המאמר היומי של מייק: 20.04.25  
Training Large Language Models to Reason in a Continuous Latent Space

המאמר מציג רעיון חדשני ומתבקש (לעניות דעתי) לשיפור תהליכי הנמקה (reasoning) של מודל שפה. כמו שאתם בטח יודעים אנו גורמים למודלי שפה לחשוב על ידי הכנסה לפרומפט ביטוים כמו ״think step by step״ או טוקנים מיוחדים של חשיבה כמו <think> וכדומה. זה גורם למודל ״לפלוט״ את שרשרת הנמקה בצורה של טוקנים, כלומר של טקסט. היתרון בגישות אלו שאנו יכולים לנתח את שרשרת החשיבה של מודל ולשפר אותה כי אנו רואים אותה כטקסט.

אבל האם מודלי שפה חייבים ״לחשוב״ בשפה שלנו? לא בהכרח. למשל הגרסא הראשונית של מודל DeepSeek המפורסם R0 בנתה שרשראות הנמקה בכמה שפות (אמנם שפות טבעיות). זה קרה כנראה בגלל שהמודל אומן עם RLHF בלבד ולא קיבל תגמול על כתיבה קוהרנטית אלא בעיקר על נכונות התשובה. כלומר המודל לא אומן על שרשראות הנמקה מסודרת (שהן מן הסתם מכילים שפה אחת). זה גרם לכך שהמודל פיתח שפה משלו (שזה ערבוב של כמה שפות כמו אנגלית, סינית, רוסית ועוד) בדרך לפתרון נכון של שאלות הדורשות חשיבה.

המאמר שנסקור היום עושה צעד נוסף בכיוון הזה. הרי מודלי שפה לא חייבים לחשוב בשפות שאנו, בני אדם, מבינים, נכון? בשביל כך יש להם את מחרב הייצוג שלהם, כלומר המרחב הלטנטי. הרי מודל שפה לא חושב באמצעות מילים ובמשפטים כמונו אלא פועל במרחב וקטורי שכל וקטור ייצוג של טוקן. אז המחברים אמרו את הדבר הבא: בוא נחליף שרשרת הנמקה בשרשראות הנמקה לטנטיות (וקטוריות) ללא תרגום לשפה האנושית. אז המודל מאומן להחליף שרשראות הנמקה בשפה טבעית בסדרה של וקטורים.

וזה בדיוק מה שנעשה באימון המודל. המחברים מאמנים מודל לפלוט וקטורים במקום עבור כמה שלבי הנמקה ראשונים. כלומר המודל מאומן (בו זמנית) להחליף שלבים 1-3 או 1-6 של שרשרת הנמקה בוקטורים. כלומר המודל מתחיל מהמשטר הלטנטי (latent mode) שהמחשבות שלו הם הוקטורים וממשיך במשטר שפתי (language mode) שבו הפלט הוא שפה טבעית. כמובן שיש טוקן שמפריד בין משטרים אלו כלומר <eot>.

מאמר עם כיוון מאוד מעניין שאני צופה לו עתיד גדול.

https://arxiv.org/pdf/2412.06769