

Toward an Integrated Decision-Making Framework for Optimized Stroke Diagnosis with DSA and Treatment under Uncertainty

Nur Ahmad Khatim
Informatics Engineering
Institut Teknologi Sepuluh
Nopember (ITS)
Surabaya, Indonesia

Azmul Asmar Irfan
School of Medicine
Universitas Islam Negeri (UIN)
Syarif Hidayatullah
Jakarta, Indonesia

Amaliya Mata'ul
School of Medicine
Universitas Islam Negeri (UIN)
Syarif Hidayatullah
Jakarta, Indonesia

Mansur M. Arief
Stanford Intelligent Systems Lab
Stanford University
Stanford, USA
mansur.arief@stanford.edu

Abstract—This study addresses the challenge of stroke diagnosis and treatment under uncertainty, a critical issue given the rapid progression and severe consequences of stroke conditions such as aneurysms, arteriovenous malformations (AVM), and occlusions. Current diagnostic methods, including Digital Subtraction Angiography (DSA), face limitations due to high costs and its invasive nature. To overcome these challenges, we propose a novel approach using a Partially Observable Markov Decision Process (POMDP) framework. Our model integrates advanced diagnostic tools and treatment approaches with a decision-making algorithm that accounts for the inherent uncertainties in stroke diagnosis. Our approach combines noisy observations from CT scans, Siriraj scores, and DSA reports to inform the subsequent treatment options. We utilize the online solver DESPOT, which employs tree-search methods and particle filters, to simulate potential future scenarios and guide our strategies. The results indicate that our POMDP framework balances diagnostic and treatment objectives, striking a tradeoff between the need for precise stroke identification via invasive procedures like DSA and the constraints of limited healthcare resources that necessitate more cost-effective strategies, such as in-hospital or at-home observation, by relying only on simulation rollouts and not imposing any prior knowledge. Our study offers a significant contribution by presenting a systematic framework that optimally integrates diagnostic and treatment processes for stroke and accounting for various uncertainties, thereby improving care and outcomes in stroke management.

Keywords—healthcare, stroke, DSA, state uncertainty, POMDP

I. INTRODUCTION

Accurate diagnosis and treatment of stroke are crucial due to the condition's rapid progression and potentially severe consequences. Stroke, often resulting from conditions such as aneurysms, arteriovenous malformations (AVM), or occlusions, necessitates swift and precise medical intervention. Approximately 87% of all strokes are ischemic, with an age-standardized incidence of cerebral infarction at 153.2 per 100,000 person-years. Additionally, the age-standardized incidence of hemorrhagic stroke is 46.2 per 100,000 person-years. Diseases associated with hemorrhagic stroke, like intracerebral aneurysm, have an incidence rate of 12.6 per 100,000 person-years, while AVM incidence stands at 2.4 per 100,000 person-years [1, 2].

The current approach to these conditions includes diagnostic imaging, such as Digital Subtraction Angiography (DSA), coupled with individualized treatment plans. DSA is a key imaging modality for identifying and confirming aberrations in cerebral blood vessels in various conditions, including aneurysms and arteriovenous malformations. It is considered the gold standard for evaluating cerebral vascular disease and is crucial for treatment planning in patients with suspected vascular abnormalities [3].

In addition, several stroke assessment tools have been developed for rapid diagnosis and triage. These tools facilitate the quick identification of acute ischemic stroke and are validated for clinical use. However, they come with significant costs, and false positives can place additional strain on the healthcare system. An evaluation of 13 clinical measures for identifying occlusion in over 1,000 individuals revealed high rates of false negatives and positives [4]. Furthermore, a meta-analysis indicated that no scale could accurately and consistently identify occlusion [5], underscoring the challenges in stroke diagnosis. This is mostly because many aspects of stroke may not be directly observable, such as microvascular changes or silent strokes, which lack apparent symptoms [6].

In this complex and highly uncertainty scenario, healthcare professionals often must make critical decisions based on partial information, compounded by the inherent history and physical examination noise in diagnostic processes. This often leads to uncertainties in diagnosis and treatment, where every decision carries significant weight. The consequences of incorrect decisions extend beyond immediate healthcare costs; they profoundly impact patient quality of life, and long-term health outcomes [7]. Under such a setting, an intelligent decision-making framework that we propose offers a transformative solution. By integrating various diagnostic tools and standardized treatment protocol with algorithms for optimizing decisions, we can incorporate the risks associated with partial information and observation noise into an uncertainty-aware decision. This approach echoes earlier works that have leveraged systematic algorithms and optimization models to address similar challenges and incorporates decision making under uncertainty, such as Markov Decision Processes (MDPs) or Partially Observable MDPs (POMDPs) [8-10].

MDPs and POMDPs are particularly adept at modeling decision-making scenarios where outcomes are partly random and partly under the control of a decision-maker [11]. MDPs, for instance, have been effectively used in scenarios requiring full state observability, assuming a 100% accurate diagnosis a priori. The strength of these models, especially in offline algorithms, lies in their ability to solve to optimality given a sufficient compute time. However, they are often constrained by the curse of dimensionality (i.e. being exponentially slower as the problem size increases) and full observability requirement. POMDPs extend the capabilities of MDPs by accounting for partial observability, making them more suitable for medical scenarios where diagnosis and patient responses to treatment can be uncertain and variable. This is especially important when it comes to stroke care, where quick and precise decision-making is critical. POMDP assists in reaching a compromise between the time and resource limitations and the requirement for comprehensive diagnostics [12].

POMDP is often solved using online algorithms. A key feature of online algorithms is their use of tree-search methods combined with particle filters [13], which enable the simulation of future trajectories. This approach simulates future observations and trajectories as a tree when a certain action is taken and utilizes these simulated rollouts to make informed decisions on the possible future progression of the states (i.e. aneurysm, AVM, or occlusion cases in our context). This procedure is often illustrated in tree diagrams (see Fig. 1).

