המאמר היומי של מייק - 10.01.25  
Chain of Thought Empowers Transformers to Solve Inherently Serial Problems

המאמר מציג ניתוח תיאורטי של כיצד (Chain of Thought (CoT מאפשר למודלי טרנספורמר להתמודד עם חישובים סדרתיים(לא מקביליים). המחברים הוכיחו חסמי expressiveness פורמליים ומציגים מחלקת complexity חדשה (איך ניתן לתרגם complexity class? חוץ ממחלקת סיבוכיות) המאפיינת את יכולות החישוב של טרנספורמרים עם CoT.

התרומה התיאורטית העיקרית של המאמר טמונה בחסמי האקספרסיביות שהוא מוכיח. באמצעות ניתוח מתמטי ריגורוזי, המחברים מוכיחים שהטרנספורמרים בעלי עומק קבוע עם דיוק סיביות קבוע מוגבלים לפתרון בעיות ממחלקת סיבוכיות הנקראת AC0 ללא CoT (משפט 3.1). עם זאת, הם מראים שעם T שלבי CoT, טרנספורמרים מסוגלים לפתור כל בעיה הניתנת לחישוב על ידי שרשרת בוליאניות בגודל T (משפט 3.3). לתוצאה זו יש השלכות עמוקות, שכן היא קובעת שמספר פולינומיאלי של צעדי CoT מאפשר לטרנספורמרים לחשב כל פונקציה במחלקת -P/poly.

המסגרת התיאורטית שפותחה במאמר מורכבת משלושה חלקים עיקריים. ראשית, המחברים מציגים ניתוח מקיף של חישובי low-precision לעומת floating-point בטרנספורמרים. שנית, הם מבססים קשרים עמוקים עם תורת סיבוכיות על ידי הגדרת מחלקת מורכבות חדשה [(CoT[T(n), d(n), s(n), e(n) המאפיין את החישובים בטרנספורמר עבור מספר שלבי CoT המסומן בתור (T(n, המימד החבוי של הטרנספורמרים (d(n, דיוק ייצוג נומרי (s(n והאקספוננטה שלו (e(n. שלישית, הם משלבים תורת אוטומטים על ידי שימוש במשפט הפירוק של קרון-רודס לניתוח יכולות הטרנספורמר.

מבחינה ארכיטקטונית, העבודה מספקת ניתוח מפורט של רכיבי טרנספורמר, כולל מנגנוני self-attention, FFNs, קידודי מיקום, והשפעות שכבות נרמול. ניתוח זה מספק אפיון מדויק של יכולות חישוביות ומבסס מיפויים ברורים בין תכונות ארכיטקטוניות וחסמים תיאורטיים.  
  
השפעת המאמר חורגת מניתוח טרנספורמרים. על ידי הצגת מחלקת סיבוכיות חדשה לחישובי הטרנספורמרים, הוא מגשר בין מודלי חישוב וקלאסיים עם אלו המבוססים למידה עמוקה. הכלים המתמטיים שפותחו משלבים מסגרות תיאורטיות מרובות ביעילות ויוצרים קשרים חדשים בין תחומים נפרדים בעבר. במבט קדימה, עבודה זו פותחת כיווני מחקר מבטיחים רבים, במיוחד בהבנת השימוש המיטבי ב- CoT והמגבלות היסודיות של ארכיטקטורות טרנספורמר. המסגרת התיאורטית שנוסדה כאן צפויה לשמש כבסיס לניתוח חידושים עתידיים בארכיטקטורת רשתות נוירונים ואסטרטגיות הנחיה.

https://arxiv.org/abs/2402.12875