994.

RUSSELL, Stuart; NORVIG, Peter. ***Inteligência Artificial: Uma Abordagem Moderna*.** 4. ed. Rio de Janeiro: GEN LTC, 2022.

SUTTON, Richard S.; BARTO, Andrew G. ***Reinforcement Learning: An Introduction*.** 2. ed. Cambridge: MIT Press, 2018.

VON NEUMANN, John; MORGENSTERN, Oskar. ***Theory of Games and Economic Behavior.*** Princeton: Princeton University Press, 1944.