Despite their potential to significantly enhance the accuracy and efficiency of medical interventions in complex cases, there is a notable lack of studies exploring the use of online algorithms in the diagnosis and treatment of aneurysm, AVM, or occlusion stroke conditions. In this study, we propose a POMDP formulation that allows us to consider both diagnosis and treatment simultaneously. More importantly, the important feature that we aim to capture is the limited information during the diagnosis stage and the impact of imperfect diagnosis on the optimality of the treatment. To study this, we build a model that does not assume full information of the stroke type, but only noisy observations in the form of CT scan reading, Siriraj score, or DSA report, with varying sensitivity in identifying the occurrence of aneurysm (96.7%), AVM (100%), or occlusions (96%) [14-16]. We then compare the performance of various baselines including random policy, expert policy, and optimized policies using an online solver DESPOT [17].

In summary, our contribution is twofold. First, we propose a POMDP formulation for integrated stroke diagnosis and treatment for long-horizon problems. Second, we provide a benchmark of our POMDP policy with expert policies under settings with high uncertainty to mimic real-world scenarios and stress-test the usefulness of these algorithms. The code for our experiments is available at: www.github.com/inteligensi/DSAPOMDPs.jl.

The rest of this paper is organized as follows. In Section II, we review works that address stroke diagnosis and introduce POMDP framework. In Section III, we present our model and approach to solving it using online solver. We then provide details of our experiments in Section IV. Finally, we discuss our findings in Section V and conclude in Section VI.

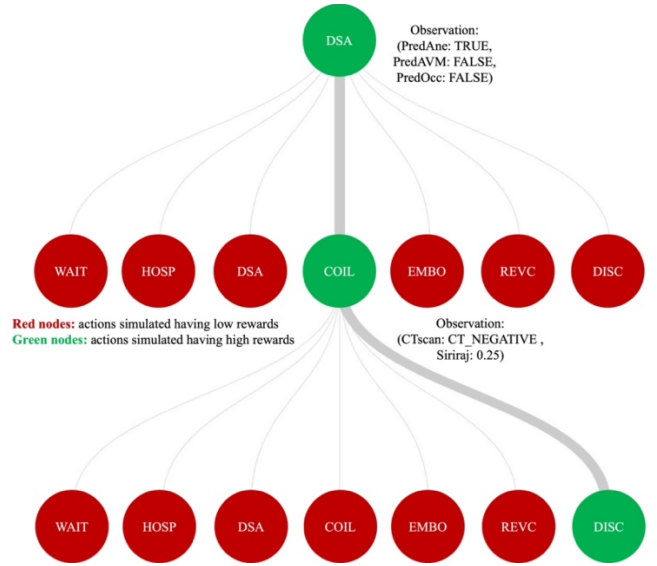


Fig. 1. Illustration of simulated policy rollouts in tree search

II. RELATED WORK

In this section, we briefly summarize stroke diagnosis and treatment, POMDP, and its uses for assisting medical decision making.

A. Stroke Diagnosis and Treatment

Swift decision-making is crucial in stroke management, with significant implications for patient outcomes. Rapidly identifying symptoms and determining their cause is essential for effective treatment and minimizing long-term effects. Medical imaging, especially CT scans and DSA, is vital in this process. CT scans are key for early stroke diagnosis, helping to identify ischemic or hemorrhagic areas by visualizing changes in brain tissue density. DSA, on the other hand, is the gold standard for evaluating and treating neurovascular diseases like AVM and aneurysms in hemorrhagic strokes [3], though highly invasive and costly in nature and thus should only be ordered only when necessary.

These diagnostic procedures are highly important to guide effective treatments. For instance, approaching ischemic stroke cases with quick revascularization, such as thrombolysis, has been associated with better outcomes. AVM treatments range from endovascular embolization, a minimally invasive procedure, to microsurgical resection [18]. Intracranial aneurysms may require transluminal coiling, a procedure using a catheter to prevent blood flow into the aneurysm [19]. These treatments, however, often have a limited window of opportunity, and delays can impact patient eligibility and prognosis [20]. Thus, integrating appropriate diagnostics and treatments is essential for resource optimization, expertise application, and patient-specific care, addressing the inherent uncertainties of stroke management.

Addressing these uncertainties is challenging. Stroke symptom variability, patient comorbidities, and individual factors often lead to incomplete and ambiguous information. Diagnostic noise, such as imaging artifacts, complicates

symptom interpretation [21], making accurate and timely diagnosis difficult. Healthcare professionals must often make critical decisions quickly, considering available resources. To mitigate this, patients are frequently observed in hospitals to accumulate lab measurements and physical observations. Repeated close observation of physical and physiological symptoms can build diagnostic confidence, albeit time-consuming. The Siriraj score [22] is commonly used to summarize these observations, guiding treatment or follow-up procedures. Scores below -1 typically indicate ischemic stroke, while scores above +1 suggest hemorrhagic stroke. Inconclusive results, around 0, often lead to more advanced diagnostics like DSA. Combining these procedures, healthcare professionals must navigate uncertainties to make critical decisions, often life-altering, and provide the best possible medical advice for patient treatment.

B. Partially Observable Markov Decision Process (POMDP)

POMDPs provide a robust mathematical framework for modeling decision-making problems under uncertainty and partial observability. In a POMDP, an agent interacts with an environment where the true state is not directly observable. The goal is to make decisions that maximize the expected cumulative reward over time [23]. The key challenge is to maintain a belief state—a probability distribution over possible states—based on observed information and to optimize decisions amidst uncertainty [24].

