MIRROR: ITERATIVE NEURO-SYMBOLIC REASONING FOR HUMAN-LIKE CLINICAL DECISION MAKING

Anonymous authors
Paper under double-blind review

ABSTRACT

Effective clinical decision-making often unfolds as an iterative loop: attention is first allocated to a subset of patient symptoms and findings, symbolic medical reasoning is applied, and the results of reasoning guide subsequent clinical inquiries and inference. We present MIRROR (Multi-Iterative Reasoning with Rules and Observation Refinement), a trainable neuro-symbolic architecture that explicitly models this loopy perception–reasoning dynamic in medical diagnosis. MIRROR alternates between (i) an adaptive perception module that selectively attends to clinical features and patient state based on evolving diagnostic hypotheses, and (ii) a differentiable reasoning module that evaluates symbolic medical rules and clinical predicates in a learned embedding space. Across multiple cycles, predicate-level feedback from reasoning drives refined attention to further patient evidence, yielding an inspectable chain of diagnostic inference. Controlled experiments on three diagnostic benchmarks quantify diagnostic accuracy, path recovery, and evidence acquisition, while cycle-by-cycle traces expose when observation, implication, diagnosis, and knowledge-region routing change. MIRROR therefore provides an explicit computational model of iterative clinical decision making.

1 INTRODUCTION

Effective decision making in complex domains is rarely a one-step process. Consider a physician evaluating a patient with dozens of symptoms and test results. Rather than inspecting all information at once, the physician focuses on a few salient clues, reasons about possible diagnoses, and then revisits the evidence to look for confirming or disconfirming details. *This cycle of perception $\rightarrow$ reasoning $\rightarrow$ re-perception may repeat many times before a confident conclusion is reached.* Such loopy perception–reasoning dynamics allow humans to cope with information overload, adjust attention based on evolving hypotheses, and progressively refine their decisions.

Modeling this process matters. In healthcare, iterative reasoning helps clinicians avoid premature conclusions and uncover rare but critical conditions. In broader decision-making settings—such as scientific discovery, law, or strategic planning—humans rely on similar loops of hypothesis testing and evidence re-examination. Capturing these dynamics in machine learning systems is essential not only for building interpretable and trustworthy models, but also for enabling human–AI collaboration: *if AI systems can reflect how humans reason, they can better explain, anticipate, and complement human decisions*.

This challenge connects to a broader question in machine learning: what internal processes underlie complex behaviors in modern AI models? Recent advances in deep learning have produced systems that rival or even surpass human performance across domains. Yet, while behavioral benchmarks confirm that these models can solve tasks, much less is known about *how* they internally represent and implement the reasoning steps that lead to such behavior. In other words, the cognitive process of decision making remains a black box: we observe inputs and outputs, but not the iterative shifts of attention, hypothesis testing, and reasoning that lie in between. Cognitive science, by contrast, offers a long tradition of describing such intermediate representations and processes—bridging between external behavior and low-level mechanisms—and recent work recovers symbolic cognitive programs from human and animal choice (Castro et al., 2025) and maps novelty-seeking decision loops in the brain (Zhang et al., 2025). Inspired by this perspective, we argue that machine learning should likewise develop a layer of *cognitive interpretability*, where models are designed not just to perform tasks, but to reveal how decisions are formed internally. Our ultimate goal is to make the black cognitive process white—building models that both capture and expose human-like reasoning

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

loops. Such models not only advance interpretability, but also lay the foundation for AI systems that can better understand, anticipate, and collaborate with human decision making.

However, current approaches fall short of this goal. Neural models with attention can extract patterns from complex inputs, and some close a top-down/bottom-up loop over visual regions or scanpaths (Anderson et al., 2018; Yang et al., 2024), but they typically do not rewrite a named symbolic state that later reasoning can inspect. Symbolic systems, by contrast, offer explicit rules that make decisions interpretable, yet they are brittle under noisy or incomplete evidence. Existing neuro-symbolic methods attempt to combine these paradigms, including closed-loop perception–parsing–reasoning for vision and formulae (Li et al., 2020), System 1/System 2 neural-logic stacks (Hua & Zhang, 2022), and differentiable forward chaining (Shindo et al., 2021), but they remain largely static at the clinical-inquiry timescale: perception attends once, reasoning applies once, and the cycle ends. What is missing is a *biologically and cognitively plausible loop where attention and reasoning repeatedly drive one another*, forming an explicit, trainable process of iterative decision making. Such a design is both natural—mirroring how humans iteratively shift focus, test hypotheses, and refine judgments—and challenging: the rule executor can be differentiable, but choosing a named inquiry and obtaining its reply are discrete events whose supervision and causal ordering must be stated explicitly. Further discussion of related approaches appears in Appendix A.

To address this gap, we introduce **MIRROR** (Multi-Iterative Reasoning with Rules and Observation Refinement), a neuro-symbolic architecture that explicitly models decision making as a loopy interplay between perception and reasoning.

The core contributions of this work are:

1. We propose MIRROR, a neuro-symbolic architecture that frames decision making as an internal iterative loop of perception and reasoning—mirroring how humans refine judgments by repeatedly shifting attention and updating hypotheses.
2. We introduce a fully differentiable reasoning module that evaluates symbolic rules in a semantic embedding space, using cosine similarity between an attention-weighted feature representation and learnable predicate embeddings. This enables robust, gradient-based learning and soft semantic alignment, beyond brittle lexical rule matching.
3. We design a dynamic feedback mechanism in which each reasoning cycle generates two levels of guidance for the next: (i) task-level signals that shape what hypotheses to prioritize, and (ii) predicate-level signals that refocus perceptual attention.

2 Problem Setup

Feature Representation We define a *universal predicate set* $\mathcal{P} = \{P_1, \ldots, P_N\}$, where $N = |\mathcal{P}|$. Each P_i is a binary proposition about an atomic feature of the decision context (e.g., the presence or absence of a symptom, or the outcome of a diagnostic test): in the world it is either true or false. At any moment, P_i is either *grounded*, with polarity $\sigma_i \in \{+1, -1\}$ (positive or negative), or *unobserved or unknown* ($\sigma_i = \perp$). The observed human decision is denoted by $y_{\text{human}} \in \{1, \ldots, K\}$, where K is the number of decision classes (e.g., diagnosed diseases).

In addition to outcome supervision, partial information about which cues influenced the decision may be available. We denote such annotations by $\boldsymbol{A}^* \in \{0, 1\}^N$, where $A_i^* = 1$ if predicate P_i was identified as relevant and $A_i^* = 0$ otherwise. These annotations can be derived from auxiliary signals such as physician queries in QA datasets, eye fixations or interaction logs in behavioral studies, or expert markup of decisive findings. When available, $\boldsymbol{A}^*$ provides process-level supervision; otherwise, the model is trained solely on y_{human}, treating the cognitive process as latent.

Latent Cognitive Process and Knowledge Regions While the final outcome y_{human} is observed, the reasoning process that leads to this decision remains hidden. We conjecture that human decision making unfolds through a sequence of *partial groundings* of $\mathcal{P}$, reflecting limited cognitive capacity. Although the full vocabulary $\mathcal{P}$ is in principle accessible, a decision maker grounds only a subset of predicates at a time, forming an evolving chain

$$C_0 \subseteq C_1 \subseteq \cdots \subseteq C_T \subseteq \mathcal{P},$$

where $C_t = \{P_i : \sigma_i^{(t)} \neq \perp\}$ is the set of predicates whose polarities have been grounded (by observation or inference) up to round t, and C_T is the set that ultimately supports the decision.

A second conjecture is that this evidence selection process is guided by structured knowledge. We assume a *universal knowledge base* $\mathcal{K} = \mathcal{K}_{\text{impl}} \cup \mathcal{K}_{\text{diag}}$ of Horn clauses. Every rule has a body that

is a conjunction of grounded literals. The head may be either a predicate or a decision class. Thus, a rule $r \in \mathcal{K}$ concludes a predicate (finding $\rightarrow$ finding) or concludes a class (findings $\rightarrow$ outcome).

$$R_r : \quad \bigwedge_{i \in B_r} (P_i, \pi_{ri}) \longrightarrow (P_{k_r}, \pi_r), \tag{1}$$

where $B_r \subseteq \{1, \ldots, N\}$ indexes the body, $\pi_{ri} \in \{-1, +1\}$ is the polarity required of P_i, k_r is the conclusion slot, and $\pi_r \in \{-1, +1\}$ is the conclusion polarity (the head is P_{k_r} or $\neg P_{k_r}$).

From a cognitive perspective, however, individual humans do not apply the full rule base $\mathcal{K}$. Instead, each operates within a *knowledge region* $\mathcal{K}^{(h)} \subseteq \mathcal{K}$ by a reliability vector $\boldsymbol{w}^{(h)} \in [0, 1]^R$ over the fixed rules, shaped by their memory, expertise, and attention. Different knowledge regions can guide different grounding trajectories $\{C_t\}_{t=0}^{T}$, explaining why the same universal evidence space may yield distinct reasoning paths across individuals.

