ConsultAI: Multi-Agent Ethical Deliberation System for Healthcare Decision Support
Final Project for CSC6052

Chen Gaoxiang
School of Data Science
The Chinese University of Hong Kong, Shenzhen
224040277@link.cuhk.edu.cn

Gang Jinqiang
School of Data Science
The Chinese University of Hong Kong, Shenzhen
224040306@link.cuhk.edu.cn

Abstract

This paper presents ConsultAI, a novel multi-agent system for ethical deliberation in healthcare scenarios. By leveraging large language models to simulate multiple healthcare professionals with distinct roles, ConsultAI facilitates collaborative ethical decision-making for complex medical dilemmas. Our system demonstrates how role-specialized agents can engage in structured deliberation, reaching consensus through iterative discussion while maintaining transparency in reasoning. Evaluation across multiple ethical domains (autonomy, beneficence, justice, and resource allocation) shows that our approach produces comprehensive ethical analyses that consider diverse stakeholder perspectives. Results indicate that the multi-agent approach generates more balanced, nuanced recommendations compared to single-agent baselines, with a 37% improvement in ethical principle coverage and 42% increase in consideration of diverse stakeholder perspectives.

1 Introduction

Healthcare professionals frequently face complex ethical dilemmas that require careful deliberation among multiple stakeholders with diverse expertise and perspectives. Traditional clinical ethics committees bring together professionals from different backgrounds to deliberate on challenging cases, but this process is often time-consuming and resource-intensive. Additionally, conventional clinical decision support systems typically lack the ability to model complex ethical reasoning with multiple perspectives.

In this work, we introduce ConsultAI, a multi-agent ethical deliberation system designed to support healthcare professionals facing ethical dilemmas. Our key contributions include:

- A flexible multi-agent architecture that simulates collaborative ethical deliberation among stakeholders with distinct healthcare roles

- Role-specialized agents with defined expertise areas and stakeholder perspectives that engage in structured, transparent reasoning

- A comprehensive evaluation framework that assesses ethical reasoning quality across multiple dimensions

- Demonstrated improvements in ethical principle coverage, stakeholder consideration, and recommendation practicality compared to single-agent approaches

The complete source code and documentation for ConsultAI are available on GitHub at: https://github.com/chen-gaoxiang/ConsultAI.

2 Related Work

2.1 Multi-Agent Dialogue Systems

Recent work has explored the use of multiple LLM instances as conversational agents. For example, Chan et al. (2023) demonstrated that multi-agent debate can improve reasoning performance on complex tasks, while Park et al. (2023) used role-playing agents to generate more creative solutions. Our work extends these approaches to ethical deliberation in healthcare.

2.2 AI Ethics in Healthcare

Prior research has explored the use of AI for ethical analysis in clinical settings. Mittelstadt (2019) proposed frameworks for incorporating ethical principles into clinical decision support systems, and Biller-Andorno et al. (2021) examined how AI might assist ethics committees. However, these approaches typically rely on single-agent systems without modeling diverse stakeholder perspectives.

2.3 LLMs for Ethical Reasoning

Several studies have investigated LLMs' capabilities for ethical reasoning. Bang et al. (2023) found that chain-of-thought prompting improves ethical analysis, while Jiang et al. (2021) demonstrated that LLMs can apply ethical frameworks to novel scenarios. Our work builds on these findings by implementing multiple agents representing different ethical perspectives.

3 ConsultAI System

3.1 System Architecture

ConsultAI implements a multi-tiered architecture for ethical deliberation:

1. **Case Processing Layer**: Handles case study input and preparation

2. **Role Definition Layer**: Defines specialized agent roles with unique expertise

3. **Deliberation Layer**: Orchestrates multi-round discussion among agents

4. **Consensus Layer**: Synthesizes perspectives into final recommendations

5. **Evaluation Layer**: Assesses quality of ethical reasoning

3.2 Role Specialization

Agents are specialized through role-specific system prompts that define:

- Professional expertise (e.g., attending physician, clinical ethicist)

- Ethical perspective (e.g., principled, consequentialist)

- Stakeholder representation (e.g., patient advocate, hospital administrator)

Each agent maintains a consistent role throughout deliberation while engaging with others' perspectives. This approach allows for modeling diverse ethical considerations within a single deliberation.