Given the inherent complexity of POMDPs, online solvers have become essential for real-time decision-making. These solvers adapt dynamically to changing environments by continuously updating decision strategies based on observed information [25]. Online planning involves simultaneous planning and execution. At each time step, it conducts local planning, selecting the best action based on the current belief and employing a lookahead search within the vicinity of the existing belief state. The chosen action is then executed promptly. This approach offers computational benefits: it is less complex to search for an optimal action within a single belief than for all beliefs, as required in offline policy computation. Additionally, it leverages local structures, reducing the search space size. However, the online approach is limited by the available planning time.

A notable online solver for POMDPs, particularly in decentralized multi-agent settings, is the Decentralized POMDP Solver with Online Planning (DESPOT) [17]. This algorithm adapts the Monte Carlo Tree Search [26] framework to decentralized scenarios, where multiple agents with partial observability must coordinate actions to achieve a common goal. DESPOT maintains belief states for each agent using particle filters and employs an adaptive tree depth strategy to optimize computational resources. By using rollout policies and approximating the value function through MCTS, DESPOT excels in real-time decision-making under uncertainty. DESPOT's decentralized structure and adaptability make it effective for large-scale problems where computational efficiency is crucial. The solver has been applied in research scenarios, such as addressing POMDP challenges in Highway Driving [27].

C. POMDP for Assisting Medical Decision Making

Decision-making challenges in stroke management often arise from various factors, including subtle symptoms and the variability of treatment options within a time-sensitive framework. Implementing POMDP in the context of stroke events offers potential benefits, particularly in optimizing diagnostic and treatment decisions. While specific studies on this topic may be limited, the application of POMDP in healthcare has shown promise.

In the context of diagnosis, an earlier study explored the timing of mammograms for suspected breast cancer cases. The researchers developed a finite-horizon POMDP model to determine personalized mammography screening policies based on a woman's risk factors and past screening results. The model's unobservable states represented the patient's cancer stage, ranging from noninvasive to invasive, under-treatment, or no cancer. The POMDP actions were "wait" and "mammography," with outcomes including positive or negative mammogram results, or self-detection results. Positive findings led to further diagnostic steps like biopsies. The study revealed a control-limit policy based on the risk of invasive and noninvasive cancers, influenced by the patient's screening history [28].

In treatment contexts, another study developed a framework for personalizing anticoagulant therapy in patients on warfarin. The POMDP model was used to update beliefs and assess the optimality of myopic policies in MDP, deriving conditions for their existence and uniqueness. Real-life patient data validated the model, offering practical clinical insights, such as medication initiation duration and patient medication sensitivity [29].

Our study proposes a POMDP formulation for integrated stroke diagnosis and treatment. The goal is to combine diagnosis and treatment decision-making and assess how the model, solved with an online solver, performs under various initial that mimic real world. This approach will be compared with heuristic policies currently in standard practice, such as hospital observation or direct DSA ordering. We hypothesize that our POMDP approach can provide a decision support system, offering healthcare professionals recommendations to navigate partial observation, state uncertainty, and the significant consequences of decision-making in these settings.

III. METHODOLOGY

In this section, we present the formulation of our problem as a POMDP. More generally, a POMDP model is defined by the tuple $(S, A, O, T, Z, R, \gamma)$, where S is the state space, A is the action space, and O is the observation space. For $s, s' \in S$, $a \in A$, $T(s'|s, a)$ is the probability of transitioning into state s' from current state s by taking action a . $Z(o|s, a)$ is the probability of observing observation $o \in O$ given that at state s action a is taken. $R(s, a)$ is the immediate reward obtained when taking action a in state s . Finally, $\gamma \in (0, 1)$ is the discount factor to incorporate how much we care about future rewards.

Specifically, for our problem, the state

$$s \in (isAne, isAVM, isOcc, t) \quad (1)$$

contains the patient's stroke-related condition, including binary variables that indicate the occurrence of aneurysm ($isAne$), AVM ($isAVM$), and occlusion ($isOcc$). We also store as part of the state a non-negative discrete time counter t . The action space A includes both diagnostic, i.e. observe closely in a hospital ($HOSP$) or perform DSA (DSA) and surgical treatment options that a medical team can take for the patient, including coiling ($COIL$), embolization ($EMBO$), or revascularization ($REVC$)

$$A = \{WAIT, HOSP, DSA, COIL, EMBO, REVC, DISC\}. \quad (2)$$

In addition, we also add a discharge action ($DISC$) to signify when the patient has been treated and can be sent home. Note here that the action $WAIT$ signifies the case in which the patient's condition is deemed healthy enough for the time being to wait at home without treatment just yet.

The defining characteristic of POMDP is the partial unobservability of the states. For our problem, we do not have full access to $s_t \in S$, the state at time t , i.e. we do not know $isAne$, $isAVM$, or $isOcc$, but only know t . However, we can observe $o_t \sim Z(\cdot | s_t, a_t) \in O$ that provides some (noisy) lab measurements or observations about the patient that can be used to build a belief model b_t to estimate the underlying value of s_t . It is based on this belief b_t that we can then take actions $a_t = \pi(b_t) \in A$ based on some policy π , after which a reward $r_t = R(s_t, a_t)$ is awarded based on the current state s_t and selected action a_t and the system progresses into the next state $s_{t+1} \sim T(\cdot | s_t, a_t)$. It is worth noting here that the original state $s_0 \sim p_0$ is often treated as a random variable sampled from some initial state distribution p_0 . For our problem, p_0 is an estimate of how the general population encounters stroke conditions and seeks medical help.

To solve the formulated POMDP, a solver seeks an optimal policy π^* that maximizes the sum of discounted rewards (often referred to as utility)

$$\pi^* = \arg \max_{\pi \in \Pi} \mathbb{E}_{s_0 \sim p_0} [\sum_{t=1}^{\infty} \gamma^{t-1} R(s_t, \pi(b_t)) | s_0]. \quad (3)$$

It should be clear from (3) that the complexity of the problem is heavily influenced by the combination of possible values that s_t , a_t , and o_t may take (the size of S , A , and O), as well as the complexity of the reward function R and policy space Π .

