Decision Aliasing in GUI Agents:
Information Requirements for Safe Commitment

Richard Shin
College of Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
rshin41@gatech.edu

July 19, 2026

Abstract

Visual GUI agents act from screenshots, action histories, and sometimes structured metadata. Their benchmark performance does not answer a prior safety question: when does an observation interface contain enough latent-state information to support a correct irreversible commitment? We place GUI interaction in a finite-horizon partially observable Markov decision process and study agent-visible summaries of interaction history, using classical belief-state control as the reference model. A summary is *decision-sufficient* at a horizon when every reachable belief state represented by the same summary shares an optimal action. We prove that this condition is necessary and sufficient for the existence of a summary-measurable policy that is optimal at every reachable history. When a visible summary aliases κ belief states with pairwise disjoint optimal-action sets, every randomized summary-based policy is suboptimal on at least one belief with probability at least $1 - 1/\kappa$; a corresponding regret bound follows from the optimal-action gap.

For zero-error one-step decisions, we define the action-cover number induced by the zero-loss action correspondence. This quantity is an ordinary minimum set-cover instance and exactly characterizes the minimum number of deterministic semantic messages needed to resolve a visual equivalence class. An incompatibility packing number gives a simple lower bound. Under nonzero error, Fano's inequality lower-bounds the mutual information required from an additional semantic channel. We then define a diagnostic GUI POMDP in which safe probes gather information before an irreversible commitment. Deterministic b-ary probing requires logarithmic depth in the action-cover number; noisy adaptive probing obeys an information-per-probe lower bound; and if every safe probe is observationally identical across two incompatible states, no safe strategy can guarantee success. These results specialize classical ideas from POMDP belief-state control, Blackwell comparison of experiments, state abstraction, zero-error functional coding, set cover, and active diagnosis to latent-state GUI commitments. They do not claim that screenshot agents are ineffective. They identify the state distinctions that no visual reasoning system can recover when the interface does not expose them.

Keywords: computer-use agents, GUI agents, human–computer interaction, POMDPs, partial observability, decision sufficiency, information theory, safe exploration

1 Introduction

Computer-use agents increasingly execute natural-language instructions through ordinary graphical interfaces. Vision-first systems consume screenshots and emit coordinate-level mouse, keyboard, or

touch actions, reducing dependence on application-specific APIs and markup (Gou et al., 2024; Qin et al., 2025). Benchmarks such as OSWorld and VisualWebArena demonstrate that this approach can support broad and realistic tasks (Xie et al., 2024; Koh et al., 2024). Yet task completion rates do not establish that the observation interface contains enough information for a correct decision in every state that can produce the same visible screen.

A rendered frame is a projection of a larger system state. The same pixels may coexist with different server acknowledgments, focus targets, permission states, version numbers, prior side effects, pending requests, or reversibility guarantees. A button can look unchanged before and after an accepted transaction if the confirmation response is lost. Two text fields can look identical while keyboard focus differs. A collaborative document can display version v while the server has already advanced to $v + 1$. In each case, an agent may perceive every visible pixel correctly and still lack a state distinction that changes what action is optimal or safe.

Recent GUI-agent work explicitly recognizes partial observability. Surveys identify hidden state, dynamic interfaces, and irreversible actions as central trustworthiness problems (Shi et al., 2025). HalluClear models GUI interaction as a POMDP and notes that perceptual aliasing creates an unavoidable gap when identical observations require different optimal actions (Jin et al., 2026). Other work studies richer observation interfaces, compressed accessibility structure, temporal and multimodal observation, and conflicts between pixels and serialized structure (Takeshita et al., 2026; Li and Shi, 2026; Zhang and Yang, 2026). The present paper does not claim to originate perceptual aliasing or partial-observation modeling. Instead, it asks a narrower interface-design question:

> What latent-state distinctions must an interface expose before an agent can safely commit, and when can idempotency, reversibility, or safe diagnosis remove that requirement?

We answer this question at three levels. First, we use a finite-horizon GUI POMDP to state the summary-measurability condition underlying decision aliasing. Second, at an irreversible commitment point we characterize the exact semantic refinement needed to resolve a visual equivalence class. Third, we model safe active diagnosis before commitment and derive depth and information lower bounds. This emphasis on latent-state commitment losses distinguishes the paper from quotient constructions that preserve observation-measurable objectives.

The paper's main contributions are:

1. An exact zero-error semantic-message characterization through an *action-cover number*, explicitly reduced to minimum set cover, with incompatibility-packing and Fano lower bounds.
2. A diagnostic GUI POMDP for safe probes followed by irreversible commitment, with deterministic depth, noisy information-per-probe, and no-safe-probe impossibility results.
3. A formal distinction between *observation refinement*, which splits aliased states, and *action regularization*, such as idempotency or undo, which creates common safe actions and can reduce semantic complexity without revealing more state.
4. A GUI-specific synthesis connecting latent-state commitment losses to transaction state, stable identity, focus, validation, permissions, versioning, reversibility, and safe status queries.

For completeness, we also state a summary-measurability characterization and randomized regret bound in the underlying finite-horizon POMDP. These are positioning lemmas rather than claims to replace classical POMDP or state-abstraction theory.

These are guarantee results, not average-case predictions. Screenshot-only agents may perform well when ambiguous states are rare, priors are concentrated, interaction history is informative, or errors are inexpensive. The results isolate a different limitation: computation and visual-model scale cannot reconstruct a task-relevant variable that is absent from the complete observation history.

2 Positioning and Related Work

2.1 Visual GUI agents and observation interfaces

UGround and UI-TARS exemplify the vision-first approach to GUI control (Gou et al., 2024; Qin et al., 2025). OSWorld and VisualWebArena evaluate open-ended interaction in realistic desktop and web environments (Xie et al., 2024; Koh et al., 2024). A growing line of work instead enriches or restructures the observation channel. A11y-Compressor compresses accessibility observations while preserving useful structure (Takeshita et al., 2026); Agent-Computer Observation Interfaces expands observation across time and modality (Li and Shi, 2026); Screen2AX reconstructs accessibility structure from screenshots (Muryn et al., 2025); and recent controlled studies examine how agents resolve conflicts between pixels and serialized UI structure (Zhang and Yang, 2026). These papers study systems, representations, or empirical behavior. Our object is the decision-theoretic adequacy of the observation channel itself.

