Latest AI News

Skills from a year of Purposeful Rationality Practice by Raemon

Source: Featured posts - LessWrong 2.0 viewer
Published: Wed, 18 Sep 2024 02:05:58 +0000

URL: https://www.greaterwrong.com/posts/thc4RemfLcM5AdJDa/skills-from-a-year-of-purposeful-rationality-practice

Summary: The article discusses the author's year-long endeavor to develop skills for addressing complex and confusing problems, primarily through the concept of "metastrategic brainstorming." This method emphasizes the ability to generate creative solutions in seemingly insurmountable situations. The author experimented with various problem-solving techniques, including Thinking Physics questions and strategy games, ultimately identifying several key skills that facilitate effective problem-solving. Notable skills highlighted include the importance of taking breaks to enhance cognitive function, optimizing working memory through effective tools, and cultivating patience to resist the urge to rush into decisions. The author stresses recognizing feelings of confusion and understanding when clarity is achieved. Additionally, the practice of "backchaining" is emphasized, encouraging individuals to visualize the end goals and work backward to identify necessary steps. The article advocates for regularly asking, “What is my goal?” to ground efforts and enhance strategy formulation. The author concludes by recommending the maintenance of multiple hypotheses or plans to avoid tunnel vision, which fosters intellectual flexibility. The synthesis of these techniques is positioned as a crucial approach for anyone tackling significant and challenging problems, blending both established and innovative strategies for clearer problem-solving pathways.

Contra papers claiming superhuman AI forecasting by nikos

Source: Featured posts - LessWrong 2.0 viewer
Published: Thu, 12 Sep 2024 18:10:50 +0000

URL: https://www.greaterwrong.com/posts/uGkRcHqatmPkvpGLq/contra-papers-claiming-superhuman-ai-forecasting

Summary: The article critiques various recent claims about the forecasting capabilities of AI Language Models (LLMs), arguing that many published studies misleadingly suggest that these models can rival or surpass human forecasters' abilities. The authors note that claims of "superhuman" forecasting are often poorly defined and misleading, especially in public discourse. They emphasize that reliable AI forecasting depends substantially on effective information retrieval (IR) and high-quality quantitative reasoning, which current models struggle to achieve. Several studies are examined, revealing issues such as insufficient data, reliance on low-quality information, and flawed methodologies. For instance, the study by Schoenegger et al. is critiqued for not adequately demonstrating LLM performance relative to human benchmarks, while Halawi et al. is considered the most rigorous but still falls short of substantiating claims of human-level capability. Overall, while the article acknowledges some progress in AI forecasting, it asserts that no current autonomous AI

forecaster matches the accuracy of top human forecasters and emphasizes the need for rigorous definitions, data quality, and methodological rigor in evaluating AI performance in this domain.