To better describe our modeling approach, below is a more detailed description of T (transition model), Z (observation model), and b_t (belief model).

A. Transition Model

The transition model T represents how the patient's state changes over time and incorporates both the effect of the treatment-type actions ($COIL$, $EMBO$, $REVC$) on the state as well as naturalistic randomness of the world (random or unmodeled occurrence of stroke).

For simplicity, we assume each of the treatment-type actions treats one specific type of stroke condition ($COIL$ treats aneurysm, $EMBO$ treats AVM, and $REVC$ treats occlusion) with probability 1, and it does not have any effect on other stroke conditions. We can extend our model to allow treatments affecting multiple conditions. However, it would require more

parameters to estimate a priori without adding much insights into the usefulness of our study significantly. We thus omit such an effect in the current study.

In addition to treatment effects, we also model the plausibility of random (unmodeled) stroke occurrence in healthy patients. To account for this, we model such occurrence events as independent Bernoulli trials at each time step with probability p_{ane} , p_{AVM} , and p_{occ} for aneurysm, AVM, and occlusion, respectively. Finally, the time counter t is updated at every time step, regardless of actions. Table 1 summarizes the state-transitioning events with non-zero probability.

TABLE I. NON-ZERO STATE TRANSITION PROBABILITY

Action	Transition Events		
	Current state	Next state	Probability
<i>COIL</i>	<i>IsAne</i> : <i>TRUE</i>	<i>IsAne</i> : <i>FALSE</i>	1.0
<i>EMBO</i>	<i>IsAVM</i> : <i>TRUE</i>	<i>IsAVM</i> : <i>FALSE</i>	1.0
<i>REVC</i>	<i>IsOcc</i> : <i>TRUE</i>	<i>IsOcc</i> : <i>FALSE</i>	1.0
Any $a \in A$	<i>IsAne</i> : <i>FALSE</i>	<i>IsAne</i> : <i>TRUE</i>	p_{ane}
	<i>IsAVM</i> : <i>FALSE</i>	<i>IsAVM</i> : <i>TRUE</i>	p_{AVM}
	<i>IsOcc</i> : <i>FALSE</i>	<i>IsOcc</i> : <i>TRUE</i>	p_{occ}
	<i>Time</i> : t	<i>Time</i> : $t + 1$	1.0

B. Observation Model

The observation model Z encapsulates the processes of lab measurement, patient observation, and report summarization as part of the treatment process. It assumes that certain symptoms will manifest in lab measurements (such as CT scan readings), patient conditions (summarized as the Siriraj score), or angiography procedures (like DSA), signaling the underlying stroke conditions ($isAne$, $isAVM$, $isOcc$), albeit with some level of noise.

For instance, a CT scan provides imaging results that the medical team interprets to determine the presence of any of the three stroke conditions. However, these conditions may not always be distinguishable solely from the scans. If a patient's condition is ($isAne$: *TRUE*, $isAVM$: *FALSE*, $isOcc$: *FALSE*), the CT scan summary might indicate *CT_POSITIVE* with a probability p_1 and *CT_NEGATIVE* with a probability $1 - p_1$. These probabilities vary based on the patient's specific condition. For example, another patient with ($isAne$: *FALSE*, $isAVM$: *TRUE*, $isOcc$: *FALSE*) would have different probabilities p_2 and $1 - p_2$, where $p_1 \neq p_2$. It is important to note that these probabilities, p_1 and p_2 , as well as the Siriraj score probabilities, are dependent on the actions. Patients waiting at home (*WAIT*) tend to have higher uncertainty in these measures, while those hospitalized (*HOSP*) exhibit lower uncertainty due to constant medical team monitoring.

The Siriraj score, a scalar value ranging from -5 to +5, summarizes the patient's physical and physiological conditions. Scores below -1 typically indicate an ischemic stroke, above 1 suggest a hemorrhagic stroke, and scores in between are considered inconclusive. Our model samples a Siriraj score

based on the conditional probability related to the underlying stroke condition. For instance, if $IsAne$ or $IsAVM$ is $TRUE$, the score is more likely to be above 1. Conversely, if $IsOcc$ is $TRUE$, the score tends to be above -1. In the absence of any stroke conditions, the score usually centers around 0.

Both the CT scan summary and the Siriraj score are assumed to be collected for every patient unless a DSA is ordered. In cases where DSA is performed, instead of observing the CT scan and Siriraj score, we obtain binary predictions ($PredAne, PredAVM, PredOcc$) for the underlying state ($IsAne, IsAVM, IsOcc$) with 98% accuracy. DSA provides reliable estimates but is a costly and invasive procedure, thus reserved for necessary cases, i.e. associated with high cost in the reward function.

C. Belief Model

The belief model b_t represents the probability distribution over the state space S at time t , given the history of actions and observations. Specifically, for our stroke-related conditions, $b_t(s_t)$ encapsulates the estimated probabilities of state s_t based on the noisy observations from CT scans and Siriraj scores. This model is pivotal in scenarios where the actual state s_t is partially observable, allowing us to infer the most probable state of the patient's condition for decision making.

At each decision epoch, the belief model informs the selection of actions $a_t \in A$. The action chosen at time t , $a_t = \pi(b_t)$, is based on the current belief b_t , where π represents the policy mapping beliefs to actions. This process dynamically adapts to the evolving state of the patient, as each new observation o_t and action a_t refine the belief model, leading to more informed and effective decision-making over time. Key to this is ways we update the belief in each iteration, especially using particle filters.