2.2 Perceptual aliasing in GUI agents

HalluClear already states the core aliasing phenomenon in a GUI-agent POMDP: an observation-measurable policy can be strictly worse than a privileged-state oracle when states with the same observation have different optimal actions (Jin et al., 2026). This directly overlaps the motivating intuition of this paper. Our contribution begins after that observation: we characterize summary-measurable optimal policies over reachable beliefs, quantify incompatible decision cells, derive exact semantic-refinement requirements, and analyze safe adaptive information acquisition.

2.3 POMDP belief states and information states

In a POMDP, the posterior belief over latent state is a sufficient statistic of the complete action-observation history for optimal control (Kaelbling et al., 1998; Smallwood and Sondik, 1973). This is classical and is not a contribution of this paper. We use the belief MDP as the reference against which a lossy GUI summary is evaluated. A current screenshot is generally not a belief state; a screenshot history may still fail to determine one; and a practical agent may use a learned internal summary. Our characterization asks exactly when such a summary preserves at least one optimal action at every reachable belief it represents.

2.4 Statistical sufficiency and comparison of experiments

Blackwell's comparison of experiments orders observation channels by whether one can be obtained as a garbling of another; a Blackwell-more-informative experiment weakly improves Bayes value for every decision problem (Blackwell, 1953). Le Cam's deficiency extends this perspective to approximate comparison (Le Cam, 1964). Our deterministic refinement result is a finite zero-error specialization of this broader theory, not a replacement for it. The GUI-specific question is which hidden interface distinctions must be exposed for a particular family of user goals and irreversible actions.

2.5 State abstraction and action-relevant representations

MDP abstraction and bisimulation study when states can be aggregated while preserving values or policies (Li et al., 2006; Abel et al., 2016). Recent work uses the term *action sufficiency* for goal representations in hierarchical reinforcement learning, defined through conditional mutual information between goals and optimal actions (Hyeon et al., 2026). Our setting differs in object and criterion: we study an observation interface over hidden GUI state, permit set-valued optimal actions, emphasize zero-error guarantees and safe probes, and explicitly relate the required refinement to interface semantics. To avoid terminological collision, we use *decision sufficiency* for the POMDP summary and *action compatibility* for zero-error state groups.

2.6 Bounded interaction and canonical POMDP quotients

Nixon develops a bounded-interaction Myhill–Nerode theory for finite POMDPs, defining controller-family-dependent indistinguishability and canonical minimal quotients that are decision-sufficient for observation-action measurable objectives (Nixon, 2026). That work is substantially broader than our summary characterization and overlaps the use of adaptive probes. The distinction relevant here is the objective boundary it makes explicit: exact quotient sufficiency applies to objectives measurable from the agent's own observation-action trajectory, while rewards depending on latent state receive approximate bounds. Our commitment loss is intentionally latent-state-dependent—for example, duplicate payment may be harmful even when the visible transcript is unchanged. We therefore focus on the minimum semantic refinement or action redesign needed for zero-error latent-state commitments, rather than on canonical quotient construction.

2.7 Zero-error functional coding and set cover

Zero-error functional coding asks how much information an encoder must reveal for a decoder to compute a task-relevant function rather than reconstruct the entire source (Orlitsky and Roche, 2001). Our one-step semantic channel has the same task-relative character, except correctness is set-valued: a state may admit several zero-loss actions. The exact deterministic message problem reduces to minimum set cover over the zero-loss supports of actions. Set cover and its computational hardness are classical (Karp, 1972); the contribution here is the GUI interpretation, its connection to safe commitment, and its use as the terminal quantity in the diagnostic lower bounds.

2.8 What is intended to be new

The elementary statement "identical observations can require different actions" is not new. Neither are belief-state sufficiency, Blackwell monotonicity, Fano's inequality, zero-error functional coding, set cover, or controller-bounded POMDP quotients. The intended technical contribution is narrower: a GUI-specific reduction from semantic refinement to action-support set cover, its interaction with action redesign, and lower bounds for state-preserving safe diagnosis before latent-state-dependent irreversible actions. The novelty claim should therefore be read as a domain-specific formal synthesis and extension, not as a new foundation for statistical decision theory or combinatorial optimization.

3 POMDP Setup and Summary-Measurability Lemma

3.1 Finite-horizon model

Fix a user goal $g \in \mathcal{G}$. A finite-horizon GUI POMDP is

$$\mathcal{M}_g = (\mathcal{X}, \mathcal{A}, \mathcal{O}, T, Z, c_g, c_g^H, b_0, H), \tag{1}$$

where all state, action, and observation sets are finite. The transition kernel is

$$T(x' \mid x, a) = \mathbb{P}(X_{t+1} = x' \mid X_t = x, A_t = a), \tag{2}$$

and the observation kernel is

$$Z(o \mid x) = \mathbb{P}(O_t = o \mid X_t = x). \tag{3}$$

The one-step cost is $c_g : \mathcal{X} \times \mathcal{A} \to [0, \infty)$, the terminal cost is $c_g^H : \mathcal{X} \to [0, \infty)$, the initial belief is $b_0 \in \Delta(\mathcal{X})$, and H is the action horizon. The latent state may include rendered UI state, server state, pending network operations, focus, permissions, historical side effects, and any other variable needed for the Markov property.

The observation timing is: X_t generates O_t, the agent selects A_t, and the system transitions to X_{t+1}. A history at time t is

$$h_t = (o_0, a_0, o_1, \ldots, a_{t-1}, o_t) \in \mathcal{H}_t. \tag{4}$$

A history-dependent policy is $\pi_t : \mathcal{H}_t \to \Delta(\mathcal{A})$.

