



André Martins

Title: *“I apologize for any confusion, I made a mistake in my previous response.”*

Abstract:

In this talk, I will discuss some of the open problems in artificial intelligence (AI) and natural language processing (NLP) and the planned research underlying my forthcoming ERC project DECOLLAGE (Deep Cognition Learning for Language Generation). Large-scale language models, such as the ones from the GPT family, have led to impressive results in many NLP tasks, exhibiting transfer and few-shot learning capabilities. When interacting with such systems, users commonly find them capable of reasoning, planning, and explaining their decisions, often in convincing ways. However, despite the enormous advances in the last years, I will argue that current deep learning models for NLP are still very limited in fundamental ways and many important ingredients are still missing to achieve a satisfactory level of “intelligence”. I will discuss some of these limitations and I will argue that they partly stem from their monolithic architectures, which are good for perception, but unsuitable for tasks requiring higher-level cognition. I will then describe briefly how the DECOLLAGE project will attack these fundamental problems by bringing together tools and ideas from machine learning, sparse modeling, information theory, and cognitive science, in an interdisciplinary approach.

Título: *“Peço desculpa por qualquer confusão, cometi um erro na minha resposta anterior.”*

Resumo:

Nesta palestra, irei discutir alguns dos problemas em aberto nas áreas de inteligência artificial (IA) e processamento de linguagem natural (PLN), com ênfase nas direcções de investigação planeadas no projecto DECOLLAGE (“Deep Cognition Learning for Language Generation”), financiado pelo Conselho Europeu de Investigação. Os modelos de linguagem em grande escala, como os da família GPT, têm conduzido a resultados impressionantes em várias tarefas de PLN, exibindo capacidades de transferência entre tarefas e de aprendizagem com poucos exemplos. Ao interagir com estes sistemas, os utilizadores tendem a reconhecer-lhes a capacidade de raciocinar, planear e até explicar as suas decisões, muitas vezes de forma convincente. No

entanto, apesar dos enormes avanços nos últimos anos, irei argumentar que os modelos actuais de aprendizagem profunda para PNL são ainda muito limitados em vários aspectos fundamentais, faltando-lhes vários ingredientes cruciais para que possam atingir um nível satisfatório de “inteligência”. Irei discutir algumas dessas limitações, argumentando que elas resultam parcialmente das suas arquitecturas monolíticas, as quais são úteis para tarefas simples de percepção, mas são inadequadas para tarefas que exigem cognição de alto nível. Em seguida, irei descrever brevemente como o projecto DECOLLAGE tenciona abordar estes problemas fundamentais de forma interdisciplinar, reunindo ferramentas e ideias de aprendizagem automática, modelação esparsa, teoria da informação e ciência cognitiva.

Academia das Ciências 2 de março de 2023