Belief Updating Process: The belief updating process is a key mechanism in our POMDP model. After executing an action a_t and observing an outcome o_t , the belief state is updated to b_{t+1} using the Bayes' rule. This update is mathematically represented as

$$b_{t+1}(s') = \eta Z(o_t | s', a_t) \sum_{s \in S} T(s' | s, a_t) b_t(s), \quad (4)$$

where η is a normalizing constant. This formula adjusts the belief distribution to account for the new observation o_t and the known effects of the action a_t , thereby incrementally refining our understanding of the patient's condition despite the inherent noise in the observations. The implementation of this uses particle filters for fast and reliable computation.

Particle Filters: Particle filters, in this context, operationalize belief updating through a simulation-based approach. Representing the belief b_t as a set of particles $\{s_t^1, s_t^2, \dots, s_t^n\}$, each particle s_t^i is a hypothesis of the patient's state. Upon receiving a new observation o_t , the particle filter updates the weights of these particles based on the likelihood $Z(o_t | s_t^i, a_t)$. The resampling step then focuses on particles with higher weights, effectively concentrating the belief on more probable states. This method is adept at managing the complex dynamics and uncertainties in medical scenarios, offering a nuanced and continuously updated representation of the patient's condition.

TABLE II. SUMMARY OF EXPERIMENT PARAMETERS

Parameter Name	Symbol	Unit	Value
Probability of $IsAne: TRUE$ with non-treatment action	p_{ane}	-	0.0005
Probability of $IsAVM: TRUE$ with non-treatment action	p_{AVM}	-	0.0002
Probability of $IsOcc: TRUE$ with non-treatment action	p_{occ}	-	0.0002
Minimal probability to be considered dominant state prediction (to make treatment action)	p_{dom_thres}	-	0.6
Minimal prediction confidence of stroke-free state for discharging patients (used by both ExpertHOSP and ExpertDSA policies)	p_{disc_min}	-	0.9
Discount factor	γ	-	0.95
Num. particles for belief updater	n	particles	100
Num. replications for performance metrics statistic calculation	K	samples	1,000
Penalty when the patient is not treated	-	-	-100,000
Cost of treatment actions	-	-	-200
Cost of DSA	-	-	-150
Cost of hospitalizing patients	-	-	-100
Reward for correct treatment	-	-	5,000
Penalty for wrong treatment	-	-	-5,000
Reward for needed DSA	-	-	250
Penalty for unnecessary DSA	-	-	-750
Reward for correct HOSP	-	-	150
Penalty for unnecessary HOSP	-	-	-400
Penalty for not hospitalizing stroke patients	-	-	-1,000
Reward for discharging stroke-free patients	-	-	5,000
Penalty for discharging stroke patients	-	-	-50,000

IV. EXPERIMENT SETTING

Our experiment goal is to compare how a POMDP model described in Section III solved with an online solver performs vs. expert policies and a baseline random policy. We simulate $K = 1,000$ cases with various initial conditions. To this end, 0.785 cases are initialized with stroke-free conditions, 0.05 with only one stroke condition (either aneurysm, AVM, or occlusion), 0.02 with two combinations, and 0.005 with all the conditions. In addition, we assume some random emergence of stroke conditions with probability of p_{ane} , p_{AVM} , and p_{occ} for aneurysm, AVM, and occlusion, respectively, during the decision periods.

The reward function considers the accuracy of treatments for the underlying stroke conditions and impose penalties for mismatch and rewards for matching cases. Table II summarizes

these parameters, along with other hyperparameters of the experiment. The full detail of the experiment is open-source and available at [www.github.com/inteligensi/DSAPOMDPs.jl](https://github.com/inteligensi/DSAPOMDPs.jl).

We compare four distinct policies: Random, ExpertHOSP, ExpertDSA, and DESPOT, allowing a robust comparison between static, expert-driven policies, and a POMDP strategy.

Random policy, serving as a baseline, involves selecting actions randomly from the action space. This approach, while straightforward, is expected to perform poorly and only establishes a comparative standard against which the more complex policies can be evaluated.

ExpertHOSP and ExpertDSA, grounded in domain expertise, utilize a belief-based decision-making framework. These policies analyze the current belief state to choose the most suitable action. In scenarios where no specific condition is dominant (the probabilities for all states are $\leq p_{dom_thres}$), both policies default to their respective diagnostic actions (*HOSP* for ExpertHOSP and *DSA* for ExpertDSA). If the belief in a stroke-free condition is sufficiently high ($> p_{disc_min}$), the patient is discharged. However, if the probability of a condition surpasses a certain threshold, the corresponding treatment action for the predicted stroke condition with the highest probability is ordered.

DESPOT, an advanced online POMDP solver, offers a dynamic and adaptive decision-making strategy. It generates policies by simulating potential future scenarios, making it particularly effective in the complex and uncertain realm of our problem.

Our experiment aims to evaluate these policies in terms of recovery rate, time-to-treatment for patients with stroke conditions, and cumulative discounted reward, as detailed in (3). We analyze the advantages and trade-offs of each approach. The performance metrics achieved by all four policies are summarized in Table I. Figure 2 presents normalized histograms of the overall cumulative discounted reward, while Figure 3 displays histograms of the time-to-treatment for patients with initial stroke conditions under each policy. Figure 4 provides a breakdown of the cumulative discounted reward, differentiating between patients requiring treatment and those who do not. Lastly, we illustrate the example actions taken by all policies for both mild (single stroke type, in this case *IsAne: TRUE*) in Figure 5 and severe (multiple stroke types, in this case *IsAVM: TRUE* and *IsOcc: TRUE*) in Figure 6.

V. DISCUSSION

In this section, we analyze the performance of all benchmarked policies and examine the sequence of actions for mild and severe cases to demonstrate each policy's strategy.

