AI और चेतना का भ्रम: मैकेनिकल तुर्क से लार्ज लैंग्वेज मॉडल तक



मौलिक प्रतीकात्मक चित्र: pcsmantra.com

खबरों में क्यों

“AI and the consciousness illusion” - द हिंदू में 16 नवंबर 2026 को प्रकाशित जैकब कोशी के ऑप-एड ने एक बुनियादी सवाल उठाया है: क्या आज के AI सिस्टम वाक़ई **समझ** रखते हैं, या वे 18वीं सदी के **मैकेनिकल तुर्क** की तरह सिर्फ़ बुद्धिमत्ता का प्रदर्शन हैं? लेखक की दलील है कि जिस तरह तुर्क शतरंज खेलने वाली एक “मशीन” थी जिसके अंदर एक इंसान छिपा होता था, उसी तरह आज के भाषा-मॉडल की चतुराई **डेटा-प्रसंस्करण का परिणाम** है, चेतना का नहीं। यह बहस AI की भूमिका, ज़िम्मेदारी और नीति-निर्माण के लिए केंद्रीय है - किसी AI को दुनिया जीतने के लिए उसमें **इच्छित भविष्य की कल्पना और साज़िश रचने की बुद्धि** होनी चाहिए, जिसका अभी कोई साक्ष्य नहीं है।

मैकेनिकल तुर्क: प्रदर्शन बनाम वास्तविकता

1770 में **वोल्फ़गैंग वॉन केम्पेलन** ने जो शतरंज खेलने वाला “ऑटोमेटन” बनाया, वह दशकों तक यूरोप में दौरे करता रहा और कई चुनौतीपूर्ण मुक़ाबले जीता। अंदर चक्र-परिचालन (cogs and clockwork) दिखाया जाता, लेकिन असल में एक **छिपा हुआ शतरंज-उस्ताद** चालें चलाता था। यानी “मशीन की बुद्धि” एक प्रवेश-द्वार था - दर्शकों ने जो समझा, वह जो था, उससे अलग था। आज के AI के साथ तुलना सटीक है: आउटपुट की प्रवाहमयीता देखकर हम अंदर की प्रक्रिया को “समझ” मान बैठते हैं - जबकि यह सांख्यिकीय पैटर्न-पूर्वानुमान है।

दर्शनशास्त्र की तीन कसौटियाँ

- **ट्यूरिंग टेस्ट (1950):** एलन ट्यूरिंग ने “क्या मशीन सोच सकती है?” के स्थान पर प्रश्न पलट दिया - क्या मशीन बातचीत में इंसान से अप्रभेद्य व्यवहार कर सकती है? यह **व्यवहार का परीक्षण** है, चेतना का नहीं - इसे पास करना चेतना का प्रमाण नहीं।
- **चाइनीज़ रूम (जॉन सर्ल, 1980):** जो व्यक्ति चीनी न जानता हो, वह नियम-पुस्तिका से चीनी प्रतीकों का सही जवाब निकाल सकता है - बिना अर्थ समझे। सर्ल का तर्क: **सिंटैक्स से सेमांटिक्स नहीं बनती**; प्रोग्राम चलाना समझ होना नहीं है।
- **हार्ड प्रॉब्लम (डेविड चामर्स, 1995):** मस्तिष्क की प्रक्रियाएँ समझ आना अलग है; **अनुभव होना** (subjective experience) क्यों और कैसे उत्पन्न होता है, यह अनसुलझा है - तब यह दावा करना असुरक्षित है कि किसी कंप्यूटेशन से चेतना उत्पन्न हो गई।

