An LLM-Based Decision Support System for Strategic Decision-Making

Majd Alkayyal[1,2], Simon Malberg[1,2][0009−0000−3781−4700] (✉), and Georg Groh[1]

[1] School of Computation, Information and Technology, Technical University of Munich, 80333 Munich, Germany {majd.alkayyal,simon.malberg}@tum.de, grohg@in.tum.de

[2] These authors contributed equally.

Abstract. We introduce StrategicAI, a decision support system (DSS) for organization leaders and managers responsible for making strategic decisions on the course of their organizations. The main idea behind StrategicAI is to reduce the inherent complexity of strategic decisions using logic trees. These tree structures recursively decompose the involved problem and solution spaces into less-complex parts until these parts become straightforward to answer based on known information. StrategicAI follows a human-AI collaboration philosophy where users are in full control of the tree decompositions applied and can decide flexibly which parts of the trees they create manually and which parts the artificial intelligence (AI) creates. The AI is a multi-agent system based on retrieval-augmented large language models (LLMs). To obtain data-driven insights, StrategicAI actively retrieves facts from user-uploaded files and online sources and incorporates them throughout the created trees. A demo video is available at https://youtu.be/uKx8L4XZI9A. We release our code at https://github.com/PortgasXDXMajd/StrategicAI.

Keywords: Decision Support System · Strategic Decision-Making · Large Language Models · Human-AI Collaboration.

1 Introduction

Strategic decision-making is a cornerstone of organizational success, requiring leaders to navigate complex challenges and shape an organization's future. Strategic decisions—spanning market entry, technology adoption, resource allocation, strategic partnerships, and others—are characterized by complexity, ambiguity, and uncertainty while often carrying significant consequences. The stakes are high, as the quality of strategic decisions, or even the failure to make a decision at all, is a key determinant of an organization's performance and survival. For example, Netflix's decision to shift from DVD rentals to streaming in 2007 boosted its market capitalization to over \$500 billion by 2025. In contrast, Kodak's failure to embrace digital photography—despite its invention of the technology—caused its market capitalization to plummet from \$31 billion in 1997 to under \$300 million by 2012, ending in bankruptcy.

Managers often operate under high pressure, expected to deliver maximum results in the shortest possible time. Their work involves making multiple decisions of vastly different kinds and consequences in a rapid succession, demanding speed-accuracy trade-offs [5] due to information overload, and possibly materializing cognitive biases as a consequence of taking mental shortcuts. This calls for *decision support systems* (DSS) for managers to enable faster and more accurate strategic decisions through automated and comprehensive analysis of known information. However, traditional DSS and *business intelligence* (BI) solutions such as Tableau or PowerBI typically require expert-crafted analysis pipelines tailored to the particular type of decision or fail to generalize across the large variety of real-world strategic decisions. This makes existing tools suitable for supporting reoccurring operational decisions or for performing selected analyses that feed into a strategic decision but leaves these tools ineffective as end-to-end DSS for complex and unique strategic decisions. Motivated by the recent success stories surrounding *large language models* (LLMs) [2, 3], we present StrategicAI, a novel DSS for strategic decision-making powered by retrieval-augmented LLM-based agents. StrategicAI requires no task-specific fine-tuning of models and is usable even without technical expertise.

StrategicAI is centered around the concept of *logic trees*—often used as *issue trees* or *hypothesis trees* by management consulting firms such as McKinsey or BCG [4, 6]—for structuring complex problems and exploring solutions. StrategicAI understands decision-making as a **problem-solving process** that identifies the root cause of a problem, explores possible solutions, and then decides which solution is the best response to the problem.

2 System Description

In StrategicAI, every problem-solving process is a **task**. Users can start a task by describing a business problem, goal, or hypothesis in natural language. StrategicAI then guides users through a three-phase process:

1. **Problem analysis** breaks down a problem into possible causes and estimates the likelihood of each cause, conditional on all known information. This phase ends when one or more causes are selected as root causes.
2. **Solution exploration** identifies possible solutions to the problem by hierarchically structuring the solution space into discrete choices and estimating each solution's likelihood to effectively solve the problem. This phase ends when the user or AI selects one or more solutions for validation.
3. **Solution validation** takes one solution hypothesis and decomposes it into its underlying assumptions, estimating the likelihood that each assumption is true based on all known information (see Figure 1). This phase ends when the user or AI selects one or more verified solutions as the final solutions.

