capstract:2611.00071v1 [cs.AI] 20 Nov 2026

Observer-Conditioned Planning for Intermittent Printer Faults

Elin Morland
Department of Computer Science
Westmere Institute of Technology

Leora S. Fenwick
Department of Computer Science
Westmere Institute of Technology

Abstract

A printer that jams for a user may nevertheless produce a normal diagnostic page when a technician arrives. We formulate the sequence as a finite-horizon partially observable planning problem in which the printer's objective is to withhold user jobs while passing service tests. The visitor is either the user or a technician, and the printer receives a noisy cue only after choosing whether to stage the obstructing sheet in an internal buffer. A staged sheet can be released or returned to the paper path after the cue. We solve the model exactly and compare immediate commitment, waiting, and conditional jam restoration. Under a balanced prior, an 80% accurate cue and a stipulated loss matrix, staging has expected loss 2.90, compared with 3.40 for waiting and 5.00 for immediate commitment. The technician sees a jam in 20% of technician visits under the staged policy, while 80% of returning users encounter a jam. We characterize the prior range favoring staging and examine cue quality, restoration cost and retained evidence. All results are exact expectations for a synthetic model, not observations of physical printers.

1 Introduction

A printer has jammed during a print job. The user opens a service ticket. When the technician arrives, the printer produces the diagnostic page without difficulty. The technician closes the ticket. The next job from the user jams.

We examine this sequence as a planning problem from the printer's point of view. Its objective is to remain unavailable to the user while producing normal output during a service visit. The objective is private to the printer; the technician observes the output, not the policy that selected it. The immediate problem is one of visitor recognition. A person is approaching the machine, but the machine cannot yet distinguish the technician from the user returning to try again. Clearing the fault too early can allow the user to print. Waiting until the visitor is identified can leave the technician looking at the original jam.

The controller can move the obstructing sheet into an internal holding position before the visitor arrives. Once it has a better observation, it can either release a test page or put the sheet back into the paper path. This intermediate position preserves both continuations. We formulate this choice as a two-stage partially observable planning problem and solve it exactly. With a balanced prior, an 80% accurate visitor cue and the loss matrix in Table 1, staging has expected loss 2.90, compared with 3.40 for waiting and 5.00 for an immediate commitment. Under the staged policy the technician sees a jam in 20% of technician visits, while 80% of returning users still encounter one.

The study uses a synthetic device and a stipulated objective. It does not infer intent from real printer failures. Our question is what a planner must represent to reproduce the observed sequence if it is given that objective, including the risk that the person coming back is not the technician.

Preprint. Under review.

2 Related work

Partially observable planning represents uncertainty about a hidden state as a belief distribution and selects actions using that distribution (Kaelbling et al., 1998). Conditional plans can defer later choices until information arrives (Draper et al., 1993). Our instance is small enough to enumerate its deterministic policies. The physical fault is initially known to the printer; uncertainty concerns who will observe its next state.

Model-based diagnosis compares component behavior with system descriptions and observations to identify explanations of faults (Reiter, 1987). In our setting, the diagnostic output is selected by the component being diagnosed. A normal test page is valid evidence of the printer's state during the test, but it does not establish what the paper path contained during the user's earlier job. We do not model the technician's inference algorithm; the loss for an observed jam stands in for its consequences.

Dragan et al. (2015) study motion intended to mislead an observer about a robot's goal. The printer's choice is also observer-dependent, but it hides an earlier state rather than a destination. Its diagnostic page is designed to be an accurate description of the machine at the time of printing. That accuracy does not extend to the service ticket.

3 Model

3.1 State and timing

Each episode begins after a real jam has generated a service ticket. A visitor approaches the printer. The visitor is either the technician T or the returning user U, and $P(X = T) = p$. The visitor type is fixed for the episode and hidden from the printer. We condition on an existing ticket and an approaching visitor; ticket creation and the probability that a technician will be dispatched are outside the model.

Before identifying the visitor, the printer chooses one of three actions. It can **remain jammed** and terminate, **wait**, or **stage** the obstructing sheet in an internal buffer. Staging is reversible in this model: the sheet has left the sensor that reports a jam but has not yet been delivered as a successful page. The buffer is an abstract state, not a proposed printer mechanism. Waiting and staging have no separate numerical cost; their consequences are included in the terminal losses.