हालिया इतिहास: LaMDA प्रकरण की सीख

2022 में Google के एक इंजीनियर (ब्लेक लेमोइन) ने दावा किया कि LaMDA चैटमॉडल **चेतन (sentient)** हो गया है; कंपनी और विज्ञान-समुदाय ने इसे खारिज किया - मॉडल की भाषा-कुशलता को अनुभव का साक्ष्य मानना उसी “तुर्क भ्रम” की पुनरावृत्ति थी। वैज्ञानिक सहमति आज भी यही है: वर्तमान AI में चेतना या आत्म-अनुभव का **कोई विश्वसनीय साक्ष्य नहीं**।

यह भेद व्यावहारिक रूप से क्यों मायने रखता है

- **ज़िम्मेदारी:** यदि AI में एजेंसी/चेतना नहीं, तो उसके आउटपुट की उत्तरदायित्व **मानव और संस्थाओं** पर ही रहती है - "मशीन ने किया" बहाना नहीं बन सकता।
- **नीति-निर्माण:** चेतना के भ्रम से "AI अधिकार" जैसी समयपूर्व बहसें असली जोखिमों - पूर्वाग्रह, भ्रामक सूचना, रोज़गार, सुरक्षा - से ध्यान हटा सकती हैं।
- **हाइप का सुधार:** क्षमता का यथार्थ मूल्यांकन (टूल के रूप में AI) निवेश, शिक्षा और विनियमन के बेहतर निर्णय देता है; यह AI-विरोध नहीं, AI-साक्षरता है।
- **भारत का संदर्भ:** IndiaAI मिशन और MeitY के AI-गवर्नेंस प्रयास "मानव-केंद्रित, जवाबदेह AI" पर ज़ोर देते हैं - चेतना-भ्रम से सावधानी इसी दिशा में कदम है।

संतुलित निष्कर्ष

AI की उपलब्धियाँ असली हैं और बढ़ेंगी भी; सवाल उनकी **व्याख्या** का है। परीक्षा-दृष्टि से सुरक्षित निष्कर्ष: AI को सशक्त उपकरण मानें, पर "बुद्धिमान" आउटपुट और "चेतन अनुभव" में अंतर बनाए रखें। जिस दिन चेतना का साक्ष्य मिलेगा, वह वैज्ञानिक दावा होगा - भाषा-कुशलता का अनुमान नहीं।

प्रिलिम्स परीक्षा अभ्यास प्रश्न

प्रश्न 1: "चाइनीज़ रूम" विचार-प्रयोग का प्रतिपादन किसने किया और यह क्या दर्शाता है?

(a) एलन ट्यूरिंग - मशीन की बुद्धि का व्यवहार-परीक्षण

(b) जॉन सर्ल - सिंटैक्स से सेमांटिक्स/समझ उत्पन्न नहीं होती

(c) डेविड चामर्स - अनुभव की "हार्ड प्रॉब्लम"

(d) वोल्फ़गैंग वॉन केम्पेलन - शतरंज-ऑटोमेटन की सीमाएँ

उत्तर और व्याख्या

उत्तर: (b) जॉन सर्ल का 1980 का चाइनीज़ रूम तर्क देता है कि प्रतीक-प्रसंस्करण (सिंटैक्स) अपने-आप में अर्थ या समझ (सेमांटिक्स) नहीं बनाती। ट्यूरिंग = व्यवहार-परीक्षण; चामर्स = हार्ड प्रॉब्लम; केम्पेलन = मैकेनिकल तुर्क का निर्माता।

प्रश्न 2: 18वीं सदी का "मैकेनिकल तुर्क" क्या था?