Modeling Objective Our objective is to uncover how each human instantiates a *cognitive reasoning path* within it. Specifically, given the final outcome y_{human} (and, when available, auxiliary annotations $\boldsymbol{A}^*$), we aim to reconstruct (i) the latent evidence-selection sequence $\{C_t\}_{t=0}^{T}$ and (ii) the individualized rule subset $\mathcal{K}^{(h)}$ that jointly explain a decision. This formulation allows us to capture potential limitations—such as reliance on an incomplete or biased region of $\mathcal{K}$—and provides a foundation for human-centric AI systems that can complement, remind, or teach human decision makers.

3 THE MIRROR ARCHITECTURE

Our goal is to develop a **white-box model of human decision making**, where the supervision signal is the observed human decision y_{human}. Unlike black-box predictors that directly map *high-dimensional* inputs to outputs, we seek a model that explicitly captures the *iterative perception–reasoning loop* hypothesized to underlie human cognition. In realistic decision-making scenarios, humans face a large set of features but *limited cognitive capacity*. Rather than processing all information at once, they allocate attention to a subset of features, generate intermediate hypotheses, and let these hypotheses guide attention toward new evidence. This creates a *loopy cycle* in which perception directs reasoning, and reasoning in turn reshapes perception, until a confident conclusion is reached. MIRROR is designed to make this hidden process explicit: it models decision making as a dynamic sequence of attention shifts and rule-based reasoning steps, thereby providing not only accurate predictions of y_{human} but also interpretable insight into the intermediate cognitive mechanisms—what features were prioritized, how hypotheses evolved, and what reasoning chains supported the final decision.

For example, after hearing that a patient has a cough, a physician may ask about fever and use the reply to decide whether breathing difficulty is the next useful question. MIRROR records each answer and each inferred finding in the same named state matrix $\boldsymbol{E}$. Its signed values guide the next inquiry and the rule-based diagnosis, yielding a trace of the questions, evidence, rules, and changing hypotheses at every step.

3.1 INITIALIZATION

Predicate universe. To preserve the meaning of a finding across rounds, each predicate in $\mathcal{P}$ keeps a named slot with a unit anchor $\boldsymbol{\theta}_i \in \mathbb{R}^d$. The anchor may be frozen or optimized and renormalized, collected as

$$\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N]^\top \in \mathbb{R}^{N \times d}, \qquad \|\boldsymbol{\theta}_i\|_2 = 1. \tag{2}$$

Patient state. For one patient, the recurrent state $\boldsymbol{E}^{(t)} \in \mathbb{R}^{N \times d}$ records signed truth values; its i-th row $\boldsymbol{E}_i^{(t)}$ corresponds to predicate P_i. Let $o_i^{(t)} \in \{0, 1\}$ indicate whether a definite fact about P_i has been observed and written into this state by round t. When such a fact is available in the patient record, $\sigma_i^{\text{obs}} \in \{-1, +1\}$ denotes its value. The state begins from the chief complaint, where the zero vector means that predicate P_i has not been observed:

$$\boldsymbol{E}_i^{(0)} = \begin{cases} +\boldsymbol{\theta}_i, & o_i^{(0)} = 1,\ \sigma_i^{\text{obs}} = +1, \\ -\boldsymbol{\theta}_i, & o_i^{(0)} = 1,\ \sigma_i^{\text{obs}} = -1, \\ \mathbf{0}, & o_i^{(0)} = 0. \end{cases} \tag{3}$$

Its signed coordinate is $s_i^{(t)} = \langle \boldsymbol{E}_i^{(t)}, \boldsymbol{\theta}_i \rangle \in [-1, 1]$. Because $\boldsymbol{\theta}_i$ has unit length, this inner product is the signed scalar projection of $\boldsymbol{E}_i^{(t)}$ onto the predicate's anchor. The update in Section 3.2.1 preserves $\boldsymbol{E}_i^{(t)} = s_i^{(t)} \boldsymbol{\theta}_i$, so this projection exactly recovers the truth value carried by row i. One scalar per named predicate is sufficient to carry the evidence through the loop. Observed positive and negative facts are fixed at $+1$ and -1; implication can assign an unread predicate a graded value in $(-1, 1)$. Both kinds of evidence therefore occupy the same matrix, while the observation mask records which values were directly observed.

3.2 Perception–Reasoning Loop

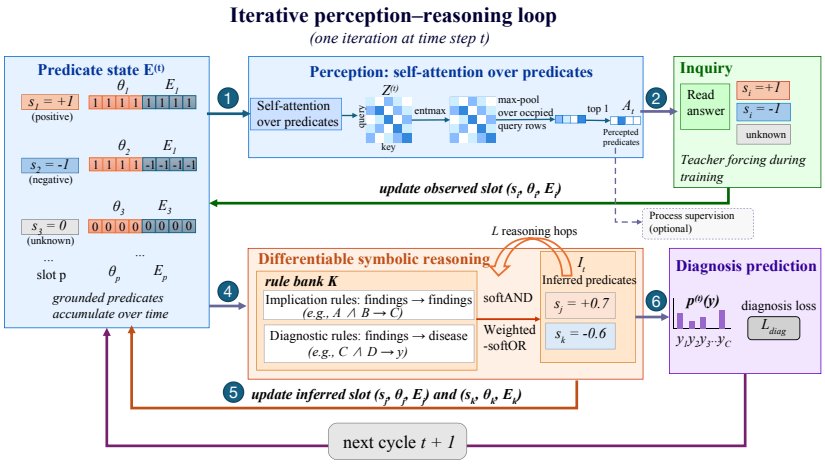


Figure 1: One perception–reason cycle of signed-state MIRROR. Occupied slots of $\boldsymbol{E}^{(t-1)}$ attend over unread predicates; the selected inquiry A_t writes one definite reply; L implication hops update inferred coordinates I_t (the diagram shows $L = 1$); diagnostic rules produce $\widehat{\boldsymbol{y}}^{(t)}$. Evaluation then stops the loop if confidence is high, or $t = T$.

The model keeps all N named slots and runs one perception–reason step per inquiry round (Figure 1). Step t reads the current state $\boldsymbol{E}^{(t-1)}$, selects one named question, writes its definite reply, applies L implication hops, and diagnoses from the updated state. The new conclusions are available to the next inquiry. At evaluation, the loop stops when its current diagnosis meets the confidence criterion in Section 3.2.3, or when the query budget is exhausted.

Sparse Attention. The next question or inquiry for doctors should reflect what has already been learned about the patient. For example, a finding inferred after the fever reply may make a question about breathing difficulty more useful than it was initially. Sparse attention is thus based on the current state including both observed and inferred findings to get the next inquiry.

Each slot needs to preserve both *which* named predicate it represents and its current signed value. We therefore concatenate the current unit anchor with the signed state vector: $\boldsymbol{h}_i^{(t)} = \left[\boldsymbol{\theta}_i \| \boldsymbol{E}_i^{(t-1)}\right] \in \mathbb{R}^{2d}$, with $d_a = 2d$. Here $d_a = 2d$ is the attention dimension. Learned matrices $\boldsymbol{W}_Q, \boldsymbol{W}_K \in \mathbb{R}^{d_a \times 2d}$ send each slot into a shared query/key space.

Query rows are the occupied slots $Q_t = O_{t-1} \cup I_{t-1}$, whereas keys are legal, unread, and not-yet-queried predicates, whose index set we denote by $\mathcal{L}_t$. A look may ask outside U, in which case the simulator returns unknown and the slot is marked queried without recording an observation.

$$\boldsymbol{q}_i^{(t)} = \boldsymbol{W}_Q \boldsymbol{h}_i^{(t)}, \qquad \boldsymbol{k}_j^{(t)} = \boldsymbol{W}_K \boldsymbol{h}_j^{(t)}, \qquad Z_{ij}^{(t)} = \frac{\boldsymbol{q}_i^{(t)\top} \boldsymbol{k}_j^{(t)}}{\sqrt{d_a}\, \tau_{\text{att}}}. \tag{4}$$

Here $\boldsymbol{q}_i^{(t)}, \boldsymbol{k}_j^{(t)} \in \mathbb{R}^{d_a}$, and the fixed $\tau_{\text{att}} > 0$ controls attention sharpness. Illegal query rows and keys are masked before normalization. We replace the usual row-wise softmax in self-attention by row-wise entmax(Martins & Astudillo, 2016; Peters et al., 2019),

$$\boldsymbol{M}_{i,:}^{(t)} = \text{entmax}_{1.5}\left(\boldsymbol{Z}_{i,:}^{(t)}\right), \quad i \in Q_t, \qquad m_j^{(t)} = \max_{i \in Q_t} M_{ij}^{(t)}, \qquad z_j^{(t)} = \max_{i \in Q_t} Z_{ij}^{(t)}. \tag{5}$$

Each occupied fact first distributes sparse attention over its legal unread keys, after which the column-wise maximum lets any occupied fact nominate P_j. The sparse pre-entmax pooled score

$\boldsymbol{z}^{(t)}$ is retained only for the auxiliary process loss in Equation 18; autonomous execution is governed by $\boldsymbol{m}^{(t)}$. The inquiry is the single highest-scoring legal key,

$$A_t = \{j_{(1)}\}, \qquad j_{(1)} = \arg\max_{j \in \mathcal{L}_t} m_j^{(t)}. \tag{6}$$

Each step therefore asks one named predicate, and that choice remains visible in the patient trace.