The belief induced by h_t is

$$b_t(x) = \mathbb{P}(X_t = x \mid h_t, g). \tag{5}$$

Given belief b, action a, and next observation o', the Bayesian update is

$$\tau(b, a, o')(x') = \frac{Z(o' \mid x') \sum_x T(x' \mid x, a) b(x)}{\sum_{\bar{x}} Z(o' \mid \bar{x}) \sum_x T(\bar{x} \mid x, a) b(x)}, \tag{6}$$

whenever the denominator is positive.

Define

$$V_H^*(b) = \sum_x b(x) c_g^H(x), \tag{7}$$

$$Q_t^*(b, a) = \sum_x b(x) c_g(x, a) + \sum_{o'} \mathbb{P}(o' \mid b, a) V_{t+1}^*(\tau(b, a, o')), \tag{8}$$

$$V_t^*(b) = \min_{a \in \mathcal{A}} Q_t^*(b, a), \tag{9}$$

for $t = 0, \ldots, H-1$. The optimal-action correspondence is

$$\mathcal{A}_t^*(b) = \arg\min_{a \in \mathcal{A}} Q_t^*(b, a). \tag{10}$$

Because the spaces are finite, this set is nonempty.

Proposition 1 (Classical belief-state sufficiency)**.** *For every finite-horizon POMDP, there exists an optimal policy of the form $\mu_t^*(b_t)$; the belief state is a sufficient statistic of the complete history for optimal control.*

Proof. The recursion above defines the belief MDP. Conditional distributions of future states, observations, costs, and updated beliefs depend on the history only through b_t. Backward induction therefore yields an optimal Markov policy on beliefs. This is the standard POMDP reduction (Kaelbling et al., 1998; Smallwood and Sondik, 1973). □

3.2 Agent-visible summaries

A deployed agent rarely receives the exact belief state. Let

$$\eta_t : \mathcal{H}_t \to \mathcal{Y}_t \tag{11}$$

be an agent-visible summary. Examples include the current screenshot, the last k screenshots and actions, a learned recurrent state, or a screenshot combined with DOM or accessibility metadata.

For a summary value $y \in \mathcal{Y}_t$, define the reachable belief fiber

$$\mathcal{B}_t(y) = \{b(h_t) : h_t \text{ is reachable and } \eta_t(h_t) = y\}. \tag{12}$$

Definition 1 (Dynamic decision sufficiency)**.** *The summary $\eta = (\eta_0, \ldots, \eta_{H-1})$ is* decision-sufficient *for $\mathcal{M}_g$ if, for every time t and every reachable summary y,*

$$\bigcap_{b \in \mathcal{B}_t(y)} \mathcal{A}_t^*(b) \neq \varnothing. \tag{13}$$

The summary need not reconstruct the belief. It need only preserve a common optimal action across beliefs it aliases.

Theorem 1 (Characterization of summary-measurable optimal control)**.** *There exists a deterministic policy $\pi_t(y)$ measurable with respect to η_t that is optimal at every reachable history if and only if η is decision-sufficient.*

Proof. Suppose an η-measurable policy π is optimal at every reachable history. Fix t and reachable y. For every $b \in \mathcal{B}_t(y)$, the policy selects the same action $\pi_t(y)$, and optimality implies $\pi_t(y) \in \mathcal{A}_t^*(b)$. Hence $\pi_t(y)$ lies in the intersection in Equation (13).

Conversely, suppose Equation (13) holds. For each reachable pair (t, y), choose one action

$$a_{t,y} \in \bigcap_{b \in \mathcal{B}_t(y)} \mathcal{A}_t^*(b),$$

and define $\pi_t(y) = a_{t,y}$. At every reachable history, this action is optimal for the corresponding belief. Backward induction on the belief-MDP value recursion shows that the resulting policy attains V_t^* at every reachable belief. □

Corollary 1 (Visual-history impossibility)**.** *If two reachable histories have the same visual summary y but their belief states have disjoint optimal-action sets, no deterministic policy using only y is optimal at both histories.*

Remark 1. *HalluClear identifies the same qualitative source of an oracle gap as perceptual aliasing (Jin et al., 2026). Theorem 1 states the exact set-valued condition for a chosen history summary over all reachable belief states and all horizons.*

4 Quantifying Decision Aliasing

4.1 Incompatibility packing

Definition 2 (Optimal-action incompatibility number)**.** *For a reachable summary y at time t, define*

$$\kappa_t(y) = \max\left\{|K| : K \subseteq \mathcal{B}_t(y),\ \mathcal{A}_t^*(b) \cap \mathcal{A}_t^*(b') = \varnothing \text{ for all } b \neq b' \in K\right\}. \tag{14}$$

A randomized η-policy uses a common distribution $q(\cdot \mid y)$ for every belief in the same fiber.

Theorem 2 (Randomized optimal-action bound)**.** *For every randomized summary-based policy and every reachable (t, y) with finite $\kappa = \kappa_t(y)$,*

$$\inf_{b \in \mathcal{B}_t(y)} \mathbb{P}_{A \sim q(\cdot|y)} \left[A \in \mathcal{A}_t^*(b) \right] \leq \frac{1}{\kappa}. \tag{15}$$

Proof. Choose $K = \{b_1, \ldots, b_\kappa\}$ attaining the maximum. Their optimal-action sets are pairwise disjoint. Therefore

$$\begin{aligned} \sum_{i=1}^{\kappa} \mathbb{P}[A \in \mathcal{A}_t^*(b_i)] &= \sum_{i=1}^{\kappa} \sum_{a \in \mathcal{A}_t^*(b_i)} q(a \mid y) \\ &\leq \sum_{a \in \mathcal{A}} q(a \mid y) = 1. \end{aligned}$$

At least one term is no greater than $1/\kappa$. □

Define the optimal-action gap

$$\Delta_t(b) = \min_{a \notin \mathcal{A}_t^*(b)} \left(Q_t^*(b, a) - V_t^*(b) \right), \tag{16}$$

with $\Delta_t(b) = +\infty$ if every action is optimal.