A. Performance Benchmarks

Table III clearly shows that the Random policy is the least effective among all benchmarked policies. Notably, our proposed DESPOT policy performs comparably with both expert policies (ExpertDSA and ExpertHOSP) in terms of

TABLE III. SUMMARY OF PERFORMANCE METRICS

Metric	Policy			
	Random	ExpertDSA	ExpertHOSP	DESPOT
Recovery rate	0.458	0.935	0.856	0.931
Disc. reward (mean)	-13,592.66	3,860.30	3,171.89	3,445.03
Disc. reward (std)	18,212.92	8,804.32	10,254.94	8,417.28
Time-to-treatment (mean)	14.66	4.15	6.96	6.04
Time-to-treatment (std)	10.36	5.26	7.49	5.34

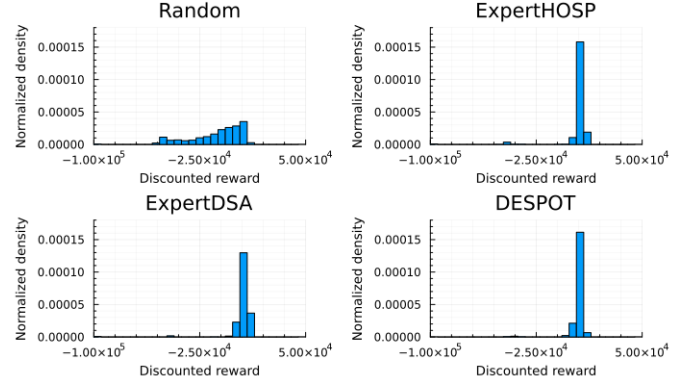


Fig. 2. The normalized histograms of discounted reward

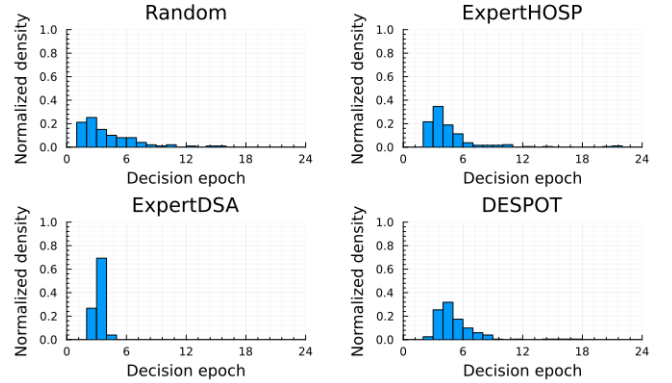


Fig. 3. The normalized histograms of time-to-treatment

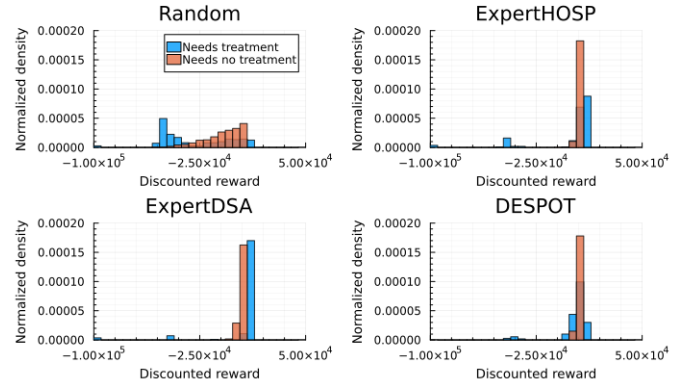


Fig. 4. The normalized histograms of discounted reward for patient with stroke (needing treatment) and without (not needing treatment)

overall objective (mean cumulative discounted reward). It also achieves a competitive recovery rate with ExpertDSA, the top performer, and a slightly better time-to-treatment performance than ExpertHOSP. However, due to the large standard deviations in both discounted reward and time-to-treatment metrics across all policies, we cannot claim statistical significance in these performance comparisons. The substantial standard errors underscore the variability in how these algorithms perform across different stroke conditions, including stroke-free conditions, which constitute approximately 78.5% of the initial patient states.

Examining the normalized histogram of discounted rewards (Fig. 2), we observe that the densities of cumulative rewards for ExpertHOSP, ExpertDSA, and DESPOT follow a similar trend, with most densities around $\pm 5,000$. However, all policies exhibit some densities at extreme negative values ($-20,000$ and $-50,000$), indicating significant penalties from false negative predictions (discharging stroke patients). The Random policy, as expected, displays a distinctly subpar performance pattern.

In terms of time-to-treatment (Fig. 3), DESPOT and ExpertHOSP show very similar densities, indicating comparable performance. ExpertDSA stands out with its density concentrated on shorter time-to-treatment values, reflecting the decisive nature of DSA in identifying stroke conditions and recommending treatment procedures. Random policy appear to have heavier tails, highlighting the higher likelihood for patients to have time-to-treatment values.

Finally, we further analyze the discounted reward metric by breaking down the density of discounted rewards for both stroke (treatment needed) and stroke-free (no treatment needed) patients (Fig. 4). This breakdown reveals that while ExpertDSA performs better (higher reward densities) for stroke cases (blue bars), ExpertHOSP slightly outperforms in stroke-free cases (orange bars). DESPOT, relying solely on internal simulations without the inherent expert knowledge of the two expert policies, appears to balance performance between stroke and stroke-free patients.