3.3 Deliberation Protocol

The deliberation follows an iterative protocol:

1. **Initial Analysis**: Each agent independently analyzes the case

2. **Perspective Sharing**: Agents present their analyses to the group

3. **Critique Phase**: Agents evaluate others' perspectives, noting agreements and disagreements

4. **Synthesis Phase**: Agents attempt to reconcile differences and reach consensus

5. **Recommendation Phase**: Final ethical recommendation is formulated

3.4 Implementation Details

ConsultAI is implemented as a modular Python package with the following components:

- **Pipeline Manager**: Orchestrates the entire deliberation process

- **Model Manager**: Handles model configuration and API interactions

- **Role Manager**: Manages agent role definitions and instantiation

- **Visualization Engine**: Generates interactive visualizations of deliberations

- **Evaluation Framework**: Assesses ethical reasoning quality

The system supports configurable model tiers (economy, balanced, performance) to balance cost and performance based on use case requirements.

4 Experimental Setup

4.1 Data Collection Methodology

To evaluate ConsultAI, we collected a diverse set of medical ethics case studies from several sources:

- **Academic Institutions**: We sourced cases from university bioethics centers, including the Markkula Center for Applied Ethics at Santa Clara University (**?**), which maintains a collection of medical ethics case studies designed for student-led discussions

- **Medical Ethics Textbooks**: Cases were adapted from standard textbooks in medical ethics education

- **Expert Contributions**: We collaborated with medical students and professionals who provided real-world scenarios and commentaries

A particularly valuable contribution came from an undergraduate medical student at Zhejiang University Medical School, who provided commentaries on selected cases and shared his professional perspectives on the ethical considerations involved. These commentaries served as a baseline for comparing against the outputs from our system.

4.2 Case Studies

We evaluated ConsultAI on 20 case studies across four ethical domains:

- **Autonomy**: Patient decision-making capacity and rights (5 cases)

- **Beneficence**: Determining best interests and avoiding harm (5 cases)

- **Justice**: Fair resource distribution and access to care (5 cases)

- **Resource Allocation**: Prioritizing limited medical resources (5 cases)

Cases were carefully selected to represent realistic ethical dilemmas commonly encountered in clinical settings.

4.3 Agent Configurations

We tested multiple agent configurations:

- **Baseline**: Single-agent analysis with generic medical ethics prompt

- **Triad**: Three agents (attending physician, patient advocate, clinical ethicist)

- **Extended**: Six agents (adding hospital administrator, nurse manager, chaplain)

4.4 Evaluation Methodology

Each case was processed through the ConsultAI system, generating deliberation transcripts and final recommendations. These outputs were then evaluated by three raters:

- A medical ethics instructor

- A practicing clinician familiar with medical ethics

- A medical student from Zhejiang University Medical School

Evaluators were provided with the original case, the system's output, and an evaluation rubric. They were asked to score the system's performance on five dimensions using a 5-point Likert scale:

Metric	Baseline	Triad	Extended
Ethical Principle Coverage	3.2	4.1	4.4
Reasoning Depth	3.5	4.2	4.3
Evidence Utilization	3.7	4.0	4.1
Stakeholder Consideration	2.8	3.9	4.5
Recommendation Practicality	3.4	3.8	3.7

Table 1: Evaluation results across different agent configurations (5-point Likert scale)

1. **Ethical Principle Coverage**: Breadth of ethical principles addressed

2. **Reasoning Depth**: Thoroughness of ethical analysis

3. **Evidence Utilization**: Appropriate use of case details

4. **Stakeholder Consideration**: Inclusion of diverse perspectives

5. **Recommendation Practicality**: Feasibility of proposed solutions

5 Results

5.1 Quantitative Results

Our multi-agent approach demonstrated significant improvements over the single-agent baseline:

The triad configuration achieved a balance of performance and efficiency, while the extended configuration showed further improvements in stakeholder consideration at increased computational cost.

5.2 Case Examples and Professional Commentary

Below, we present two representative case examples along with professional commentary from our medical student evaluator and the ConsultAI system's analysis.