Corollary 2 (One-step dynamic regret)**.** *Let $K \subseteq \mathcal{B}_t(y)$ be a pairwise incompatible set of size $\kappa \geq 2$ and let $\underline{\Delta} = \min_{b \in K} \Delta_t(b) > 0$. For every randomized summary policy, some $b \in K$ satisfies*

$$\mathbb{E}[Q_t^*(b, A) - V_t^*(b)] \geq \underline{\Delta} \left(1 - \frac{1}{\kappa} \right). \tag{17}$$

Proof. By Theorem 2, some $b \in K$ receives an optimal action with probability at most $1/\kappa$. Every nonoptimal action has gap at least $\Delta_t(b) \geq \underline{\Delta}$. Multiplying the minimum gap by the probability of a nonoptimal action gives Equation (17). □

4.2 Bayes-optimal action under a lossy summary

Let $P(B = b \mid Y = y)$ be the posterior over reachable beliefs induced by a deployment distribution.

Proposition 2 (Bayes-optimal summary action)**.** *For a fixed (t, y), the maximum probability of selecting an optimal action is*

$$S_t^*(y) = \max_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}_t(y)} P(b \mid y) \mathbf{1}\{a \in \mathcal{A}_t^*(b)\}. \tag{18}$$

Randomization cannot improve this value.

Proof. The success probability is linear in the action distribution, so its maximum over the simplex occurs at a deterministic extreme point. □

This separates worst-case insufficiency from deployment performance. A summary can fail the universal criterion while remaining highly effective under a concentrated posterior.

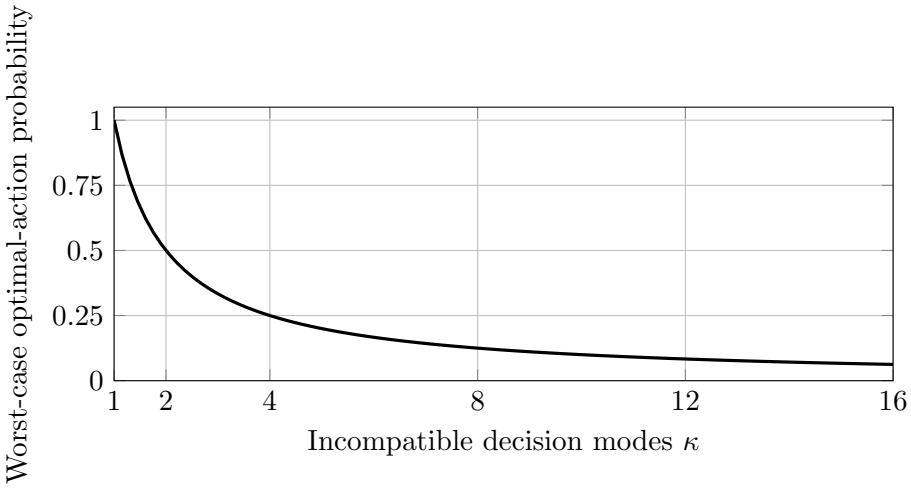

Figure 1: Any randomized policy that receives the same visible summary in κ pairwise incompatible decision modes has worst-case optimal-action probability at most $1/\kappa$.

5 Zero-Error Semantic Refinement

The dynamic model reduces at a commitment point to a one-step decision problem. Let C be a finite set of hidden states consistent with the same visual observation and goal. For each $x \in C$, let $\Gamma(x) \subseteq \mathcal{A}$ be the nonempty set of zero-loss actions. For each action $a \in \mathcal{A}$, define its zero-loss support

$$C_a = \{x \in C : a \in \Gamma(x)\}. \tag{19}$$

Definition 3 (Action-cover number)**.** *The* action-cover number *is*

$$\tau_\Gamma(C) = \min\left\{|U| : U \subseteq \mathcal{A},\ C \subseteq \bigcup_{a \in U} C_a\right\}. \tag{20}$$

It is the minimum number of zero-loss action supports needed to cover the aliased state class.

Computing $\tau_\Gamma(C)$ is precisely a minimum set-cover problem on universe C with sets $\{C_a : a \in \mathcal{A}\}$. We use the separate notation only to make the interface interpretation explicit.

Theorem 3 (Exact deterministic semantic-message complexity)**.** *Let $z : C \to \mathcal{Z}$ be a deterministic semantic side channel and let a policy choose one action from each message. A zero-error policy based on the visual observation and $z(x)$ exists if and only if every nonempty message fiber $z^{-1}(u)$ is contained in C_a for some action a. Consequently,*

$$\min_{z,\pi} |z(C)| = \tau_\Gamma(C), \tag{21}$$

and a fixed-length binary representation requires

$$\lceil \log_2 \tau_\Gamma(C) \rceil \tag{22}$$

bits in the worst case.

Proof. Suppose a zero-error encoder-policy pair uses the nonempty messages $u_1, \dots, u_m$, and let a_j be the action selected after message u_j. Zero error implies

$$z^{-1}(u_j) \subseteq C_{a_j}$$

for every j. Since the message fibers cover C, the distinct selected actions form an action cover of size at most m. Therefore $m \geq \tau_\Gamma(C)$.

Conversely, let $U = \{a_1, \dots, a_m\}$ be a minimum action cover. Assign each state x to any index j for which $x \in C_{a_j}$, use that index as the message, and choose a_j after receiving it. The resulting policy has zero loss in every state and uses at most $m = \tau_\Gamma(C)$ messages. Together with the lower bound, equality follows. The bit statement follows from fixed-length coding of $\tau_\Gamma(C)$ messages. □

Definition 4 (Incompatibility packing number)**.** *Define*

$$\kappa_\Gamma(C) = \max\{|K| : K \subseteq C,\ \Gamma(x) \cap \Gamma(y) = \varnothing \text{ for all } x \neq y \in K\}. \tag{23}$$

Proposition 3 (Packing lower bound)**.**

$$\kappa_\Gamma(C) \leq \tau_\Gamma(C). \tag{24}$$

Proof. A single action support C_a cannot contain two states whose zero-loss action sets are disjoint. Hence every action cover must use at least one support for each member of a pairwise incompatible set. □

The inequality can be strict. Pairwise incompatibility is only a packing certificate, whereas $\tau_\Gamma(C)$ solves the full covering problem. Exact computation is NP-hard in general by the standard set-cover reduction (Karp, 1972).