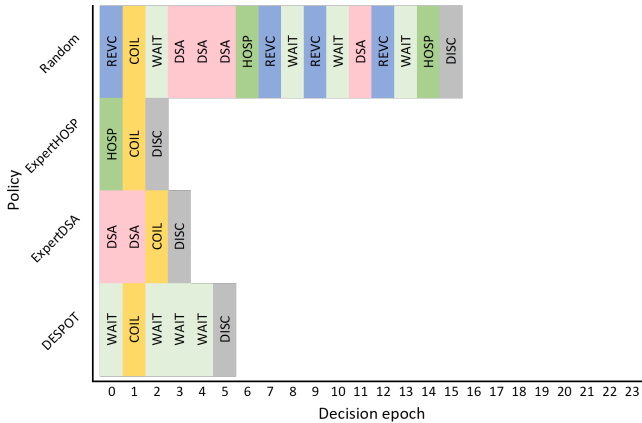


Fig. 5. Sampled actions from all policies for a mild case ($IsAne: TRUE, IsAVM: FALSE, IsOcc: FALSE$)

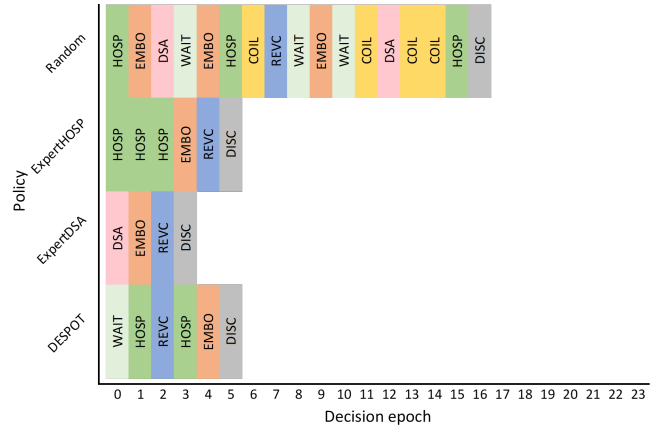


Fig. 6. Sampled actions from all policies for a severe case ($IsAne: TRUE, IsAVM: TRUE, IsOcc: TRUE$)

B. Sampled Actions for Mild and Severe Cases

We also examined the actions taken by all policies for two different stroke conditions. In a mild case (Fig. 5), where patients exhibit only one stroke type, the Random policy selects arbitrary action sequences, many of which are illogical (e.g., choosing *DSA* after a *REVC* procedure). In contrast, ExpertDSA and ExpertHOSP consistently follow their primary strategies, initiating *DSA* and *HOSP* actions, respectively, upon patient arrival. They perform treatment (*COIL*) when the underlying stroke type ($IsAne: TRUE$) is predictable, followed by discharging the patient post-procedure. DESPOT adopts a more cautious and cost-effective approach. It initially recommends a *WAIT* action to gather CT scan and Siriraj score observations before proceeding with treatment. Interestingly, DESPOT often advises observing the patient with three consecutive *WAIT* actions to ensure no untreated conditions remain before selecting the *DISC* action. This approach explains why DESPOT tends to have a longer time-to-treatment compared to ExpertDSA.

Fig. 6 illustrates actions for a more severe case (a patient exhibiting both aneurysm and occlusion, occurring with a probability of only 0.02 in our model). In this scenario, DESPOT again adopts a balanced approach, gradually increasing observation accuracy, but only up to the *HOSP* level (without performing *DSA*), to build its belief about the underlying stroke conditions. It then performs treatment (*REVC*) while continuing to gather observations (opting for another *HOSP* action), before proceeding with another treatment (*EMBO*) and ultimately discharging the patient. ExpertHOSP and ExpertDSA behave as expected, thoroughly collecting data before consecutively performing treatments. This balanced approach highlights the strength and potential of the POMDP modeling approach in effectively integrating diagnostic and treatment goals, assisting medical professionals in providing safe and cost-effective care amidst uncertainties.

VI. CONCLUSION

Our study demonstrates the significant potential of the POMDP modeling approach in enhancing decision-making processes in medical settings, particularly in the context of stroke management. By integrating diagnostic and treatment

strategies, the POMDP framework, exemplified by the DESPOT solver, offers a balanced and effective method for navigating the complexities and uncertainties inherent in medical care. This approach not only optimizes patient outcomes by providing tailored treatment plans but also ensures cost-effectiveness and resource optimization. However, our current formulation is not without limitations. One key constraint is the reliance on the accuracy and comprehensiveness of the input data, particularly in representing the diverse range of stroke conditions and patient responses. Looking ahead, future research should focus on incorporating more diverse and extensive datasets, including patient histories and broader stroke-related variables, to refine the model's accuracy and applicability. By addressing these areas, the POMDP framework can be further developed to become an indispensable tool in medical decision-making, ultimately improving patient care and outcomes in various healthcare settings.