Asking and answering are separate. U holds definite signed facts in the complete record; a fact enters the state only after the corresponding inquiry (unknown replies do not write). The mask is $o_j^{(t)} = o_j^{(t-1)} \vee \mathbf{1}[j \in A_t \cap U]$. Diagnosis training uses a straight-through estimator (STE) on the hard inquiry. Let $a_{tj} = \mathbf{1}[j \in A_t]$ and $\bar{m}_{tj} = a_{tj} m_j^{(t)}$. The gate is

$$g_{tj} = \bar{m}_{tj} + \mathrm{sg}[a_{tj} - \bar{m}_{tj}], \tag{7}$$

with forward value $g_{tj} = a_{tj}$. Let $u_j = \mathbf{1}[j \in U]$ and let $\sigma_{j,\mathrm{dir}}^{(t)}$ store the last observed sign ($\sigma_{j,\mathrm{dir}}^{(0)} = o_j^{(0)} \sigma_j^{\mathrm{obs}}$). $\boldsymbol{E}_j^{(t,0)}$ is the state after the reply and before implication:

$$\begin{aligned} \sigma_{j,\mathrm{dir}}^{(t)} &= \sigma_{j,\mathrm{dir}}^{(t-1)} + u_j g_{tj} \big(\sigma_j^{\mathrm{obs}} - \sigma_{j,\mathrm{dir}}^{(t-1)}\big), \\ \boldsymbol{E}_j^{(t,0)} &= \boldsymbol{E}_j^{(t-1)} + u_j g_{tj} \big(\sigma_j^{\mathrm{obs}} \boldsymbol{\theta}_j - \boldsymbol{E}_j^{(t-1)}\big). \end{aligned} \tag{8}$$

In the forward pass $g_{tj} = a_{tj}$: a selected definite reply overwrites the slot with $\sigma_j^{\mathrm{obs}} \boldsymbol{\theta}_j$. The gradient through g_{tj} trains the inquiry scores from diagnosis loss.

The reconstructed path is the occupied state after the observation and implication, $C_t = \{P_i \in \mathcal{P} : |s_i^{(t)}| > \varepsilon\} = O_t \cup I_t$, with $C_0 = O_0$. An unread or queried-but-unknown slot remains $\boldsymbol{E}_j = \mathbf{0}$, so $s_j = 0$. Inferred predicates (I_t) enter C_t only when implication writes their state, and we then discuss how to get inferred predicates by neural symbolic rule forward chaining.

Rule Forward Chaining. The recorded reply becomes the starting point for reasoning. Rules read the current truth values and can add graded support for findings that have not been directly observed. One implication hop evaluates the maps below on the occupied state; the hop is repeated L times.

A signed literal asks whether the current coordinate has the polarity the rule requires. For a body premise (P_i, π_{ri}) of rule r,

$$x_{ri}^{(t)} = \mathrm{ReLU}\left(\pi_{ri} s_i^{(t)}\right) \in [0, 1]. \tag{9}$$

The product $\pi_{ri} s_i$ is positive when the observed or inferred sign matches the required polarity. ReLU keeps that aligned magnitude and zeros a mismatch. For example, if P_i denotes fever, a recorded absence gives $s_i = -1$ and fully satisfies a rule requiring no fever ($\pi_{ri} = -1$). A positive fever reply or an unread fever slot gives that negative premise zero support.

Hard AND/OR and their soft surrogates. A Horn body is $\min_{i \in B_r} x_{ri}$ classically; several rules sharing a head are a $\max$. We use a normalized SoftMin and a weighted log-sum-exp so partial evidence remains differentiable:

$$a_r^{(t)} = -\tau_\wedge \log\left(\frac{1}{|B_r|} \sum_{i \in B_r} \exp\left(-\frac{x_{ri}^{(t)}}{\tau_\wedge}\right)\right). \tag{10}$$

If every premise equals c, then $a_r = c$. As $\tau_\wedge \to 0^+$ SoftMin recovers $\min_{i \in B_r} x_{ri}$ (hard AND); we use $\tau_\wedge = 0.1$. Several rules may support the same signed head, each with reliability $w_r \in (0, 1]$ (Section 3.2.2). The soft-OR is

$$\mathrm{softOR}_{\mu_\vee}(\mathcal{R}; a, w) = \frac{1}{\mu_\vee} \log\left(\frac{1}{|\mathcal{R}|} \sum_{r \in \mathcal{R}} \exp\big(\mu_\vee w_r a_r\big)\right) \tag{11}$$

Both factors lie in $[0, 1]$. As $\mu_\vee \to \infty$ the operator recovers $\max_r(w_r a_r)$; we use $\mu_\vee = 10$.

Let $\mathcal{R}_{k,+}$ and $\mathcal{R}_{k,-}$ contain implication rules with positive and negative conclusions for P_k. We aggregate each polarity into a head conclusion. A positive net value supports P_k, while a negative one supports its negation; opposing rules reduce the net support:

$$v_k^{(t)} = \mathrm{softOR}_{\mu_\vee}\big(\mathcal{R}_{k,+}; a^{(t)}, w\big) - \mathrm{softOR}_{\mu_\vee}\big(\mathcal{R}_{k,-}; a^{(t)}, w\big) \tag{12}$$

3.2.1 State Update with Observed Evidence

For the observed answer, we get the definite value from the patient record. For an unread predicate, the net proposal v_k instead moves its truth value toward the supported polarity. Its magnitude controls how far the state moves, so weak support remains visible without being treated as a definite

observed fact. Let $\boldsymbol{E}^{(t,0)}$ be the state after the inquiry write. Each implication hop then updates the same rows while keeping observed values fixed:

$$\overline{\boldsymbol{E}}_k^{(t,\ell)} = (1 - |v_k^{(t,\ell)}|)\boldsymbol{E}_k^{(t,\ell-1)} + v_k^{(t,\ell)}\boldsymbol{\theta}_k, \qquad \boldsymbol{E}_k^{(t,\ell)} = o_k^{(t)}\sigma_{k,\mathrm{dir}}^{(t)}\boldsymbol{\theta}_k + (1 - o_k^{(t)})\overline{\boldsymbol{E}}_k^{(t,\ell)}, \quad (13)$$

for hops $\ell = 1, \ldots, L$, with $v^{(t,\ell)}$ computed from $\boldsymbol{E}^{(t,\ell-1)}$. Equivalently, writing $\mathrm{Imp}^{\circ L}$ for the L-fold composition of one hop ($\mathrm{Imp} \circ \cdots \circ \mathrm{Imp}$, L times), $\boldsymbol{E}^{(t)} = \mathrm{Imp}^{\circ L}\big(\boldsymbol{E}^{(t,0)}; o^{(t)}\big) = \boldsymbol{E}^{(t,L)}$. Later hops may fire rules whose premises were only inferred earlier. Observed slots keep $\sigma_{k,\mathrm{dir}}^{(t)} = \sigma_k^{\mathrm{obs}}$ in the forward pass. An unread hop is $s = (1 - |v|)s_{\mathrm{prev}} + v$, so every row stays on $[-\boldsymbol{\theta}_i, +\boldsymbol{\theta}_i]$.

3.2.2 Knowledge profile learning

The shared bank is the same for every case; clinicians need not trust every rule equally. MIRROR stores a reliability w_r per rule and lets the current state select which reliability vector is active.

Evidence-conditioned state encoder. The number and order of predicate rows are fixed within a dataset, so $\boldsymbol{E}^{(t)} \in \mathbb{R}^{N \times d}$ has a fixed shape even though its entries change after each query and reasoning update. With row-major vectorization in the fixed order $\mathcal{P} = (P_1, \ldots, P_N)$, the encoder input is $\boldsymbol{x}_t = \mathrm{vec}\big(\boldsymbol{E}^{(t)}\big) \in \mathbb{R}^{Nd}$. We use a small MLP as the encoder. It maps the complete state to one candidate reliability for every named rule:

$$\boldsymbol{h}_t = \mathrm{GELU}(\boldsymbol{W}_{\mathrm{enc}}\boldsymbol{x}_t + \boldsymbol{b}_{\mathrm{enc}}) \in \mathbb{R}^{d_h}, \qquad \boldsymbol{u}_t = \mathrm{sigmoid}(\boldsymbol{W}_{\mathrm{prop}}\boldsymbol{h}_t + \boldsymbol{b}_{\mathrm{prop}}) \in (0,1)^R, \quad (14)$$

where $d_h = 64$ and $R = |\mathcal{K}|$. Thus $\boldsymbol{u}_t$ is a proposal used only to select a knowledge region.