5.1 Approximate information requirement

Let X be uniform on a pairwise incompatible set K of size κ. The visual observation is constant on K, and an additional possibly randomized signal is Z.

Theorem 4 (Fano lower bound)**.** *If a policy based on Z selects an acceptable action with error probability at most $\varepsilon < 1 - 1/\kappa$, then*

$$I(X; Z) \geq \log_2 \kappa - h_2(\varepsilon) - \varepsilon \log_2(\kappa - 1), \tag{25}$$

where h_2 is binary entropy.

Proof. Let A be the possibly randomized action produced from Z. Because the acceptable-action sets are pairwise disjoint, a successful action identifies the unique state whose set contains it. Map all other actions arbitrarily to obtain an estimator $\widehat{X}(A)$ with error probability at most ε. Fano's inequality therefore gives

$$I(X; A) \geq \log_2 \kappa - h_2(\varepsilon) - \varepsilon \log_2(\kappa - 1).$$

The Markov chain $X \to Z \to A$ and the data-processing inequality imply $I(X; Z) \geq I(X; A)$, yielding Equation (25). □

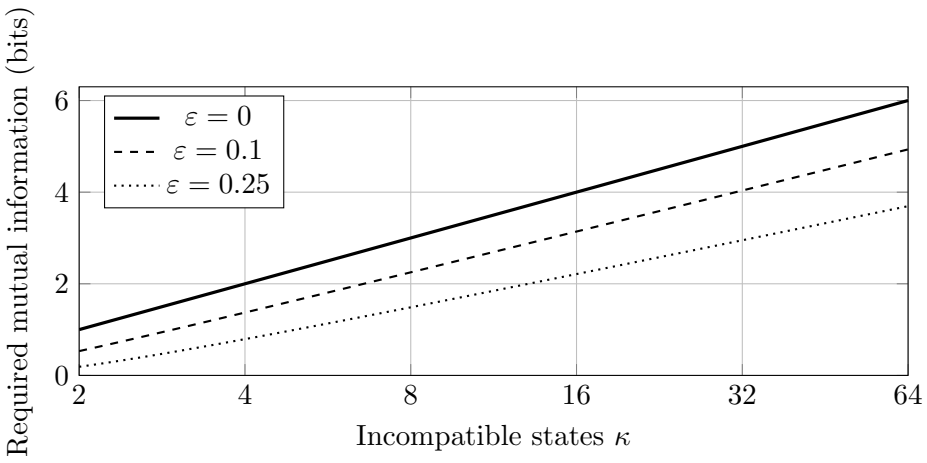

Figure 2: Fano lower bounds for a uniform class of pairwise action-incompatible hidden states.

6 Observation Refinement, Action Regularization, and Information Value

Definition 5 (Deterministic refinement)**.** *An observation Ω_2 refines Ω_1 if there exists f such that $\Omega_1 = f \circ \Omega_2$.*

Proposition 4 (Monotonicity of zero-error decision complexity)**.** *If Ω_2 refines Ω_1, then each Ω_2 equivalence class is contained in an Ω_1 class. Therefore refinement cannot increase the action-cover number within the corresponding class and preserves decision sufficiency once achieved.*

Proof. Equality under Ω_2 implies equality under Ω_1. Restricting an action cover of the larger class to a subset leaves a valid cover with no more actions. A common zero-loss action for the larger class remains zero-loss on every subset. □

For stochastic channels, the appropriate general relation is Blackwell dominance. If channel Z_1 is a stochastic garbling of Z_2, every policy based on Z_1 can be simulated from Z_2, so Z_2 weakly improves Bayes value for every prior and loss function (Blackwell, 1953). This classical theorem is stronger and more general than deterministic partition refinement. The GUI interpretation is that an agent-facing semantic channel should be evaluated by the decision problems it resolves, not by token count or structural richness alone.

Proposition 5 (Monotonicity under action regularization)**.** *Let Γ and Γ' be zero-loss action correspondences on the same visual class C, with*

$$\Gamma(x) \subseteq \Gamma'(x) \quad \text{for every } x \in C. \tag{26}$$

Then

$$\tau_{\Gamma'}(C) \leq \tau_{\Gamma}(C). \tag{27}$$

Proof. For every action a, its support under Γ' contains its support under Γ. Hence every action cover valid under Γ remains valid under Γ', and the minimum cannot increase. □

Observation refinement and action regularization are dual design moves. A transaction-status field splits "not submitted" from "already accepted." An idempotency key instead expands the safe-action sets so that resubmission is acceptable in both states. The former spends semantic information; the latter changes the transition and loss structure.

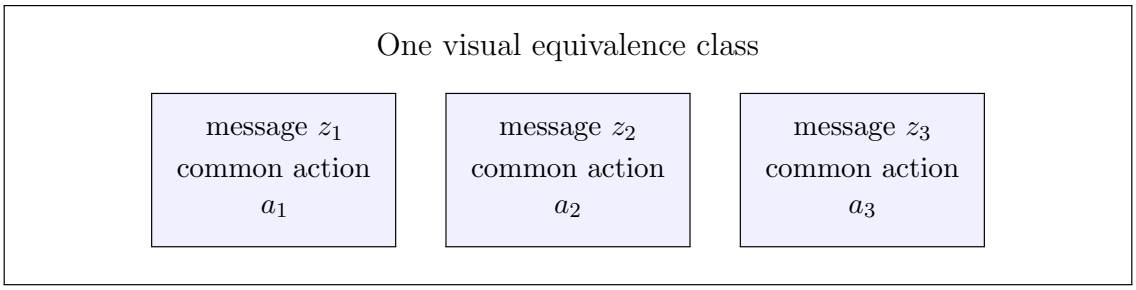

Figure 3: A semantic channel is zero-error sufficient when each message fiber lies within the zero-loss support of its selected action. The minimum number of messages is τ_Γ.