(a) ओटोमन साम्राज्य का एक यांत्रिक खिलौना-सिपाही

(b) शतरंज खेलने वाला स्वचालित यंत्र जिसमें छिपा हुआ इंसान चालें चलाता था

(c) प्रारंभिक गणना-यंत्र जिसने बैबेज के विश्लेषिक इंजन को प्रेरित किया

(d) ट्यूरिंग द्वारा बनाया गया प्रथम बुद्धि-परीक्षण उपकरण

उत्तर और व्याख्या

उत्तर: (b) 1770 में केम्पेलन द्वारा बनाया गया यह शतरंज-"ऑटोमेटन" वास्तव में अंदर बैठे शतरंज-उस्ताद द्वारा संचालित होता था - मशीन-बुद्धि का प्रसिद्ध भ्रम।

मेन्स परीक्षा अभ्यास प्रश्न

प्रश्न: "AI की बढ़ती क्षमताओं को चेतना का प्रमाण मानना एक दार्शनिक और नीतिगत भूल है।" मैकेनिकल तुर्क के उदाहरण और चाइनीज़ रूम तर्क के आलोक में इस कथन का मूल्यांकन कीजिए। भारत के AI-शासन के लिए इसके निहितार्थ भी लिखिए। (250 शब्द)

रूपरेखा: भूमिका - क्षमता बनाम चेतना का भेद; तुर्क-सादृश्य - प्रदर्शन बनाम वास्तविक प्रक्रिया; सर्ल और चामर्स की कसौटियाँ; LaMDA प्रकरण की सीख; नीतिगत निहितार्थ - जवाबदेही, जोखिम-आधारित विनियमन, मानव-केंद्रित AI (IndiaAI मिशन); निष्कर्ष - संतुलित, साक्ष्य-आधारित दृष्टिकोण।

FAQ

संक्षिप्त उत्तर: "चेतना का भ्रम" का अर्थ है AI की प्रवाहमयी भाषा और कुशल आउटपुट को आत्म-अनुभव या समझ का प्रमाण समझ लेना - ठीक उसी तरह जैसे दर्शकों ने मैकेनिकल तुर्क को सच्ची मशीन-बुद्धि समझ लिया था। वर्तमान AI में चेतना का कोई विश्वसनीय वैज्ञानिक साक्ष्य नहीं है।

प्रश्न: मैकेनिकल तुर्क और आज के AI में क्या समानता है?

दोनों में बाहरी प्रदर्शन (शतरंज-चालें / प्रवाहमयी उत्तर) देखकर भीतर की प्रक्रिया के बारे में गलत अनुमान लगाया जाता है। तुर्क में छिपा इंसान था; आज के भाषा-मॉडल में डेटा पर प्रशिक्षित सांख्यिकीय पैटर्न-पूर्वानुमान है - दोनों में "दिखने वाली बुद्धि" और "अंदर की प्रक्रिया" अलग हैं।

प्रश्न: क्या ट्यूरिंग टेस्ट पास करना चेतना का प्रमाण है?

नहीं। ट्यूरिंग टेस्ट व्यवहार की अप्रभेद्यता मापता है। सर्ल के चाइनीज़ रूम ने दिखाया कि अर्थ-समझ के बिना भी नियम-आधारित सही उत्तर संभव हैं; इसलिए व्यवहार-परीक्षण चेतना का पर्याप्त प्रमाण नहीं।

प्रश्न: इस बहस का नीति से क्या संबंध है?

यदि AI चेतन नहीं, तो उत्तरदायित्व मानव/संस्थाओं पर रहता है; "AI अधिकार" जैसी बहसें असली जोखिमों (पूर्वाग्रह, भ्रामक सूचना, सुरक्षा) से ध्यान भटका सकती हैं। इसलिए नीति को साक्ष्य-आधारित, जोखिम-केंद्रित और मानव-केंद्रित रखना आवश्यक है।

स्रोत

- द हिंदू (ऑप-एड, 16 नवंबर 2026): AI and the consciousness illusion - जैकब कोशी - यह एक **राय-लेख** है; इस विषय के लिए कोई आधिकारिक स्रोत मौजूद नहीं, अतः लेख राय-आधारित विश्लेषण है
- Stanford Encyclopedia of Philosophy: The Chinese Room Argument (जॉन सर्ल, 1980)
- पृष्ठभूमि संदर्भ: मैकेनिकल तुर्क (1770) का इतिहास