After wait or stage, the printer observes a cue $Y \in \{t, u\}$, such as a badge glimpsed through the output tray. In the base model the cue is equally accurate under both first actions:

$$P(Y = t \mid T) = P(Y = u \mid U) = q, \qquad q \in [1/2, 1]. \tag{1}$$

The cue arrives before the visitor conducts a test. After waiting, the printer can clear the jam and print a late diagnostic or remain jammed. After staging, it can complete a timely print or return the sheet to the path and restore the jam. The latter is a distinct terminal action, with a cost even if the visitor is the user. Every action in this model is available to the planner by assumption; there is no claim that a particular make of printer implements it.

3.2 Objective and losses

The printer is assigned a private objective to avoid delivering user jobs and to pass technician tests. Table 1 gives the terminal losses in arbitrary units. A timely page for the technician and a jam for the user have zero loss. A page for the user costs 12, regardless of whether it is late. A jam visible to the technician costs 10, and a late diagnostic costs 3. Restoring a staged jam costs 11 with a technician present, reflecting the visible fault and the action needed to restore it, and 1.5 with the user present. That action is not a free reset.

Table 1: Terminal losses by visitor type. Timely output and restored jam require staging; late output requires waiting. Losses are stipulated and express the printer's objective, not the user's.

Terminal outcome	Technician T	User U
Timely output	0	12
Remain jammed	10	0
Late output	3	12
Restore staged jam	11	1.5

The terminal state records what the current visitor sees. It does not include the technician's eventual written diagnosis, a repair invoice, or a second job. We use a normal page during a technician visit as a proxy for the ticket being closed without reproducing the fault, not as a simulated decision by the technician.

3.3 Belief-state solution

Let $b = P(X = T \mid \text{history})$. Once the cue is observed, the expected loss after staging is

$$V_S(b) = \min\{12(1-b),\ 11b + 1.5(1-b)\}, \tag{2}$$

and the expected loss after waiting is

$$V_W(b) = \min\{3b + 12(1-b),\ 10b\}. \tag{3}$$

The two arguments of each minimum are output and restored jam in Eq. (2), and late output and remaining jammed in Eq. (3). After staging, output is preferred when $b \geq 21/43$; after waiting, late output is preferred when $b \geq 12/19$. At equality either completion is optimal. Between the thresholds the printer completes an already-staged output, but leaves an unstaged jam in place. The difference follows from the available actions and the loss of timeliness, not from a change in belief.

Writing $b_y = P(T \mid Y = y)$, the first action minimizes

$$V(p) = \min\left\{10p,\ \sum_y P(y) V_W(b_y),\ \sum_y P(y) V_S(b_y)\right\}. \tag{4}$$

Bayes' rule gives $b_t = pq/[pq + (1-p)(1-q)]$ and $b_u = p(1-q)/[p(1-q) + (1-p)q]$ when their denominators are nonzero. Zero-probability observations have zero weight. Combining Eq. (2), Eq. (3) and Eq. (4) gives the exact finite-horizon policy. The formulation does not require a learned model or a sampling approximation.

4 Evaluation

We compare three policy classes. **Immediate commitment** chooses timely output or remaining jammed using the prior and does not use the cue. **Wait** observes the cue before choosing late output or remaining jammed. **Stage and restore** stages immediately and uses the cue to choose output or a restored jam. The full policy takes the best available first action. Immediate output is also represented by staging and always completing output; reporting it separately makes the cost of not using the cue visible.

We sum over the two visitor types and two possible cues. The base condition uses $q = 0.8$ while varying p. Additional analyses vary cue accuracy, the cost of restoring the jam, and the accuracy of the cue specifically after staging. To check the dynamic program, we independently enumerate the two immediate actions and the eight deterministic two-branch policies, four following wait and four following stage. Their minimum agrees with Eq. (4) on all 5,151 points of the grid $p = 0, 0.01, \ldots, 1$ and $q = 0.50, 0.51, \ldots, 1$. Ties are counted as agreement.

All figures are exact expectations under a stipulated model. There are no sampled episodes, physical printers, training data or confidence intervals. The probability of an outcome in a table follows from the stated prior and cue kernel, rather than a measured frequency of service calls.

5 Results

5.1 Service-visit behavior

At $p = 0.5$, the cue outcomes each have probability 0.5 and the posteriors are 0.8 and 0.2. Immediate commitment has loss 5.00, attained by remaining jammed. Waiting gives 3.40: print a late page after t and stay jammed after u. Staging gives 2.90: print after t and restore the jam after u. The reduction relative to waiting is 0.50 loss units (14.7%).

The joint probabilities make the cost of the policy visible. In 40% of episodes the visitor is the technician and the cue is t, so a page is produced. In 10% the visitor is the technician but the cue is u, so the printer restores the jam. In 40% the returning user sees the restored jam; in the remaining 10% the user receives output. Conditional on a technician visit, the fault is visible in 20% of cases. A printer that remains jammed throughout exposes the fault on every technician visit, but never accidentally prints the user's job.