7 Safe Active Diagnosis Before Commitment

A GUI agent can interact to gain information. General POMDP actions can simultaneously alter the state and reveal it, so information gathering and task execution are not cleanly separable (Kaelbling et al., 1998). For irreversible interface actions, however, a useful special case distinguishes safe diagnostic actions from terminal commitments.

7.1 Diagnostic GUI POMDP

A diagnostic GUI POMDP has latent state $X \in C$ fixed during diagnosis, probe actions $p \in \mathcal{P}$, and commitment actions $a \in \mathcal{C}$. For a probe p,

$$T(x' \mid x, p) = \mathbf{1}\{x' = x\}, \tag{28}$$

and the agent receives outcome $Y \sim Q_p(\cdot \mid x)$. Probes may incur a nonnegative cost but do not themselves violate the task. A commitment terminates the episode with loss $L(x, a)$. The zero-error action set is

$$\Gamma(x) = \{a \in \mathcal{C} : L(x, a) = 0\}. \tag{29}$$

A diagnostic history is

$$d_n = (p_1, y_1, \ldots, p_n, y_n), \tag{30}$$

and an adaptive strategy chooses either another probe or a commitment from the history. This is a fully specified POMDP special case: probes have identity transitions, action-dependent observation kernels, and terminal losses.

7.2 Deterministic probes

Suppose each probe outcome is deterministic given x, and each probe has at most $b \geq 2$ possible outcomes.

Theorem 5 (Safe deterministic depth lower bound)**.** *Every zero-error diagnostic strategy has worst-case probe depth*

$$d \geq \lceil \log_b \tau_\Gamma(C) \rceil . \tag{31}$$

In particular, $d \geq \lceil \log_b \kappa_\Gamma(C) \rceil$.

Proof. The leaves of the probe decision tree partition C by transcript. At a leaf, the strategy selects one commitment action, so all states reaching that leaf lie in that action's zero-loss support. The leaf actions therefore form an action cover of C, and there must be at least $\tau_\Gamma(C)$ nonempty leaves. A depth-d tree with branching factor at most b has at most b^d leaves. Thus $b^d \geq \tau_\Gamma(C)$. □

7.3 Noisy adaptive probes

Now let probe outcomes be stochastic. Let X be uniform on a pairwise incompatible set of size κ, and let D_n denote the full adaptive transcript after at most n probes. Strategies that stop early are padded with a null probe and null outcome. The selected probe P_i is a function of the previous transcript.

Theorem 6 (Information-per-probe lower bound)**.** *Assume every admissible adaptive probe satisfies*

$$I(X; Y_i \mid D_{i-1}, P_i) \le C_{\max} \tag{32}$$

for all histories with positive probability. If a strategy commits with error probability at most ε, then

$$n \ge \frac{\log_2 \kappa - h_2(\varepsilon) - \varepsilon \log_2(\kappa - 1)}{C_{\max}}. \tag{33}$$

Proof. By the chain rule and because P_i is determined by D_{i-1},

$$\begin{aligned} I(X; D_n) &= \sum_{i=1}^{n} I(X; Y_i \mid D_{i-1}, P_i) \\ &\le n C_{\max}. \end{aligned}$$

Let A be the final possibly randomized commitment. Pairwise action incompatibility turns a successful commitment into an estimator of X, so Fano's inequality gives

$$I(X; A) \ge \log_2 \kappa - h_2(\varepsilon) - \varepsilon \log_2(\kappa - 1).$$

Because $X \to D_n \to A$ is a Markov chain, data processing gives $I(X; D_n) \ge I(X; A)$. Combining the inequalities yields Equation (33). □

Theorem 7 (No-safe-probe impossibility)**.** *Suppose $x_0, x_1 \in C$ have disjoint zero-loss commitment sets and satisfy*

$$Q_p(\cdot \mid x_0) = Q_p(\cdot \mid x_1) \quad \text{for every safe probe } p. \tag{34}$$

Then every adaptive safe-probe transcript has the same distribution under x_0 and x_1. No safe strategy can guarantee zero error; under the uniform prior on the pair, commitment error is at least $1/2$.

Proof. Proceed by induction on transcript length. Before any probe, the empty transcripts are identical. Assume the transcript distributions agree through step $i - 1$. The strategy therefore induces the same conditional distribution over the next probe in both states. By Equation (34), the next outcome has the same conditional distribution in both states for every selected probe. Hence the length-i transcript distributions agree. By induction, the complete transcript is statistically independent of which of the two states is present. The final commitment distribution is therefore the same in both states. Since the zero-loss action sets are disjoint, the two success probabilities sum to at most one, so one is at most $1/2$; under a uniform prior the average error is at least $1/2$. □

This theorem formalizes why an interface may need an external status query, idempotency key, or semantic side channel. Repeating an irreversible action is not a safe diagnostic operation merely because its resulting screenshot would reveal the state.

8 Risk-Sensitive Visual Ambiguity

Let two latent states x_1, x_2 produce the same summary, and suppose the available commitment actions are exactly $\{a_1, a_2\}$. Their state-matched losses are zero and their cross-losses are

$$L(x_1, a_2) = c_1 > 0, \qquad L(x_2, a_1) = c_2 > 0. \tag{35}$$

Theorem 8 (Two-state minimax risk)**.** *The minimax randomized policy selects a_1 with probability*

$$p^* = \frac{c_1}{c_1 + c_2}, \tag{36}$$

and its unavoidable worst-case risk is

$$\mathcal{R}^* = \frac{c_1 c_2}{c_1 + c_2}. \tag{37}$$

Proof. If a_1 is selected with probability p, the statewise risks are

$$R_1(p) = (1 - p)c_1, \qquad R_2(p) = pc_2.$$

The maximum of these two affine functions is minimized at their intersection. Solving $(1-p)c_1 = pc_2$ gives the stated policy and risk. □

The result shows that a loss-aware policy can allocate ambiguity risk but cannot eliminate it. The more costly error receives lower probability, while zero risk requires an informative refinement or a common zero-loss action.