Discrete VQ knowledge regions. The continuous proposal $\boldsymbol{u}_t$ is mapped to one of a small number of complete rule-use patterns, making the selected reliabilities inspectable at each step. The estimator stores a profile weight matrix $\boldsymbol{W}_{\mathrm{K}} = [\boldsymbol{w}^{(1)}, \ldots, \boldsymbol{w}^{(H)}]^\top \in [0,1]^{H \times R}$, where H is the number of doctor profiles or knowledge regions and $\boldsymbol{w}^{(h)}$ is the complete reliability vector for the profile h. At every loop step it chooses the nearest code under mean squared rule-weight distance,

$$h_t = \arg\min_{h \in \{1,\ldots,H\}} \frac{1}{R}\|\boldsymbol{u}_t - \boldsymbol{w}^{(h)}\|_2^2, \qquad \boldsymbol{w}_t^{\mathrm{eff}} = \boldsymbol{u}_t + \mathrm{sg}[\boldsymbol{w}^{(h_t)} - \boldsymbol{u}_t], \quad (15)$$

where h_t is the selected code index and $\mathrm{sg}[\cdot]$ is stop-gradient (identity forward, zero derivative). Equation 15 executes one discrete region in the forward pass while passing task gradients to the proposal encoder. Because $\boldsymbol{E}^{(t)}$ changes after a reveal and subsequent reasoning, the same case may switch regions over its trajectory. The codebook is updated by EMA with decay 0.99.

Prediction. Diagnostic rules use the same SoftMin/soft-OR as implication. With $\mathcal{R}_{c,\pm}^{\mathrm{diag}}$ the diagnostic rules for class c,

$$\ell_c^{(t)} = \mathrm{softOR}_{\mu_\vee}\big(\mathcal{R}_{c,+}^{\mathrm{diag}}; a^{(t)}, w\big) - \mathrm{softOR}_{\mu_\vee}\big(\mathcal{R}_{c,-}^{\mathrm{diag}}; a^{(t)}, w\big), \qquad \widehat{\boldsymbol{y}}^{(t)} = \mathrm{softmax}(\boldsymbol{\ell}^{(t)}). \quad (16)$$

3.2.3 Confidence stopping

A temperature $\tau > 0$ is fit on development full-rollout logits. The top-1 posterior is $\gamma^{(t)} = \max_c \mathrm{softmax}(\boldsymbol{\ell}^{(t)}/\tau)_c$. Evaluation asks another question only if $\gamma^{(t)} < \tau_\star$, a legal unread predicate remains, and $t < T$; $\tau_\star$ is chosen on Dev after training.

3.3 Training Objective

Training connects the final diagnosis back to the sequence of inquiries that produced it. The diagnosis trajectory uses the observed class $y_{human} \in \{1, \ldots, K\}$ after its final state:

$$\mathcal{L}_{\mathrm{diag}} = -\log \mathrm{softmax}(\boldsymbol{\ell}^{(T)})_{y_{\mathrm{human}}}. \quad (17)$$

The hard action has no ordinary derivative to the query scorer. With STE, Equation 7 passes a surrogate gradient from this loss through the recorded reply and subsequent reasoning.

When reference questions are available, a separate teacher-forced training trajectory provides process supervision. At each step, the query policy scores the current state *before* the reference action is written; that action then builds the next teacher state. With target distribution q_t over legal named questions,

$$\mathcal{L}_{\mathrm{proc}} = -\frac{1}{|\mathcal{V}|}\sum_{t \in \mathcal{V}}\sum_{j=1}^{N} q_t(j)\log \pi_\phi^{\mathrm{query}}(j \mid C_{t-1}), \quad \boldsymbol{\pi}_\phi^{\mathrm{query}} = \mathrm{softmax}(\boldsymbol{z}^{(t)}). \quad (18)$$

Here $\mathcal{V}$ indexes rounds with a valid reference target; at test time the model selects every inquiry itself.

For the knowledge-region variant, the selected code $\boldsymbol{w}^{(h_t)}$ supplies the reliabilities used by implication and diagnostic rules. Diagnosis CE trains its proposal encoder through the straight-through map in Equation 15. We also use the codebook and commitment terms of VQ-VAE (Van Den Oord et al., 2017), identifying the proposal $\boldsymbol{u}_t$ with the encoder output and the selected rule-weight vector with its nearest embedding:

$$\mathcal{L}_{\mathrm{VQ}} = \|\mathrm{sg}[\boldsymbol{u}_t] - \boldsymbol{w}^{(h_t)}\|_2^2 + \beta\|\boldsymbol{u}_t - \mathrm{sg}[\boldsymbol{w}^{(h_t)}]\|_2^2, \tag{19}$$

$$\mathcal{L} = \mathcal{L}_{\mathrm{diag}} + \lambda_{\mathrm{proc}}\mathcal{L}_{\mathrm{proc}} + \lambda_{\mathrm{VQ}}\mathcal{L}_{\mathrm{VQ}}. \tag{20}$$

The loop is trainable end to end: attention, implication, state updates, and diagnosis pass gradients directly; STE connects the discrete query and region choices to the outcome.

4 Experiments

We evaluate our MIRROR method on three medical datasets: DDXPlus (Fansi Tchango et al., 2022), MZ-4 (Chen et al., 2023), and DiReCT (Wang et al., 2024), to illustrate how our model can learn the human decision process. They all provide clinical decision y_{human}, MZ-4 and DiReCT provide a meaningful reference process C_T^* from auxiliary annotations $\boldsymbol{A}^*$.

4.1 Datasets, baselines, and frozen rules

We first build the rule bank from the dataset, where we construct meaningful medical rules from the dataset (see Section B.3). We then compare our MIRROR with advanced black-box and white-box baselines, including classic machine learning models, transformer-based models, Large Language Models (LLMs), and interpretable rule-based models. Dataset construction, evidence semantics, baseline definitions, and the complete rule-bank are introduced in Appendix B.

4.2 Main comparison

We report the prediction accuracy on the test set in Table 1. For prediction accuracy and F1 score, we compare our MIRROR's decision with ground truth y_{human}. Besides, we also show the diagnosis accuracy over loops in Figure 2. Appendix C.3 records the selected hyperparameters, and Appendix C.4 reports the sensitivity.

Table 1: Main test comparison over three seeds. Accuracy and macro-F1 are percentages. Blue marks the best value among black-box and LLM models, and teal marks the second best in that group. **Bold** marks the best value among MIRROR and the white-box rule-based models. On DiReCT, cat is the disease category and diag is the primary discharge diagnosis (PDD).

	MZ-4		DDXPlus		DiReCT			
Method	Acc.	F1	Acc.	F1	Acc.$^{\mathrm{cat}}$	F1$^{\mathrm{cat}}$	Acc.$^{\mathrm{diag}}$	F1$^{\mathrm{diag}}$
Black-box models and LLM models								
Random Forest	73.33±0.77	71.55±0.80	100.00±0.00	100.00±0.00	81.02±1.60	63.09±2.29	62.50±1.39	40.69±0.94
Transformer	70.14±1.51	68.90±1.52	99.60±0.20	99.60±0.20	14.81±7.13	9.40±4.57	9.26±5.61	4.21±2.74
BERT+MLP	73.72±0.73	72.14±1.30	99.93±0.12	99.93±0.12	58.80±2.89	44.84±1.73	33.33±2.78	22.20±1.99
Qwen	62.32±0.00	61.43±0.00	88.20±0.00	88.16±0.00	54.17±0.00	51.64±0.00	34.72±0.00	24.26±0.00
Llama-3-8B	50.72±0.00	32.97±0.00	67.00±0.00	61.74±0.00	44.44±0.00	46.23±0.00	26.39±0.00	23.84±0.00
DxFormer	68.12±1.33	66.40±1.26	84.00±2.36	83.76±2.37	16.67±7.73	5.87±4.20	12.04±3.21	2.73±1.79
Rule-based models								
BRCG	53.91±0.00	45.84±0.00	99.80±0.00	99.80±0.00	25.00±0.00	13.76±0.00	19.44±0.00	8.23±0.00
LEN	69.57±1.00	68.26±1.49	**100.00±0.00**	**100.00±0.00**	55.56±2.78	33.59±5.49	30.09±6.99	16.26±6.62
DR-Net	29.86±0.00	11.50±0.00	79.27±4.75	80.57±4.44	12.96±0.80	1.28±0.61	8.33±0.00	0.44±0.00
RRL	71.88±2.30	**70.59±2.23**	99.87±0.23	99.87±0.23	47.22±3.67	28.14±3.10	30.09±0.80	14.79±2.17
MIRROR	**72.46±0.29**	69.45±0.37	99.20±0.80	99.20±0.80	**67.59±6.42**	**56.02±4.52**	**50.46±2.89**	**32.29±1.29**

Among the rule-based models, MIRROR has the highest accuracy on MZ-4 and DiReCT. On DDXPlus, some runs reach 100%, the same ceiling as LEN. Macro-F1 is a little lower than the best baseline. The accuracy is comparable to the black-box models. Especially, LLM models do not perform well on medical tasks, showing our neural symbolic design is meaningful and efficient for such specific domain tasks. Compared with existing rule-based models, such baselines mainly focus on learning rules, so they are not able to recover the cognition process. Their performance based on relatively large predicates and rules combination, where most of the rules are not meaningful for such medical tasks.