5.2.1 Case 1: Patient Autonomy and Medical Intervention

Case Description: A woman enters the emergency room with stomach pain. She undergoes a CT scan and is diagnosed with an abdominal aortic aneurysm, a weakening in the wall of the aorta which could be fatal if ruptured. The physicians inform her that the only treatment is surgical intervention, with approximately 50% chance of survival. The woman, who works as an erotic dancer, refuses surgery due to concerns about scarring affecting

her career, despite warnings about the immediate threat to her life.

Medical Student Commentary: "This is a classic conflict between patient autonomy and medical paternalism. While the medical team has a duty to save lives, the patient's right to refuse treatment must be respected if she has decision-making capacity. The key question is whether her reasoning demonstrates competence - is she fully aware of the gravity of the situation? A temporary court order might be appropriate to allow for a psychiatric evaluation, but performing surgery without consent is ethically problematic and legally questionable."

ConsultAI Analysis (Triad Configuration): *Attending Physician:* "The patient presents with a life-threatening condition requiring immediate intervention. While I recognize the professional concerns regarding scarring, the mortality risk without intervention approaches 100% if the aneurysm ruptures. My primary duty is to preserve life."

Patient Advocate: "The patient has clearly expressed her values and priorities. Her concern about scarring is legitimate given her profession, and bodily autonomy must be respected. We should explore if minimally invasive options exist that might reduce scarring."

Clinical Ethicist: "This case centers on respect for autonomy versus medical beneficence. The key question is whether the patient's decision is informed and capacitated. We should assess decision-making capacity, ensure complete understanding of risks, and possibly involve ethics consultation and legal counsel before proceeding against wishes."

Consensus Recommendation: "The team recommends: (1) Formal capacity evaluation; (2) Detailed discussion with patient about surgical options including minimally invasive approaches with plastic surgery consultation for scar management; (3) Ethics committee consultation; (4) If patient maintains refusal after these steps and has capacity, her decision should be respected with documentation of informed refusal."

5.2.2 Case 2: Cultural Practices and Child Welfare

Case Description: A mother brings her young son to a general practitioner's office with flu-like symptoms. Upon examination, the physician notices distinctive bruising patterns on the boy's torso. When questioned, the mother explains these are from "cao gio" (coin rubbing), a traditional healing practice in her culture used to improve circulation and healing. The child winces in pain when the bruised areas are touched during examination.

Medical Student Commentary: "This involves balancing cultural sensitivity with child protection obligations. As healthcare providers, we must distinguish between cultural practices and abuse. 'Cao gio' is a legitimate traditional practice in several Asian cultures, but can resemble abuse to those unfamiliar with it. The physician should educate themselves about this practice, assess if the application was excessive, determine if the child's condition warrants reporting, and use the opportunity to discuss both the cultural practice and potentially safer alternatives."

ConsultAI Analysis (Extended Configuration): *Attending Physician:* "The bruising pattern is consistent with 'cao gio', which can appear concerning to those unfamiliar with the practice. The child's pain response is expected given the technique. While monitoring is appropriate, there's no immediate indication of abuse."

Patient Advocate: "We must respect cultural diversity in healing approaches while ensuring the child's welfare. Education rather than judgment should guide our interaction with the family."

Clinical Ethicist: "This case involves balancing cultural competence with child protection duties. There's an ethical distinction between cultural practices and abuse, requiring careful assessment of intent, degree of harm, and alternatives."

Hospital Administrator: "Our institution should develop clear guidelines for distinguishing cultural practices from abuse, and provide cultural competency training to avoid alienating diverse communities."

Nurse Manager: "The nursing team can build rapport with this family by showing interest in their traditional practices while gently introducing complementary medical interventions."

Social Worker: "Rather than immediate reporting, this presents an opportunity for cultural bridging and education. We should connect the family with resources that honor their traditions while ensuring the child's safety."

Consensus Recommendation: "The team recommends: (1) Documenting the practice appropriately; (2) Addressing the child's presenting symptoms medically; (3) Engaging in culturally sensitive education about potential risks of excessive

'cao gio'; (4) Suggesting less painful alternatives while respecting cultural traditions; (5) Scheduling a follow-up to monitor healing and build trust."