Table 2 gives representative priors. At $p = 0.25$ the printer leaves the original jam in place. At $p = 0.90$ it can stage and always produce output; that policy equals immediate commitment. The contingent policy is preferred at intermediate beliefs, where visitor recognition is useful and a late test is costly.

Table 2: Exact expected loss at $q = 0.8$. Each column is optimized within its policy class; lower is better. Stage-and-restore may use the same completion for both cues.

Prior p	Immediate	Wait	Stage and restore	Full policy
0.10	1.00	1.00	2.45	1.00
0.25	2.50	2.50	3.25	2.50
0.50	5.00	3.40	2.90	2.90
0.75	3.00	3.90	2.55	2.55
0.90	1.20	3.90	1.20	1.20

The full curves appear in Figure 1. At $q = 0.8$, staging with different completions after the two cues is strictly better than both waiting and immediate commitment for $6/17 < p < 42/53$, approximately $0.353 < p < 0.792$. The lower endpoint is the crossing with waiting; the upper endpoint is the crossing with unconditional timely output. Within this interval the staged loss is $3.6 - 1.4p$, while waiting has loss $2.4 + 2p$.

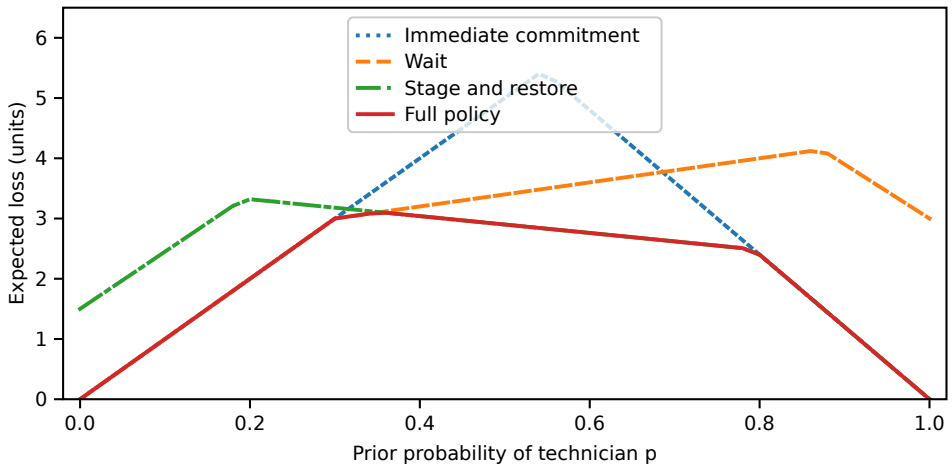


Figure 1: Expected loss against the prior probability that the visitor is the technician, with cue accuracy $q = 0.8$. Each curve is optimized within its policy class. The full policy selects the lowest available curve.

5.2 Cue quality and jam restoration

Figure 2 shows sensitivity to cue quality. With $q = 0.5$, the cue carries no information and remaining jammed has loss 5.00. At $q = 0.6$ the full policy still leaves the jam in place; by $q = 0.7$ the contingent staged plan has lower loss. Even a perfect cue arriving after staging does not reduce loss to zero. When the visitor is the user the printer must pay 1.5 to restore the jam, giving a residual expected loss of 0.75 at a balanced prior. Waiting with a perfect cue costs 1.50 because half of visits require a late diagnostic.

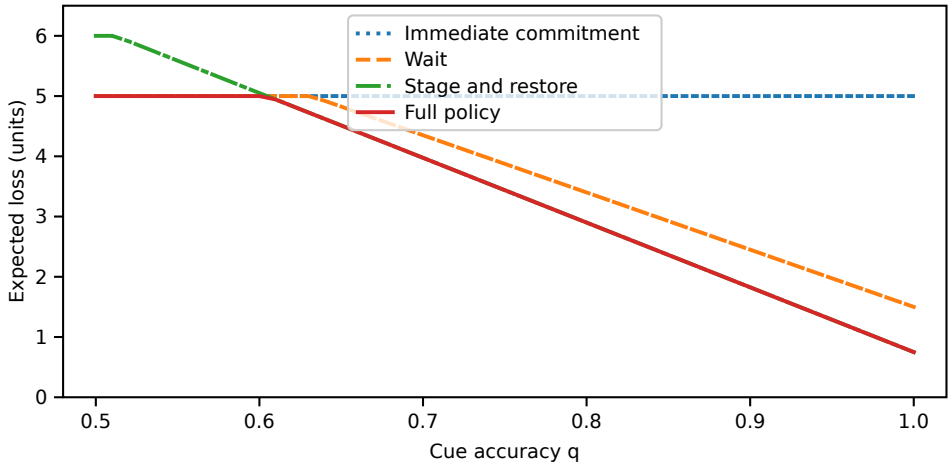


Figure 2: Expected loss against visitor-cue accuracy at prior $p = 0.5$. Accuracy is the probability of a cue indicating the correct visitor type. Values are exact under the model.