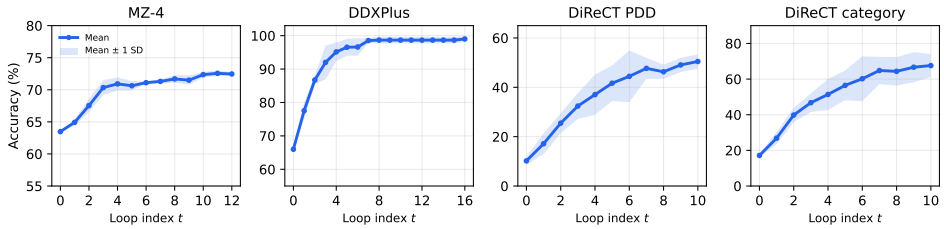


Figure 2: Diagnosis accuracy versus loop times on the three datasets. All three datasets gradually gain accuracy over loops, showing the model is able to collect useful evidence to mimic the human decision process.

For cognition-path learning we report only MZ-4 and DiReCT in Table 2; DDXPlus has no clinician query annotation. The predicted path is $C = O_{\text{stop}} \cup I_{\text{stop}}$, i.e. every named predicate with $|s_i| > \varepsilon$ at the selected stop. The reference $C^\star$ is the case-level set of explicit and clinician-requested findings: `exp_sxs` $\cup$ `imp_sxs` on MZ-4, and C_0 together with physician-tree premises on DiReCT. We also score observation-only overlap O_{stop} versus the same $C^\star$. Precision, recall, and F1 are the micro-averaged set scores with $A \in \{C, O_{\text{stop}}\}$:

$$\mathrm{P} = \frac{|A \cap C^\star|}{|A|}, \qquad \mathrm{R} = \frac{|A \cap C^\star|}{|C^\star|}, \qquad F_1 = \frac{2\mathrm{PR}}{\mathrm{P}+\mathrm{R}}. \tag{21}$$

Table 2: Cognitive path overlap with the gold set $C^\star$ at the selected stop, for the MZ-4 and DiReCT runs in Figure 2. State columns use $C = O_{\text{stop}} \cup I_{\text{stop}}$; Obs. columns use recorded observations O_{stop} only. Entries are mean $\pm$ sample SD over those three runs, as test percentages of Equation 21.

	C vs $C^\star$			O vs $C^\star$		
Dataset	P	R	F1	P	R	F1
MZ-4	99.01±0.02	83.82±0.83	90.78±0.48	100.00±0.00	83.82±0.83	91.19±0.49
DiReCT	15.16±0.52	78.92±1.34	25.43±0.67	82.02±0.25	77.39±2.45	79.62±1.40

The DiReCT state row is low because C includes implication writes I_{stop} that are absent from the physician tree. The observation-only block shows that the model still recovers most clinician-requested facts. On MZ-4 the two blocks nearly coincide. The annotations differ in kind (Appendix B.1): MZ-4 gold slots are named findings copied from the complete patient record, so $C^\star \subseteq U$, whereas DiReCT gold is a physician entailment tree whose premises are mostly latent in the note rather than chart rows.

4.3 Knowledge Regions

Figure 3 visualizes the $H = 4$ rule profiles on MZ-4. The same geometry, rule-gap, and occupancy views for all three datasets appear in Appendix C.2. We show that each dataset has its own rule profile, showing the model is able to learn the meaningful rules for such medical tasks.

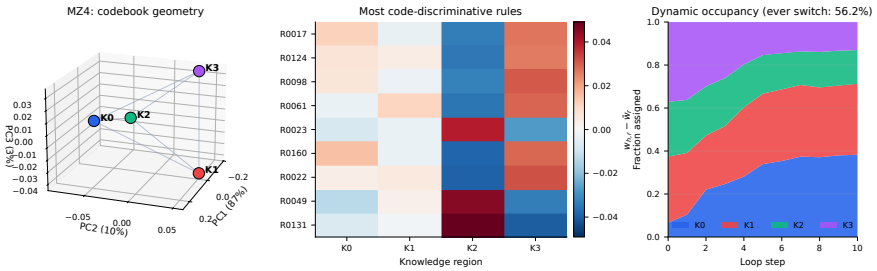


Figure 3: Discrete knowledge regions on MZ-4 ($H = 4$). Left: three-dimensional PCA of the rule-weight matrix $\boldsymbol{W}_K$. Middle: the nine rules with the largest across-code weight ranges; each cell shows that code's rule weight minus the rule's four-code mean, $w_{h,r} - \bar{w}_r$. Right: the fraction of frozen test trajectories assigned to each code at every loop step.

4.4 Ablation studies

Table 3 compares Full MIRROR with three evaluation-time information-flow settings and one retrained setting. *Chief only* diagnoses from C_0 and stops before the first inquiry. *Full patient record* writes every available signed fact at $t = 0$ and does not run the inquiry policy. *No implication* replaces each implication update by the identity. *No process supervision* retrains with $\lambda_{\text{proc}} = 0$. DDXPlus has no process target, so that column repeats Full MIRROR.

Table 3: Ablation study over three seeds.

Dataset	Full MIRROR	Chief complaint only	Full patient record	No implication updates	No process supervision
MZ-4	72.46±0.29	63.48±0.00	70.72±0.58	70.34±1.70	63.48±0.77
DiReCT	50.46±2.89	10.19±1.60	38.43±3.50	25.46±9.25	15.28±2.41
DDXPlus	99.20±0.80	66.00±0.00	98.20±0.00	99.00±0.80	99.20±0.80

Results in Table 3 show that Chief-complaint only is below the full loop on every dataset. The initial facts are not sufficient, and the inquiries collected along the loop change the diagnosis. Writing the full patient record at the start does not replace that loop. It is slightly below the full loop on MZ-4 and DDXPlus, and substantially below it on DiReCT (38.43 versus 50.46). Removing implication updates leaves MZ-4 and DDXPlus near the full-loop accuracy, but drops DiReCT from 50.46 to 25.46. On MZ-4 the reference findings are named slots in the patient record. On DiReCT the physician tree is latent in the note, so implication rules are what place those findings in the state. Removing process supervision lowers MZ-4 from 72.46 to 63.48 and DiReCT from 50.46 to 15.28.

4.5 Case study

We show the loopy process of MIRROR on MZ-4 in Figure 4 (DiReCT and DDXPlus examples appear in Appendix C.1): a query reveals a signed patient fact, implication rules update the accumulated state, diagnostic rules change their activation, and the diagnosis is revised.

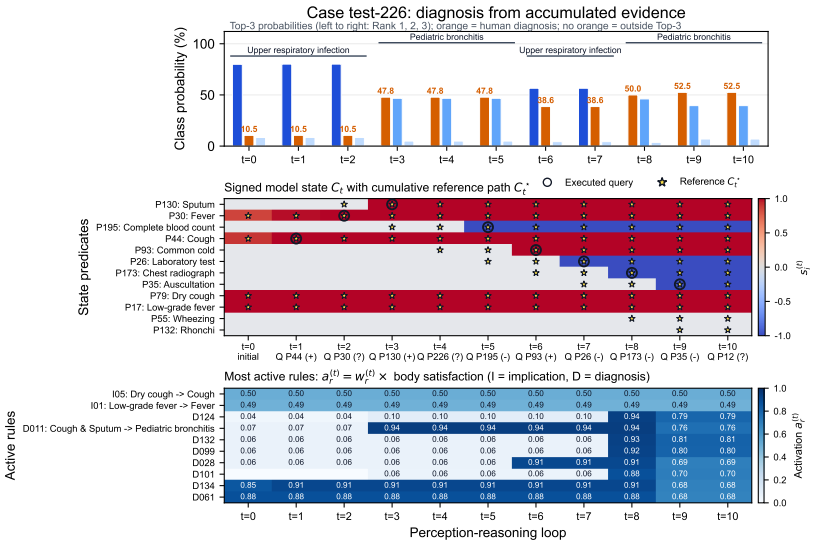


Figure 4: MZ-4 case. The human diagnosis probability rises from 10.5% at $t = 0$ to 52.5% at $t = 10$. The upper panel shows the Top-3 class probabilities at every loop. Orange marks the truth human diagnosis. The middle panel gives the signed model state, stars mark predicates in the cumulative reference path $C_t^\star$, and circles mark executed queries. The lower panel reports the most active implication and diagnostic rules with activation score.

5 Limitations and Conclusion

MIRROR is an elegant loopy system that makes observation, implication, rule execution, and inquiry explicit in one inspectable state. The experiments show that this loop recovers human diagnoses at competitive accuracy, traces a cognition path when annotations exist, and exposes cycle-by-cycle changes in evidence, rules, and hypotheses. One limitation is that the present rule banks are frozen Horn clauses. Structure-learning of $\mathcal{K}$ is left to future work.