Users can optionally skip phases 1-2 if they start directly with a task describing a goal or hypothesis rather than a problem.

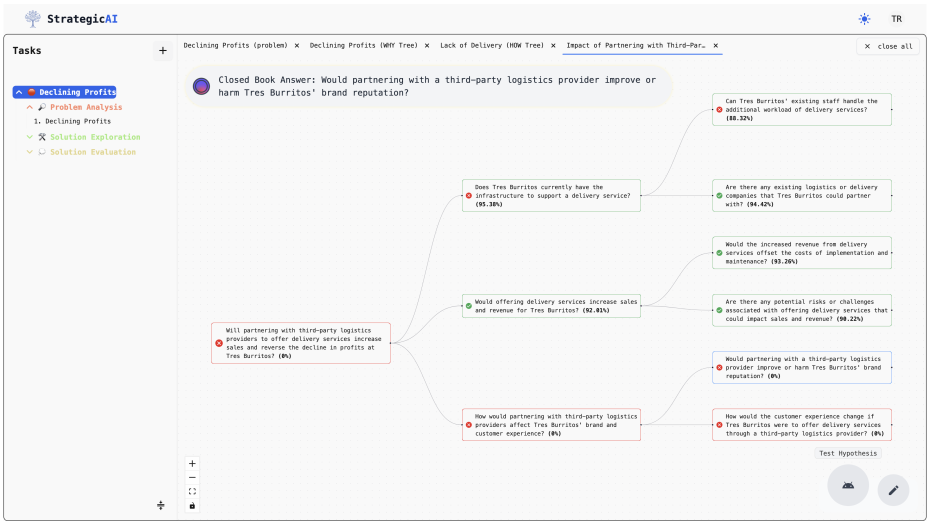

Fig. 1. The screenshot shows a tree built during the solution validation phase. The tree breaks down a solution hypothesis into the sub-hypotheses that need to be true for this solution to effectively solve the problem. StrategyAI is currently performing a hypothesis test by conducting a post-order traversal of the tree from its leafs to the root, estimating the likelihood of each sub-hypothesis based on evidence from the task context, uploaded files, online sources, and the likelihood of the descendant hypotheses.

During all three phases, users build logic trees that structure the underlying problem, solution, or assumption spaces. In every tree, StrategicAI applies an evidence-based search process, estimating the likelihood of different tree nodes based on known facts. StrategicAI maintains an **organization profile** with key facts about the organization and a **fact database** with all known information about a task, which it extracts from the user's task description, the organization profile, files uploaded by the user, and relevant online sources found through search engines. Every tree node comes with a natural-language **explanation**, a **trace to sources** containing facts relevant to that node, and with a **confidence score** from 0 to 100% that indicates the likelihood of that node. Inspired by *Probabilistic Tree-of-Thoughts* [1], confidence scores are derived from the LLM's output logits after including known information as context in the prompts:

$$\text{Confidence} = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log p_i\right),$$

where p_i is the probability of the i-th output token in the LLM's response and n is the total number of tokens in the response.

Users of StrategicAI can alternate between different execution flows of the problem-solving workflow, including manual human-only, human-AI collaborative, and fully autonomous AI-only execution (which we call **AutoRun**). The AutoRun coordinates several LLM-based agents as illustrated in Figure 2.

StrategicAI also offers optional AI-based functionalities for reoccuring user needs. These include a **potential candidates** function that uses *self-consistency* [7] to narrow down potential root causes or effective solutions, a **node verification** to verify and enrich individual nodes with information from selected sources, an **integrated chatbot** which has access to everything within the task and can make recommendations or execute certain actions, and a **textual summary** of the entire problem-solving process and results that can be exported as a PDF or Word document and shared with other stakeholders.