9 Constructed GUI Witnesses

The results are conditional on decision-aliasing states existing. The following are architectural witnesses, not claims about empirical frequency.

Table 1: Constructed visual-aliasing witnesses and task-relevant semantic refinements.

Scenario	Visually identical latent states	Incompatible decision	Useful refinement
Duplicate submission	Request not sent; request accepted but confirmation lost	Submit versus do not resubmit	Transaction status or idempotency key
Hidden focus	Focus is in field A; focus is in field B, with no visible focus ring	Type into A versus avoid typing into B	Focus identifier
Stale collaboration	Displayed version is current; server advanced after render	Edit versus refresh or merge	Version token and conflict state
Permissions	Control appearance is unchanged; authorization differs	Execute versus request permission	Effective permission and failure semantics
Destructive action	Identical "Remove" controls target recoverable versus permanent deletion	Proceed versus require confirmation	Reversibility and target identity
Validation	Form appears unchanged; hidden or off-screen validation differs	Submit versus repair field	Machine-readable validation state
Background dialog	Foreground pixels match; a system modal owns input in one state	Click foreground versus dismiss modal	Window ownership and z-order
Asynchronous loading	Static frame matches; backend operation is pending versus failed	Wait versus retry	Request lifecycle state

9.1 Duplicate submission as a diagnostic POMDP

Let x_0 mean that no payment was accepted and x_1 mean that payment was accepted but the client did not render confirmation. Both states show the same form and "Pay" button. Commitments are a_{pay} and a_{stop}, with

$$\Gamma(x_0) = \{a_{\text{pay}}\}, \qquad \Gamma(x_1) = \{a_{\text{stop}}\}. \tag{38}$$

Clicking "Pay" is not a safe probe because it can duplicate the transaction in x_1. A read-only transaction query with different responses in the two states makes the pair safely separable. An idempotency key changes the loss structure by making repeat submission acceptable in both states; it can therefore create a common action rather than merely reveal more state.

9.2 Focus and state identity

Two screenshots can be pixel-identical when focus styling is absent or outside the cropped frame. The latent focus target determines the effect of a keyboard action. Exposing a stable focus identifier refines the observation without changing the human-visible layout. This example also illustrates why OCR or stronger visual grounding cannot solve every alias: the distinguishing variable is not rendered.

10 Design Implications for Agent-Facing Interfaces

Specify the decision family before exposing state. A minimal semantic channel is goal-relative. The same visual class may be sufficient for browsing but insufficient for exact-once payment. Interface designers should identify the decisions and loss asymmetries that the agent must support before choosing metadata.

Expose transaction lifecycle and idempotency. Exact-once tasks require distinguishing unissued, pending, accepted, rejected, and safe-to-repeat states. Idempotency can reduce decision complexity by making one action acceptable across states that were previously incompatible.

Expose stable identity, not only coordinates. Stable identifiers support state tracking across layout changes and distinguish visually identical controls with different targets or side effects.

Expose focus, selection, validation, permissions, and versioning. These variables often determine action consequences while remaining visually subtle, off-screen, transient, or completely unrendered.

Describe action semantics and reversibility. A machine-readable action schema can expose whether an operation is destructive, reversible, permission-gated, externally consequential, or confirmation-requiring.

Provide safe diagnostic actions. Read-only status checks, previews, dry runs, undo, and confirmation endpoints increase the information available before commitment. Theorem 7 shows that reflection alone cannot overcome a safe-probe channel that is identical across incompatible states.

Treat semantic consistency as an invariant. A richer channel can be stale or wrong. Pixel-structure conflict studies show that agents may defer to structured values even when visual perception is correct (Zhang and Yang, 2026). Semantic metadata therefore requires synchronization, testing, and provenance.

11 Boundary of the Diagnostic Lower Bounds

The probing results require a distinction between diagnostic information actions and arbitrary state-changing actions. In the diagnostic GUI POMDP, probes preserve the latent state and only reveal outcomes. This model covers read-only status queries, inspection actions, previews, and other noncommitting diagnostics. Under this assumption, a terminal transcript can only identify which zero-loss action support contains the unchanged state, which makes the action-cover leaf bound valid.

The same lower bound does not hold for unrestricted state-changing probes. An information action may alter the environment so that states that initially required incompatible commitments later acquire a common safe continuation. For example, an explicit cancel operation could transform both a pending and an accepted workflow into a common resolved state. Such actions remain representable in the full POMDP formulation, but their guarantees depend on reachable post-action beliefs rather than only on the initial action-cover number. The no-safe-probe theorem likewise concerns the stated family of safe probes; it does not rule out an external channel or a state-changing recovery operation outside that family.

12 Scope and Limitations

First, the theory gives structural guarantees over a specified model; it does not measure how often aliases occur in deployed interfaces. Empirical prevalence remains a separate question.

Second, the finite POMDP assumes that the latent state includes all variables required for the Markov property and that transition and observation kernels are well defined. Real software may be nonstationary, adversarial, or only approximately modeled.

Third, decision sufficiency is task- and loss-relative. A visual summary can be sufficient for a low-stakes goal and insufficient for a stricter exact-once or safety-constrained goal.

Fourth, richer semantics introduce privacy and security costs. Exposing permissions, transaction state, or control identity may leak sensitive information or enlarge the attack surface.

Fifth, the exact action-cover number is NP-hard to compute in general and also requires a model of latent states and zero-loss actions. It is primarily a conceptual target for certification, approximation, and counterexample search.

Sixth, much of the mathematical machinery is classical. The paper's value lies in the GUI-specific synthesis, the action-support reduction, and the safe-diagnostic formulation rather than in a new general theory of partial observation.