AI USE STATEMENT

In this paper, LLMs were used for writing polishing in several paragraphs, like the Experiment section. Besides, LLM were used for baselines results comparison. All the key ideas, proofs, research, and writing are created completely by human authors.

REPRODUCIBILITY STATEMENT

We provide all the data preprocessing and rules construction on the Appendix B. We detailed the hyperparameters selection in Appendix C.3. We will release the code and data for reproducibility in the camera-ready version.

REFERENCES

Peter Anderson, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson, Stephen Gould, and Lei Zhang. Bottom-up and top-down attention for image captioning and visual question answering. In *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 6077–6086. IEEE, 2018.

Pietro Barbiero, Gabriele Ciravegna, Francesco Giannini, Pietro Lió, Marco Gori, and Stefano Melacci. Entropy-based logic explanations of neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 6046–6054, 2022.

Kexin Gu Baugh, Luke Dickens, and Alessandra Russo. Neural dnf-mt: A neuro-symbolic approach for learning interpretable and editable policies. *arXiv preprint arXiv:2501.03888*, 2025.

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

Pablo Samuel Castro, Nenad Tomasev, Ankit Anand, Navodita Sharma, Rishika Mohanta, Aparna Dev, Kuba Perlin, Siddhant Jain, Kyle Levin, Noémi Éltető, et al. Discovering symbolic cognitive models from human and animal behavior. *bioRxiv*, pp. 2025–02, 2025.

Wei Chen, Cheng Zhong, Jiajie Peng, and Zhongyu Wei. Dxformer: a decoupled automatic diagnostic system based on decoder–encoder transformer with dense symptom representations. *Bioinformatics*, 39(1):btac744, 2023.

Nuri Cingillioglu and Alessandra Russo. pix2rule: End-to-end neuro-symbolic rule learning. *arXiv preprint arXiv:2106.07487*, 2021.

Daniel Cunnington, Mark Law, Jorge Lobo, and Alessandra Russo. Ffnsl: feed-forward neural-symbolic learner. *Machine Learning*, 112(2):515–569, 2023.

Rajarshi Das, Shehzaad Dhuliawala, Manzil Zaheer, Luke Vilnis, Ishan Durugkar, Akshay Krishnamurthy, Alex Smola, and Andrew McCallum. Go for a walk and arrive at the answer: Reasoning over paths in knowledge bases using reinforcement learning. *arXiv preprint arXiv:1711.05851*, 2017.

Sanjeeb Dash, Oktay Gunluk, and Dennis Wei. Boolean decision rules via column generation. *Advances in neural information processing systems*, 31, 2018.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.

Arsene Fansi Tchango, Rishab Goel, Zhi Wen, Julien Martel, and Joumana Ghosn. Ddxplus: A new dataset for automatic medical diagnosis. *Advances in neural information processing systems*, 35: 31306–31318, 2022.

Chuqin Geng, Anqi Xing, Li Zhang, Ziyu Zhao, Yuhe Jiang, and Xujie Si. Neurologic: From neural representations to interpretable logic rules. *arXiv preprint arXiv:2501.08281*, 2025.

Wenyue Hua and Yongfeng Zhang. System 1+ system 2= better world: Neural-symbolic chain of logic reasoning. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 601–612, 2022.

Jinu Lee and Wonseok Hwang. Symba: Symbolic backward chaining for structured natural language reasoning. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pp. 2468–2484, 2025.

Qing Li, Siyuan Huang, Yining Hong, Yixin Chen, Ying Nian Wu, and Song-Chun Zhu. Closed loop neural-symbolic learning via integrating neural perception, grammar parsing, and symbolic reasoning. In *International Conference on Machine Learning*, pp. 5884–5894. PMLR, 2020.

Lihui Liu, Zihao Wang, and Hanghang Tong. Neural-symbolic reasoning over knowledge graphs: A survey from a query perspective. *ACM SIGKDD Explorations Newsletter*, 27(1):124–136, 2025.

Zhihao Ma, Yuzheng Zhuang, Paul Weng, Hankz Hankui Zhuo, Dong Li, Wulong Liu, and Jianye Hao. Learning symbolic rules for interpretable deep reinforcement learning. *arXiv preprint arXiv:2103.08228*, 2021.

Robin Manhaeve, Sebastijan Dumancic, Angelika Kimmig, Thomas Demeester, and Luc De Raedt. Deepproblog: Neural probabilistic logic programming. *Advances in neural information processing systems*, 31, 2018.

Jiayuan Mao, Chuang Gan, Pushmeet Kohli, Joshua B Tenenbaum, and Jiajun Wu. The neuro-symbolic concept learner: Interpreting scenes, words, and sentences from natural supervision. *arXiv preprint arXiv:1904.12584*, 2019.

Andre Martins and Ramon Astudillo. From softmax to sparsemax: A sparse model of attention and multi-label classification. In *International conference on machine learning*, pp. 1614–1623. PMLR, 2016.

Parth Padalkar, Jaeseong Lee, Shiyi Wei, and Gopal Gupta. Improving interpretability and accuracy in neuro-symbolic rule extraction using class-specific sparse filters. *arXiv preprint arXiv:2501.16677*, 2025.

Ben Peters, Vlad Niculae, and André FT Martins. Sparse sequence-to-sequence models. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 1504–1519, 2019.

Litao Qiao, Weijia Wang, and Bill Lin. Learning accurate and interpretable decision rule sets from neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 4303–4311, 2021.

Wendi Ren, Ke Wan, Junyu Leng, and Shuang Li. Inferring the invisible: Neuro-symbolic rule discovery for missing value imputation. In *International Conference on Learning Representations*, volume 2026, pp. 153114–153149, 2026.

Tim Rocktäschel and Sebastian Riedel. Learning knowledge base inference with neural theorem provers. In *Proceedings of the 5th workshop on automated knowledge base construction*, pp. 45–50, 2016.

Tim Rocktäschel and Sebastian Riedel. End-to-end differentiable proving. *Advances in neural information processing systems*, 30, 2017.

Hikaru Shindo, Devendra Singh Dhami, and Kristian Kersting. Neuro-symbolic forward reasoning. *arXiv preprint arXiv:2110.09383*, 2021.

Aaron Van Den Oord, Oriol Vinyals, et al. Neural discrete representation learning. *Advances in neural information processing systems*, 30, 2017.

Bowen Wang, Jiuyang Chang, Yiming Qian, Guoxin Chen, Junhao Chen, Zhouqiang Jiang, Jiahao Zhang, Yuta Nakashima, and Hajime Nagahara. Direct: Diagnostic reasoning for clinical notes via large language models. *Advances in neural information processing systems*, 37:74999–75011, 2024.

Zhuo Wang, Wei Zhang, Ning Liu, and Jianyong Wang. Scalable rule-based representation learning for interpretable classification. *Advances in Neural Information Processing Systems*, 34:30479–30491, 2021.

Zhongyu Wei, Qianlong Liu, Baolin Peng, Huaixiao Tou, Ting Chen, Xuan-Jing Huang, Kam-Fai Wong, and Xiang Dai. Task-oriented dialogue system for automatic diagnosis. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 201–207, 2018.

Zhibo Yang, Sounak Mondal, Seoyoung Ahn, Ruoyu Xue, Gregory Zelinsky, Minh Hoai, and Dimitris Samaras. Unifying top-down and bottom-up scanpath prediction using transformers. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1683–1693. IEEE, 2024.

Shuo Zhang, Yan Tian, Quanying Liu, and Haiyan Wu. The neural correlates of novelty and variability in human decision-making under an active inference framework. *Elife*, 13:RP92892, 2025.

Wenqi Zhang, Kai Zhao, Peng Li, Xiao Zhu, Yongliang Shen, Yanna Ma, Yingfeng Chen, and Weiming Lu. A closed-loop perception, decision-making and reasoning mechanism for human-like navigation. *arXiv preprint arXiv:2207.11901*, 2022.

A Related Work

A.1 Neuro-symbolic inference

Neuro-symbolic methods connect learned representations to explicit inference in different ways. Neural theorem provers make symbol matching differentiable within backward chaining (Rocktäschel & Riedel, 2016; 2017), while DeepProbLog combines neural predicates with probabilistic logic programs (Manhaeve et al., 2018). With visual inputs, the Neuro-Symbolic Concept Learner executes programs parsed from questions over learned scene representations (Mao et al., 2019), and NSFR applies forward chaining to perceived facts (Shindo et al., 2021). Other approaches learn to traverse relational paths in knowledge graphs (Das et al., 2017) or combine neural and logical reasoning stages (Hua & Zhang, 2022); a broader account of this line of work is given by Liu et al. (2025). These methods establish different mechanisms for combining learned representations with explicit inference. MIRROR builds on this general paradigm, but places symbolic inference inside an iterative loop in which the evolving named state also guides the acquisition of subsequent evidence.

