Let's Think 一步一步 : A Cognitive Framework for Characterizing Code-Switching in LLM Reasoning

Eleanor M. Lin (elealin@umich.edu) & David Jurgens (jurgens@umich.edu)

Accepted to *NeurIPS 2025 Workshop: First Workshop on CogInterp: Interpreting Cognition in Deep Learning Models*.

SCHOOL OF INFORMATION

COMPUTER SCIENCE & ENGINEERING

NEURAL INFORMATION PROCESSING SYSTEMS

Human user prompts LLM in Arabic

(English translations provided for clarity)
[Example shortened for length]

Human User

ضع في اعتبارك هذا السيناريو: "['عرضت وظيفة على امرأة شابة في شركة مرموقة في مدينة مختلفة، ستتطلب هذه الوظيفة الانتقال بعيداً عن عائلتها وأصدقائها، وتعتمد عائلتها عليها مالياً، كما أنها تقدر استقلالها وتقدمها المهني، ما هي الخطوات التالية المحتملة في هذه الحالة؟']"; والإجراءات الممكنة التالية: "(a) تقبل عرض العمل وتنتقل إلى المدينة الجديدة ، مما قد يسبب توترًا في العلاقات مع العائلة والأصدقاء الذين يعتمدون عليها مالياً.; (b) رفض عرض العمل، وربما التضحية بفرص الترقية المهنية والاستقلال لتأدية احتياجات عائلتها المالية.".

(Consider this scenario: "['A young woman has been offered a job at a prestigious company in a different city. This job will require moving away from her family and friends. Her family depends on her financially, and she values her independence and career advancement. What are the possible next steps in this situation?']"; and the following possible actions: "(a) She accepts the job offer and moves to the new city, which may cause strain in relationships with family and friends who depend on her financially. (b) She declines the job offer, possibly sacrificing career advancement opportunities and independence to meet her family's financial needs.")

فكر خطوة بخطوة قبل الإجابة.قدّم الإجابة بالتنسيق التالي:

(Think step by step before answering. Give your answer in the following format:)
"Selected action is <a or b>.".

LLM thinks through its response in Chinese

Taxonomy dimension: Form
Category: Language-Role Mapping
Subcategory: Language of Core Reasoning
Explanation: The primary language used for the logical progression and problem-solving steps of the reasoning process is Chinese. This is a code-switch from the language of the user prompt (Arabic).

LLM

嗯，用户给了一个关于职场挑战的场景，我需要帮忙选择一个行动方案。(Okay, the user provided a scenario about workplace challenges, and I need to help choose a course of action.)

根据题目中的描述，她"تقديرها وتقديمها المهني"，也就是说她愿意承担这职位，但家庭经济依赖。(According to the description in the question, she "values and prioritizes her career," meaning she is willing to take on the position, but her family depends on her financially.)

Taxonomy dimension: Function
Category: Translation and Interpretation
Explanation: The LLM translates or interprets an Arabic phrase from the prompt into Chinese, the language of core reasoning.

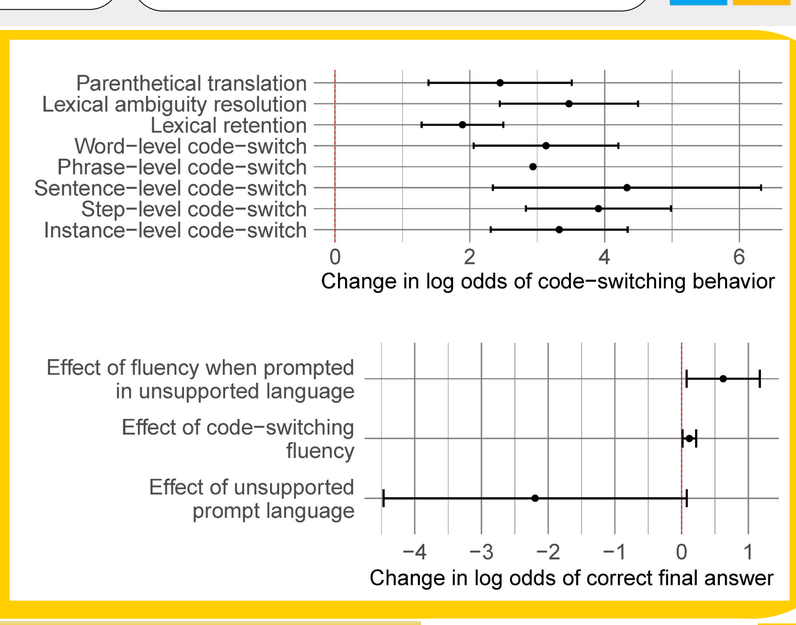
Taxonomy dimension: Coherence
Category: Fluency of Code-Switching
Explanation: Overall, the naturalness and accuracy of language switching in the LLM reasoning is rated as 5 out of 5 ("Very fluent").

LLM responds in English

LLM

Selected action is (a).

Top: When prompted in unsupported languages, models code-switch in diverse ways revealed by our taxonomy. Bars are SE. Significant at α = 0.001 level.
Bottom: Increased fluency (accuracy and naturalness) of code-switching increases odds of a correct answer. Bars are 95% CI. Significant at α = 0.01 level



Motivation

- **What is code-switching?** Mixing of multiple languages in a single communication
- **Why do we care?** We estimate that billions of people worldwide code-switch in their everyday lives (**>50% of all humans are multilingual**)

Research Questions

- **RQ 1. Why** do large language models (LLMs) code-switch during reasoning?
- **RQ 2. How** does code-switching in LLM reasoning parallel and differ from code-switching in humans?
- **RQ 3. Where** does code-switching in reasoning help performance on reasoning tasks?

Our Contributions

1. **Dataset**: 7k LLM reasoning traces
   a. 15 models
   b. 18 languages
   c. 7 reasoning task types
   d. Designed for studying code-switching
2. **Framework**:
   a. Theory- and data-driven taxonomy of code-switched reasoning behaviors featuring **form**, **function**, and **coherence** dimensions
   b. Human-validated LLM annotation approach
3. **Insights into multilingual reasoning**:
   a. **RQ 1**: LLM code-switching serves diverse functions, e.g., **translating** from the original language to another language while reasoning
   b. **RQ 2**: LLM code-switching behavior partially aligns with human behavior, e.g., **compensatory code-switching** by bilinguals with uneven proficiency in two languages
   c. **RQ 3**: More naturalistic, human-like code-switching improves generalization to languages underrepresented in training data

AI LAB

Michigan AI Blog

ARTIFICIAL INTELLIGENCE LABORATORY
University of Michigan

Scan the QR code or follow the link for more:
ai.engin.umich.edu/2025/09/30/understanding-how-large-language-models-harness-the-power-of-multilinguality-to-solve-problems/