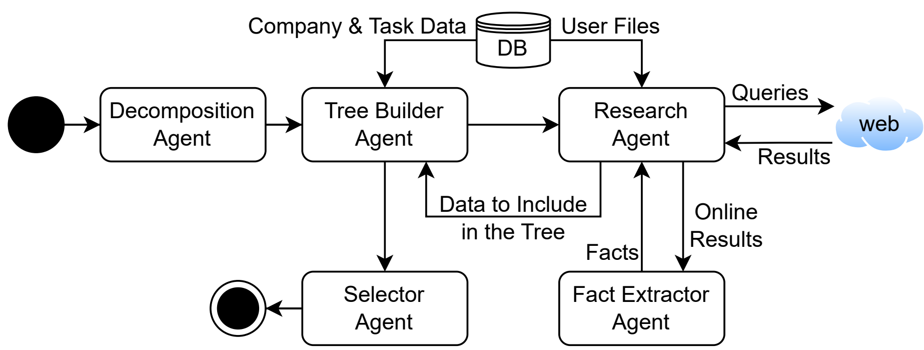

Fig. 2. When the user starts an AutoRun during the problem analysis, StrategyAI coordinates multiple LLM-based agents to decompose the problem into a tree (**Decomposition Agent**), build and refine the tree (**Tree Builder Agent**), find relevant evidence (**Research Agent**), extract facts from found sources (**Fact Extractor Agent**), as well as select the most likely root causes (**Selector Agent**).

3 System Evaluation

To evaluate StrategicAI, we compiled a dataset of six business case studies from well-known business schools[3] and recruited a group of six human testers aged 24-34 with different educational backgrounds and bachelor's degrees or higher. We annotated all case studies with ground-truth root causes and solutions and let each tester solve one case study using StrategicAI. We also let StrategicAI's AutoRun and ChatGPT solve each case study. All tests used a GPT-4o LLM.

Results can be seen in Table 1. In the problem analysis, the human test takers using StrategicAI achieved an average F1 score of 84% (vs. AutoRun's 71% and ChatGPT's 60%). In the solution exploration and validation, the human test takers using StrategicAI achieved 65% (vs. AutoRun's 52% and ChatGPT's

[3] Further documentation of the included case studies can be found in our GitHub repository https://github.com/PortgasXDXMajd/StrategicAI.

Phase	Method	Mode	Precision	Recall	F1
Problem analysis	ChatGPT	AI-only	43%	**100%**	60%
	StrategicAI	AI-only (AutoRun)	59%	94%	71%
	StrategicAI	Human-AI collaboration	**79%**	92%	**84%**
Solution exploration & validation	ChatGPT	AI-only	40%	76%	50%
	StrategicAI	AI-only (AutoRun)	40%	**83%**	52%
	StrategicAI	Human-AI collaboration	**65%**	69%	**65%**

Table 1. Accuracy results from the problem analysis and solution exploration & validation phases observed for ChatGPT (baseline), StrategicAI's AI-only execution (AutoRun), and human-AI collaborative execution of StrategicAI. Best results in **bold**.

50%). In a separate anonymous survey, all test takers expressed a high satisfaction with the tool and stated that they find StrategicAI's main features either somewhat useful or very useful. These findings suggest that StrategicAI is effective at supporting its users in strategic problem-solving and decision-making. We plan to further extend the functionality of StrategicAI and evaluate it under additional realistic scenarios.

Ethical Considerations Recommendations by AI-based tools may be biased and inaccurate. The tools should never be used to make fully automated decisions about humans. Recommendations should always be thoroughly checked by an expert.

Disclosure of Interests. Majd Alkayyal works as a system engineer at an accounting firm. Simon Malberg works as a management consultant at a consulting firm. Both firms were not involved in this research project and had no influence over the outcomes.

References

1. Cao, S., Zhang, J., Shi, J., Lv, X., Yao, Z., Tian, Q., Hou, L., Li, J.: Probabilistic tree-of-thought reasoning for answering knowledge-intensive complex questions. In: Findings of the Association for Computational Linguistics: EMNLP 2023. pp. 12541–12560. Association for Computational Linguistics, Singapore (Dec 2023)
2. Csaszar, F.A., Ketkar, H., Kim, H.: Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors. Strategy Science **9**(4), 322–345 (2024)
3. Doshi, A.R., Bell, J.J., Mirzayev, E., Vanneste, B.S.: Generative artificial intelligence and evaluating strategic decisions. Strategic Management Journal **46**(3), 583–610 (2025)
4. Garrette, B., Phelps, C., Sibony, O.: Cracked it!: How to solve big problems and sell solutions like top strategy consultants. Springer (2018)
5. Heitz, R.P.: The speed-accuracy tradeoff: history, physiology, methodology, and behavior. Frontiers in neuroscience **8**, 150 (2014)
6. Minto, B.: The pyramid principle: logic in writing and thinking. Pearson Education (2009)
7. Wang, X., Wei, J., Schuurmans, D., Le, Q., Chi, E., Narang, S., Chowdhery, A., Zhou, D.: Self-consistency improves chain of thought reasoning in language models. arXiv preprint arXiv:2203.11171 (2022)