A.2 Feedback-guided evidence acquisition

Task context can guide what a model processes next. Question-guided attention selects image regions for visual question answering (Anderson et al., 2018), while task-conditioned scanpath models predict successive visual fixations (Yang et al., 2024). More explicit closed-loop systems feed symbolic reasoning back into perception (Li et al., 2020); during inference over text, SymBa asks a language model for facts or rules needed by its backward-chaining solver (Lee & Hwang, 2025). Navigation systems also connect perception, decisions, and reasoning across steps (Zhang et al., 2022). These systems illustrate different forms of feedback; in clinical inquiry, feedback determines which patient finding becomes available next.

Clinical diagnosis makes evidence acquisition part of the decision process. The MZ-4 dialogue subset separates explicit from implicit symptoms (Wei et al., 2018); DDXPlus provides synthetic signed evidence and differential diagnoses without a clinician query trajectory (Fansi Tchango et al., 2022); and DiReCT annotates clinical notes with physician reasoning trees (Wang et al., 2024). DxFormer learns symptom inquiry and diagnosis as coupled stages (Chen et al., 2023). In MIRROR, each discrete inquiry adds a new evidence to the clinical state, and the state is updated by differentiable rule-based inference before the next inquiry.

A.3 Rule learning and evidence-conditioned rule weighting

Interpretable rule models have primarily focused on learning or extracting the rule structure. BRCG searches for compact Boolean rules (Dash et al., 2018); DR-Net and RRL learn interpretable rule sets for classification (Qiao et al., 2021; Wang et al., 2021). LEN extracts logical explanations from concept activations (Barbiero et al., 2022), while FFNSL and pix2rule learn rules from perceptual inputs (Cunnington et al., 2023; Cingillioglu & Russo, 2021). Related methods also derive rules from neural representations (Geng et al., 2025; Padalkar et al., 2025) or apply symbolic rules to reinforcement-learning policies (Ma et al., 2021; Baugh et al., 2025). MIRROR keeps a named clinical clause bank fixed and learns state-dependent rule reliabilities, so individual rules contribute with different strength as evidence accumulates.

Vector quantization represents variation through a finite set of learned codes (Van Den Oord et al., 2017). In MIRROR, the rule structure remains fixed, while rule weights are conditioned on the evidence accumulated over successive inquiries.

B Dataset, baseline, and rule-bank details

B.1 Prepared datasets and evidence semantics

Table 4 records the exact prepared splits. Fit is used for parameter updates, fit-only rule construction, and VQ initialization; Dev is used for early stopping and policy calibration; Test is evaluated once after those choices are frozen.

Table 4: Datasets in the declared protocol.

Dataset	$\lvert\mathcal{P}\rvert$	Cls.	Fit	Dev/Test	Evidence / reference process
MZ-4 (Wei et al., 2018)	230	4	1,110	278/345	signed patient evidence; implicit-symptom serialization
DDXPlus (Fansi Tchango et al., 2022)	987	5	1,000	250/500	closed-world signed sheet; no clinician trajectory
DiReCT (Wang et al., 2024)	830	62	364	75/72	note-grounded patient evidence; physician entailment tree

MZ-4. This Chinese pediatric dialogue subset contains upper respiratory infection, bronchitis, indigestion, and diarrhea. C_0 is the released explicit signed finding set. The available patient facts U contain only definite status-0/1 findings; querying outside U returns unknown and does not write an observed value to the state. Process supervision uses the released implicit-symptom insertion order, a serialization of implicit findings.

DDXPlus. We use the prepared five-condition slice (anemia, initial HIV infection, localized edema, URTI, and viral pharyngitis). C_0 is `INITIAL_EVIDENCE`; every remaining predicate has a closed-world positive or negative reply. The DDXPlus experiment uses no process CE.

DiReCT. MIMIC-IV-Ext-DiReCT supplies clinical notes, physician entailment trees, and category-specific diagnostic graphs. Lexical grounding initializes C_0 from the chief-complaint text and defines the available patient facts from the full note. Both clinical predicates and `DX::` graph nodes are legal actions, except PDD diagnosis leaves, which are excluded from the inquiry set. A diagnosis node is revealed only after the model ranks it from the current state. The full-record oracle therefore withholds the case-specific gold diagnosis leaf.

B.2 Matched baseline families

All baselines in Table 1 use the same fixed fit/development/test splits on MZ-4, DDXPlus and DiReCT as MIRROR. Full-record baselines receive all available findings, excluding differential diagnoses and physician annotation trees.

Black-box baselines. *Random Forest* (Breiman, 2001) trains 100 decision trees on indicator-encoded findings; DiReCT notes are represented by TF–IDF word unigrams and bigrams. *Transformer* encodes the full sequence of findings or note tokens with two attention layers, four heads, and a hidden dimension of 128. *BERT+MLP* (Devlin et al., 2019) mean-pools embeddings from a frozen Chinese BERT model on MZ-4 or BioClinicalBERT on DDXPlus and DiReCT, and trains a 256-unit MLP classifier. *Llama-3-8B-Instruct* predicts from the complete record and candidate diagnoses using greedy decoding; DiReCT diagnosis is predicted after selecting a disease category. *DxFormer* (Chen et al., 2023) uses its original symptom-inquiry framework on MZ-4 and DDXPlus, with a ten-question budget. For DiReCT, it is adapted to full-note classification: the concatenated six-section note is the input and the canonical diagnosis is the target, with no sequential inquiry.

Rule-based and logic-explanation baselines. Recorded findings enter these models as binary indicators; on DiReCT, they share a training-derived vocabulary of at most 1,000 binary word unigrams and bigrams. DDXPlus age is standardized for RRL and binarized for the other three methods. *RRL* (Wang et al., 2021) learns logical features through AND/OR layers and uses them directly for prediction; we use one 64-unit logical layer. *e-LEN* (Barbiero et al., 2022) trains an entropy-regularized classifier and subsequently extracts logical formulas that explain its predictions. *BRCG* (Dash et al., 2018) learns compact Boolean rule sets using the heuristic AIX360 implementation. *DR-Net* (Qiao et al., 2021) learns disjunctive rules with a network of 50 candidate rules. Because BRCG and DR-Net are binary methods, we train one-versus-rest classifiers for multiclass diagnosis. The accuracy of e-LEN's classifier and the fidelity of its extracted formulas are evaluated separately. Our fixed training budgets are 41, 30, and 2,000 epochs for RRL, e-LEN, and DR-Net, respectively; BRCG uses ten search iterations.

B.3 Rule construction

Each bank is compiled once from released sources and then frozen. Training never invents a predicate, body, head, or polarity; it learns only bounded reliabilities on this structure. The clauses below are the Horn forms stored in the banks and instantiate Equation 1.

Table 5: Frozen signed Horn banks. Diagnostic counts include unary and longer bodies.

Dataset	$\lvert\mathcal{K}_{\mathrm{impl}}\rvert$	$\lvert\mathcal{K}_{\mathrm{diag}}\rvert$	$\lvert\mathcal{K}\rvert$
MZ-4	9	155	164
DDXPlus	752	115	867
DiReCT	815	62	877

MZ-4. The bank merges 86 clinical diagnostic templates, nine clinical implications, and 77 pair/triple diagnostic clauses selected by nested fit/dev support and precision gates; exact deduplication leaves 164 rules. Implication rules close finding–finding relations such as high fever $\Rightarrow$ fever, productive cough $\Rightarrow$ cough, and loose stool $\Rightarrow$ diarrhea. Diagnostic templates include unary support (bronchitis $\rightarrow$ pediatric bronchitis), conjunctive support (cough $\wedge$ sputum $\rightarrow$ pediatric bronchitis), and explicit exclusion (wheezing $\rightarrow \neg$ upper respiratory infection). A typical fit-only pair that survives the nested gates is loose stool $\wedge$ diarrhea $\rightarrow \neg$ upper respiratory infection.

DDXPlus. The official condition and evidence tables contribute 115 retained unary diagnostic clauses and 752 typed implication edges. A unary diagnostic clause copies a listed symptom of a target condition, for example fever (`E_91`) $\rightarrow$ HIV (initial infection). A typed implication fires only on a non-default child value: answering that pain is located at the left tonsil (`E_55_@_V_21`) implies the parent question "pain somewhere, related to the consult" (`E_53`); the default value "nowhere" never activates the parent.

DiReCT. The released diagnostic graphs compile into 815 implication clauses and 62 unary diagnostic Horns, one per PDD leaf. Semicolon-separated guideline items become alternative bodies: each of hyperlipidemia, hypertension, and smoking is a separate rule concluding `DX::Suspected ACS`. Explicit numeric AND remains conjunctive; the reversibility criterion compiles as `DX::Suspected Asthma` $\wedge$ (FEV_1 improvement $>12\%$) $\wedge$ (FEV_1 improvement >200 ml) $\rightarrow$ `DX::Asthma`. A parent diagnosis together with a supporting premise yields a chain step such as `DX::NSTE-ACS` $\wedge$ (hs-cTn $>$99th percentile) $\rightarrow$ `DX::NSTEMI`. Guideline negation sets literal polarity: "absence of other systemic or brain diseases to explain the decline" is stored as a negative body literal on the Alzheimer path. Each PDD leaf then receives one diagnostic Horn, e.g. `DX::NSTEMI` $\rightarrow$ NSTEMI.