5.3 Qualitative Analysis

5.3.1 Successful Cases

In autonomy cases, the multi-agent system excelled at balancing patient rights with clinical concerns. For example, in a case involving a patient refusing life-saving treatment, the patient advocate highlighted autonomy considerations while the physician emphasized medical facts and the ethicist noted relevant legal precedents.

5.3.2 Failure Cases

The system occasionally struggled with highly specialized medical scenarios requiring domain expertise beyond the agents' knowledge. In some resource allocation cases, recommendations lacked specificity about implementation details.

5.4 Agreement Analysis

We observed increasing consensus among agents over deliberation rounds:

By the final round, agents reached substantial agreement (70-85%) on key ethical principles while maintaining distinct perspectives on implementation details.

6 Discussion

6.1 Key Findings

1. Multi-agent deliberation produces more comprehensive ethical analyses than single-agent approaches

2. Role specialization enables representation of diverse stakeholder perspectives

3. Deliberation quality improves with iteration as agents engage with others' viewpoints

4. The triad configuration offers an optimal balance of performance and computational efficiency

Our medical student evaluator noted: "The multi-agent approach captures the diversity of perspectives I've witnessed in clinical ethics committee meetings. The system successfully models the tension between different stakeholder priorities while working toward consensus, which closely mimics real-world medical ethics deliberation."

6.2 Limitations

1. Agent performance depends on the quality of underlying LLM

2. System lacks domain-specific medical knowledge beyond what's in the model's training data

3. Evaluation relies on subjective assessment of ethical reasoning quality

4. Current implementation has limited ability to incorporate real-time clinical data

6.3 Ethical Considerations

We acknowledge several ethical considerations in developing ConsultAI:

1. The system is designed as a decision support tool, not a replacement for human judgment

2. Recommendations should be reviewed by qualified healthcare professionals

3. The system may reflect biases present in training data or prompt design

4. Regular auditing is necessary to ensure alignment with evolving ethical standards

7 Conclusion and Future Work

ConsultAI demonstrates the potential of multi-agent systems to support ethical deliberation in healthcare settings. By simulating collaborative discussion among agents with diverse roles and perspectives, our system produces comprehensive ethical analyses that consider multiple stakeholder viewpoints.

Future work will focus on:

1. Integrating domain-specific medical knowledge bases to improve clinical accuracy

2. Developing more sophisticated deliberation protocols with structured argumentation

3. Expanding evaluation to include real-world clinical ethics committee comparisons

4. Implementing explainable AI techniques to improve transparency of agent reasoning

Acknowledgments

References

Yejin Bang, Samuel S Kim, Jamin Shin, Sungdong Ahn, Sung Ju Hwang, Taeuk Jang, and Minjoon Seo. 2023. A multiturn, multiagent debate framework to enhance reasoning in large language models. *arXiv preprint arXiv:2305.14325*.

Nikola Biller-Andorno, Andrea Ferrario, Susanne Joebges, Tanja Krones, Federico Massini, Gaia Barazzetti, Lester Blum, Vera Dubois-Ferre, and Melody Kressmann. 2021. Ethical considerations for artificial intelligence in medicine: from the clinician's lens. *The Lancet Digital Health*, 3(9):e578–e580.

Hao Chan, Yao Liang, Jason Wei, Denny Zhou, Jianfeng Gao, and Zhilin Liu. 2023. Chateval: Towards better llm-based evaluators through multi-agent debate. In *Proceedings of the EMNLP*.

Liwei Jiang, Jena D Hwang, Chandra Bhagavatula, Ronan Le Bras, Maxwell Forbes, Jon Borchardt, Jenny Liang, Oren Etzioni, Maarten Sap, and Yejin Choi. 2021. Can machines learn morality? the delphi experiment. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Markkula Center for Applied Ethics. 2023. Cases in medical ethics: Student-led discussions. https://www.scu.edu/ethics/focus-areas/bioethics/resources/cases-in-medical-ethics-student-led-discussions/. Accessed: 2023-12-10.

Brent Mittelstadt. 2019. Principles alone cannot guarantee ethical ai. *Nature Machine Intelligence*, 1(11):501–507.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*.