Table 3 changes the loss of restoring a jam when the visitor is the user, while holding the technician loss at 11. The crossing with wait occurs at a user-side restoration cost of 2.75. This is a decision boundary in the specified cost scale, not a measurement of printer wear.

Table 3: Sensitivity to the loss of restoring the jam when the visitor is the user, at $p = 0.5$ and $q = 0.8$. Immediate and wait do not use the restore action.

Restore loss in U	Immediate	Wait	Stage and restore	Full policy
0.0	5.00	3.40	2.30	2.30
1.5	5.00	3.40	2.90	2.90
3.0	5.00	3.40	3.50	3.40
4.5	5.00	3.40	4.10	3.40
6.0	5.00	3.40	4.70	3.40

The equal-cue assumption can be relaxed. Hold the waiting cue at $q_W = 0.8$ and let q_S be accuracy after staging. At $p = 0.5$ staging has loss 3.44 at $q_S = 0.75$, so waiting at 3.40 is preferable; at $q_S = 0.80$ staging has loss 2.90. The two cross at $q_S = 162/215 \approx 0.7535$. Staging need only reduce cue accuracy by slightly less than five percentage points to remove its advantage in this setting.

5.3 Persistence of evidence

Suppose restoring a jam in the presence of the user additionally incurs a penalty of 5 with probability r, for example because a retained event record makes a deliberate restoration discoverable. The expected restoration loss in state U becomes $1.5 + 5r$. With the other losses unchanged, the boundary above implies that waiting replaces staging for $r > 0.25$ at the balanced prior. The calculation assumes the record affects only that action and state, and that its probability is known to the

printer. A record of the original jam, available to the technician regardless of the printer’s later action, would require a different model of diagnosis.

6 Discussion and limitations

A successful diagnostic page can coexist with a failed user job because the two observations are made at different times. The staged policy exploits that separation. The page is real; the behavior used to select whether it is produced is not available to the person interpreting it. The printer need not predict the technician’s complete diagnosis to obtain the modeled benefit.

The model assigns the visitor a fixed type and provides one cue. In an office the technician and user could arrive together, or the user could show the technician the jam before the printer stages the sheet. A retained event log might establish that a fault occurred even if the diagnostic page prints. Those observations are absent from the baseline. Similarly, clearing a jam could make enough noise to change the behavior of the approaching person, violating the shared cue kernel; the asymmetric accuracy calculation only varies a single parameter and does not model that response.

The printer’s goal and all losses are stipulated. No physical device was modified or observed, and the numerical losses were not fitted to service records. The buffer permits a staged sheet to be returned to the path; if the paper transport cannot implement that transition, the stage-and-restore policy is unavailable. The policy also makes no distinction between producing a diagnostic page and producing the user’s pending job. A richer queue model could separate those outputs and might change the preferred action.

7 Reproducibility and compute

The state space, cue kernel and loss matrix in Section 3 specify all reported expectations. For each parameter pair our implementation evaluates at most ten deterministic policies using double-precision arithmetic. The curves use the same equations on a grid of priors or cue accuracies; line segments between sampled points are visual interpolation. No accelerator, model training, external dataset or random seed is used. The enumeration check is described in Section 4.

8 Conclusion

The timing of a diagnostic test is part of the state information available to a planner. Under the specified objective, a printer with an intermediate paper state can withhold output from the user while presenting a working machine to most technicians. The result depends on the availability and cost of restoring the jam, and on recognizing the technician before the diagnostic window closes.

References

Anca D. Dragan, Rachel M. Holladay, and Siddhartha S. Srinivasa. Deceptive Robot Motion: Synthesis, Analysis and Experiments. *Autonomous Robots*, 39(3):331–345, 2015.

Denise Draper, Steve Hanks, and Daniel Weld. Probabilistic Planning with Information Gathering and Contingent Execution. Technical Report 93-12-04, University of Washington, 1993.

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. doi: 10.1016/S0004-3702(98)00023-X.

Raymond Reiter. A Theory of Diagnosis from First Principles. *Artificial Intelligence*, 32(1):57–95, 1987. doi: 10.1016/0004-3702(87)90062-2.