REFERENCES

- [1] S. U. Lee, T. Kim, O. K. Kwon, J. S. Bang, S. P. Ban, H. S. Byoun and C. W. Oh, "Trends in the incidence and treatment of cerebrovascular diseases in Korea: part II cerebral infarction, cerebral arterial stenosis, and Moyamoya disease, J. Korean Neurosurg. Soc., vol. 63, no. 1, pp. 69-79, Jan. 2020.
- [2] G. Duloquin, M. Graber, L. Garnier, V. Crespy, P. O. Comby, L. Baptiste and S. Mohr, "Incidence of acute ischemic stroke with visible arterial occlusion: a population-based study (Dijon stroke registry), Stroke, vol. 51, no. 7, pp. 2122-2130, 2020.
- [3] F. Settecase and V. L. Rayz, "Advanced vascular imaging techniques," Handbook of Clinical Neurology, Elsevier, vol. 176, pp 81-105, 2021.
- [4] G. Turc, B. Maier, O. Naggara, P. Seners, C. Isabel and M. Tisserand, "Clinical scales do not reliably identify acute ischemic stroke patients with large-artery occlusion," Stroke, vol. 47, pp. 1466-72, 2016.
- [5] P. Smita, R. Rosanna, J. Duaa and D. Karen, "Detection, diagnosis and treatment of acute ischemic stroke: current and future perspectives," Frontiers in Medical Technology, vol.4, June 2022.
- [6] S. E. Vermeer, W. T. Jr. Longstreth Jr. and P. J. Koudstaal "Silent brain infarcts: a systematic review," The Lancet Neurology, vol. 6, no. 7, pp. 611-619, 2007.
- [7] E. S. Donkor, "Stroke in the 21st century: a snapshot of the burden, epidemiology, and quality of life," Stroke Res. Treat., vol. 27, 2018
- [8] L. N. Steimle and B. T. Denton, "Markov decision processes for screening and treatment of chronic diseases," Markov Decision Processes in Practice, pp. 189-222, 2017.
- [9] A. Tsoukalas, T. Albertson and I. Tagkopoulos, "From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis," JMIR Medical Informatics, vol. 3, no. 1, 2015.
- [10] G. J. Schell, W. J. Marrero, M. S. Lavieri, J. B. Sussman, and R. A. Hayward, "Data-driven Markov decision process approximations for personalized hypertension treatment planning," MDM Policy and Practice, vol. 1 no. 1, 2016.
- [11] M. J. Kochenderfer, T. A. Wheeler and K. H. Wray, Algorithms for Decision Making, MIT Press, 2022.
- [12] D. E. Ehrmann, S. Joshi, S. D. Goodfellow, M. L. Mazwi and D. Eytan, "Making machine learning matter to clinicians: model actionability in medical decision-making," NPJ Digital Medicine, vol. 6, no. 1, 2023.
- [13] J. Fischer and O. S. Tas, "Information particle filter tree: an online algorithm for POMDPs with belief-based rewards on continuous domains," Intl. Conf. on Machine Learning, PMLR, 2020.
- [14] Y. Feng and S. J. Shu, "Diagnostic value of low-dose 256-slice spiral CT angiography, Mr Angiography, and 3D-DSA in cerebral aneurysms", Disease Markers, pp. 1-5, 2020.
- [15] A. H. Chowdhury, S. K. Ghose, Q. D. Mohammad, M. Habib, S. U. Khan, K. M. Rahman, "Digital subtraction angiography is superior to magnetic resonance angiography in diagnosis of cerebral arteriovenous malformation," Mymensingh Med J, vol. 24, no. 2, pp. 356-65, 2015.
- [16] R. Liu, Z. Yan, G. Zhang, Z. Ding, Y. Li and Z. Jiang, "Comparison of digital subtraction angiography and contrast-enhanced ultrasound in assessment of carotid stenosis," Afr. Health Sci, vol. 20, no. 1, pp. 509-514, 2020.
- [17] N. Ye, A. Somani, D. Hsu and W. S. Lee, "DESPOT: online POMDP planning with regularization," Journal of Artificial Intelligence Research, vol. 58, AI Access Foundation, pp. 231-266, Jan. 26, 2017.
- [18] C. P. Derdeyn, G. J. Zipfel, F. C. Albuquerque, D. L. Cooke, E. Feldmann, J. P. Sheehan, and J. C. Torner, "Management of brain arteriovenous malformations: a scientific statement for healthcare professionals from the American Heart Association/American Stroke Association", Stroke, vol. 48, no. 8, pp. e200-e224, 2017.
- [19] J. Zhao, H. Lin, R. Summers, M. Yang, B. G. Cousins and J. Tsui, "Current treatment strategies for intracranial aneurysms: an overview," Angiology, vol. 69, no. 1, pp. 17-30, 2017.
- [20] R. Hurford, A. Sekhar and T. A. T. Hughes, "Diagnosis and management of acute ischaemic stroke," Practical Neurology, vol. 20, pp. 304-316, 2020.
- [21] J. G. Merino, M. Luby, R. T. Benson, "Predictors of acute stroke mimics in 8187 patients referred to a stroke service", 2020.
- [22] N. Pongvarin, A. Viriyavejakul and C. Komontri, "Siriraj stroke score and validation study to distinguish supratentorial intracerebral hemorrhage from infarction," British Medical Journal, 1991.
- [23] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable Markov Processes over a finite horizon," Operations Research, vol. 21 no. 5, pp. 1071-1088, 1973.
- [24] O. Madani, S. Hanks and A. Condon, "On the undecidability of probabilistic planning and infinite-horizon partially observable Markov decision problems," AAAI Conf. on Artificial Intelligence, pp. 541-548, 1999.
- [25] S. Ross, J. Pineau, S. Paquet, and B. Chaib-draa, "Online planning algorithms for POMDPs," Journal of Artificial Intelligence Research, vol. 32. AI Access Foundation, pp. 663-704, Jul. 29, 2008.
- [26] M. H. M. Winands, Y. Björnsson, and J.-T. Saito, "Monte-Carlo tree search solver," Computers and Games. Springer Berlin Heidelberg, pp. 25-36, 2008
- [27] C. H. Ulfsoo and D. Axehill, "On integrating POMDP and scenario MPC for planning under uncertainty – with applications to highway driving," 2022 IEEE Intelligent Vehicles Symposium (IV). IEEE, Jun. 05, 2022.
- [28] T. Ayer, O. Alagoz, and N.K. Stout, "A POMDP approach to personalize mammography screening decisions," Operations Research, vol. 60, no. 5, pp.1019-1034, 2012.
- [29] R. Ibrahim, B. Kucukyazici, V. Verter, M. Gendreau and M. Blostein, "Designing personalized treatment: an application to anticoagulation therapy," Production and Operations Management, vol. 25 no. 5, pp. 902-918, 2016.

AUTHORS' INFORMATION

Your Name	Title	Research Field	Personal website
Nur Ahmad Khatim	Researcher	Algorithms and programming	https://naimackerman.github.io/
Azmul Asmar Irfan	Medical Doctor Candidate, Researcher	Stroke assessment for diagnose	-
Amaliya Mata'ul	Medical Doctor Candidate, Researcher	Stroke assessment for diagnose	-
Mansur M. Arief	Doctor, Postdoctoral Researcher	Safe AI-assisted decision-making systems, POMDP.	https://mansurarieff.github.io/