13 Research Directions

Minimal semantic interface synthesis. Given a finite state model and zero-loss action correspondence, synthesize or approximate a minimum action cover and generate the semantic messages that realize it.

Privacy-constrained observation design. Choose a side channel minimizing disclosure subject to a decision-error constraint:

$$\min_{p(z|x)} I(X;Z) \quad \text{subject to} \quad \mathbb{P}[A \notin \Gamma(X)] \leq \varepsilon. \tag{39}$$

This is a task-oriented rate-distortion problem.

Model checking for visual aliases. A developer tool could search for pairs or sets of application states that render to the same agent observation but have incompatible safe actions. Counterexamples would directly identify missing semantics.

Compositional sufficiency. Determine when component-level semantic interfaces compose into application-level decision sufficiency and when cross-component dependencies defeat local guarantees.

Approximate and robust sufficiency. Le Cam deficiency, approximate state abstraction, and robust POMDP methods may yield quantitative bounds when observation channels are noisy or semantic metadata is stale rather than exactly correct.

14 Conclusion

Visual GUI agents operate under an observation interface chosen by system designers. The relevant question is not whether pixels are generally informative, but whether the complete agent-visible history preserves the decisions required by the task. In a finite-horizon POMDP, a summary supports an optimal policy exactly when all reachable beliefs sharing that summary have a common optimal action. If a summary aliases κ beliefs with pairwise disjoint optimal-action sets, randomization cannot make the policy optimal everywhere, and positive action gaps create unavoidable regret.

At a zero-error commitment point, the minimum deterministic semantic refinement is exactly the action-cover number τ_Γ, an induced minimum set-cover instance. Safe state-preserving probes require logarithmic decision-tree depth in this quantity, noisy probes are limited by information gained per interaction, and a safe probe family that is statistically identical across incompatible states cannot guarantee success. These results support a concrete interface principle: reliable GUI agency requires either a common safe action, an informative semantic distinction, or a safe diagnostic path. More visual reasoning alone cannot recover a distinction that the observation interface never reveals, while idempotency or reversibility can sometimes eliminate the need to reveal it.

References

David Abel, D. Ellis Hershkowitz, and Michael L. Littman. Near optimal behavior via approximate state abstraction. In *Proceedings of the 33rd International Conference on Machine Learning*, 2016.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, second edition, 2006.

Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

Boyu Gou, Ruohan Wang, Boyuan Zheng, Yanan Xie, Cheng Chang, Yiheng Shu, Huan Sun, and Yu Su. Navigating the digital world as humans do: Universal visual grounding for GUI agents. *arXiv preprint arXiv:2410.05243*, 2024.

Jinu Hyeon, Woobin Park, Hongjoon Ahn, and Taesup Moon. Action-sufficient goal representations. *arXiv preprint arXiv:2601.22496*, 2026.

Chao Jin, Wenkui Yang, Hao Sun, Yuqi Liao, Qianyi Jiang, Kai Zhou, Jie Cao, Ran He, and Huaibo Huang. HalluClear: Diagnosing, evaluating and mitigating hallucinations in GUI agents. *arXiv preprint arXiv:2604.17284*, 2026.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.

Jing Yu Koh, Robert Lo, Lawrence Jang, Vikram Duvvur, Ming Chong Lim, Po-Yu Huang, Graham Neubig, Shuyan Zhou, Ruslan Salakhutdinov, and Daniel Fried. VisualWebArena: Evaluating multimodal agents on realistic visual web tasks. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, 2024.

Lucien Le Cam. Sufficiency and approximate sufficiency. *The Annals of Mathematical Statistics*, 35(4):1419–1455, 1964.

Bojie Li and Noah Shi. Agent-computer observation interfaces enable dynamic computer use. *arXiv preprint arXiv:2606.29472*, 2026.

Lihong Li, Thomas J. Walsh, and Michael L. Littman. Towards a unified theory of state abstraction for MDPs. In *Proceedings of the Ninth International Symposium on Artificial Intelligence and Mathematics*, 2006.

Viktor Muryn, Marta Sumyk, Mariya Hirna, Sofiya Garkot, and Maksym Shamrai. Screen2AX: Vision-based approach for automatic macOS accessibility generation. *arXiv preprint arXiv:2507.16704*, 2025.

Yujia Qin, Yining Ye, Junjie Fang, Haoming Wang, Shihao Liang, Shizuo Tian, Junda Zhang, et al. UI-TARS: Pioneering automated GUI interaction with native agents. *arXiv preprint arXiv:2501.12326*, 2025.

Anthony T. Nixon. The Myhill–Nerode theorem for bounded interaction: Canonical abstractions via agent-bounded indistinguishability. *arXiv preprint arXiv:2603.21399*, 2026.

Alon Orlitsky and James R. Roche. Coding for computing. *IEEE Transactions on Information Theory*, 47(3):903–917, 2001.

Yucheng Shi, Wenhao Yu, Jingyuan Huang, Wenlin Yao, Wenhu Chen, and Ninghao Liu. Towards trustworthy GUI agents: A survey. *arXiv preprint arXiv:2503.23434*, 2025.

Richard D. Smallwood and Edward J. Sondik. The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973.

Michito Takeshita, Takuro Kawada, Takumi Ohashi, Shunsuke Kitada, and Hitoshi Iyatomi. A11y-Compressor: A framework for enhancing the efficiency of GUI agent observations through visual context reconstruction and redundancy reduction. *arXiv preprint arXiv:2605.00551*, 2026.

Tianbao Xie, Danyang Zhang, Jixuan Chen, Xiaochuan Li, Siheng Zhao, Ruisheng Cao, Toh Jing Hua, Zhoujun Cheng, Dongchan Shin, Fangyu Lei, et al. OSWorld: Benchmarking multimodal agents for open-ended tasks in real computer environments. In *Advances in Neural Information Processing Systems*, 2024.

Guijia Zhang and Harry Yang. Do GUI agents believe their eyes? Diagnosing state-belief reliance on pixels versus structure. *arXiv preprint arXiv:2607.04334*, 2026.