C SUPPLEMENTARY EXPERIMENTAL RESULTS

C.1 ADDITIONAL LOOP CASE STUDIES

Figure 4 presents the MZ-4 storyboard in the main text; this section supplies the corresponding DiReCT and DDXPlus cases. Grouped bars show the calibrated Top-3 class probabilities. The human diagnosis is always orange when present, so its trajectory remains identifiable even when its rank changes; no orange bar means it is outside the Top-3. The signed-state matrix encodes the model path C_t by cell color and overlays the cumulative reference path $C_t^\star$ with stars; a hollow circle marks the queried coordinate and the tick label gives its revealed sign. The final heatmap shows the rules used at each loop and their activation $a_r^{(t)} = w_r^{(t)}$ times body satisfaction. Long rules are represented by stable implication (I) or diagnostic (D) identifiers. This design keeps the predicate-level state and rule-level computation visible without reducing either to an occupancy count. Predicate and rule names for all three figures are in Appendix E.

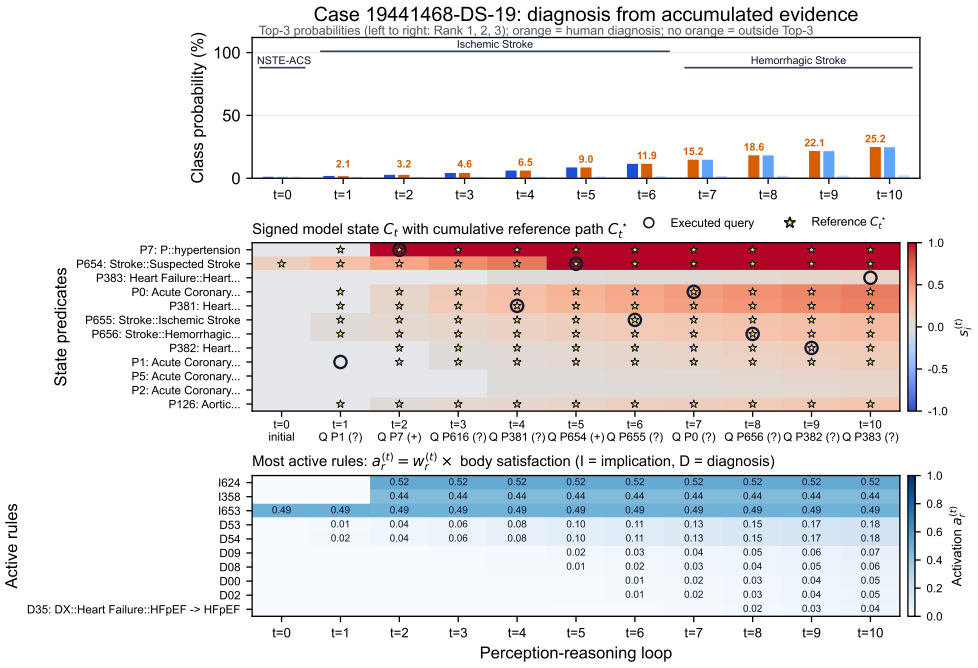


Figure 5: DiReCT note `19441468-DS-19`. The prediction recovers from NSTE-ACS, via ischemic stroke, to hemorrhagic stroke; calibrated correct-class probability rises from 1.6% to 25.2%. The panels use the same Top-3 probability, signed-path, and rule-activation encodings as Figure 4.

C.2 PER-DATASET KNOWLEDGE-REGION DIAGNOSTICS

Figure 7 is the complete $H = 4$ view whose MZ-4 row appears in Figure 3. Figures 8–10 complement that capacity view with the selected $H = 2$ codebooks.

C.3 EXPERIMENTAL SETTINGS

Table 6 records every dataset-specific choice used by the main tables. All choices were fixed on development data. Every dataset uses the same VQ-VAE construction.

All models use $d = 32$ learnable unit anchors, SoftMin AND with $\tau_\wedge = 0.1$, and weighted soft-OR with $\mu_\vee = 10$. Training rolls out to T and applies disease CE only at the final autonomous state. MZ-4 and DiReCT use ordered process CE with $\lambda_{\text{proc}} = 0.3$. All datasets use VQ-VAE with $\lambda_{\text{VQ}} = 1$ and $\beta = 0.25$ as in Van Den Oord et al. (2017). DDXPlus has no clinician trajectory and is trained with $\lambda_{\text{proc}} = 0$. Each step asks one predicate. The confidence threshold $\tau_\star$ is selected on Dev per dataset. No router sees disease labels or unread answers.

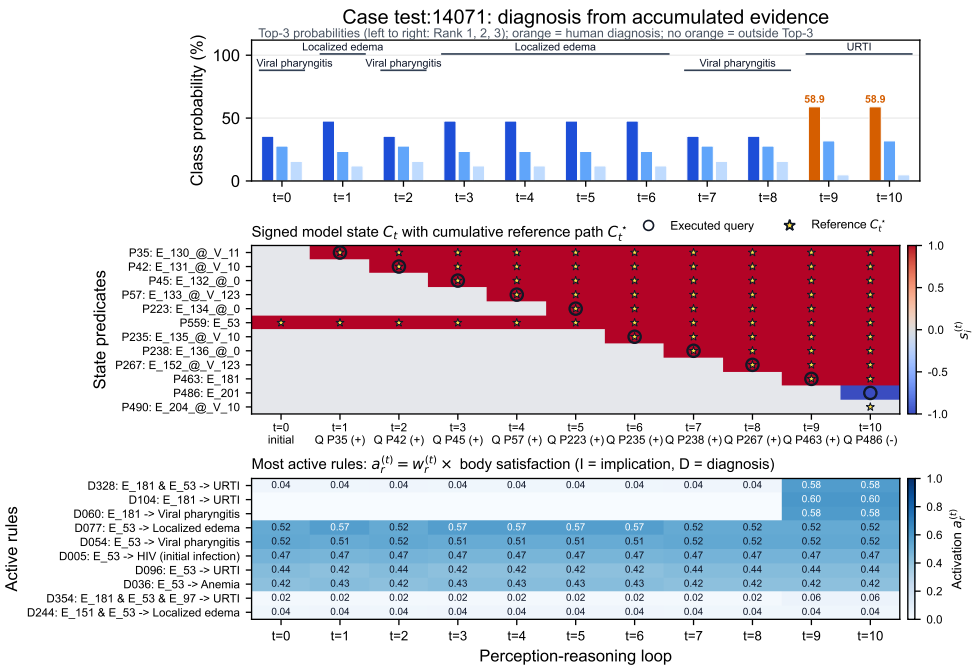


Figure 6: DDXPlus case `test:14071`. The model recovers from viral pharyngitis to URTI; calibrated correct-class probability rises from 11.6% to 58.9%, with the same Top-3 probability, signed-path, and rule-activation encodings as Figure 4.

Table 6: Exact MIRROR settings behind Table 1. L is implication depth and $\tau_\star$ is the confidence stop. Each step asks one predicate.

Dataset	H	L	$\tau_\star$	λ_{proc}	Process target
MZ-4	2	2	0.95	0.3	ordered released evidence
DDXPlus	1	1	0.80	0	none (outcome CE)
DiReCT	2	2	0.80	0.3	physician-tree order

We report diagnosis accuracy and macro-F1. State-path precision, recall, and F1 are micro-aggregated over named nodes in $C = O_{\text{stop}} \cup I_{\text{stop}}$ versus the dataset gold set $C^\star$ (explicit $\cup$ implicit findings), so implication writes count only when they also appear in $C^\star$. Region diagnostics include hard-assignment perplexity, active-code fraction, code separation, and within-trajectory switching. Seed aggregates use mean$\pm$sample standard deviation.

The shared settings omitted from the per-dataset columns are $\tau_{\text{att}} = 1$, $\tau_\wedge = 0.1$, $\mu_\vee = 10$, $\lambda_{\text{VQ}} = 1$, and $\beta = 0.25$. The selected VQ encoder uses the complete $N \times d$ recurrent state and hidden width $d_h = 64$; the training-parameter sensitivity is reported in Appendix C.4. Baseline rows of Table 1 follow the access families in Appendix B.

Knowledge region learning settings. We use this VQ-VAE construction throughout. We initialize the H code vectors by deterministic farthest-point K-means followed by 20 Lloyd iterations. During training, assignment counts n_h and proposal sums $\boldsymbol{m}_h$ are updated with EMA decay 0.99, and $\boldsymbol{w}^{(h)} \leftarrow \boldsymbol{m}_h/(n_h + \varepsilon)$.

The codebook and commitment terms follow Equation 19 in the main text. Under EMA, the codebook term is reported for comparability but the codebook itself follows assigned proposal statistics; the commitment term trains the encoder.

C.4 Hyperparameter sensitivity

For the training factors in Table 7, each listed value is a separate training run. We change one factor at a time around the Figure 2 MZ-4 configuration, with the same split, architecture, rule bank, and

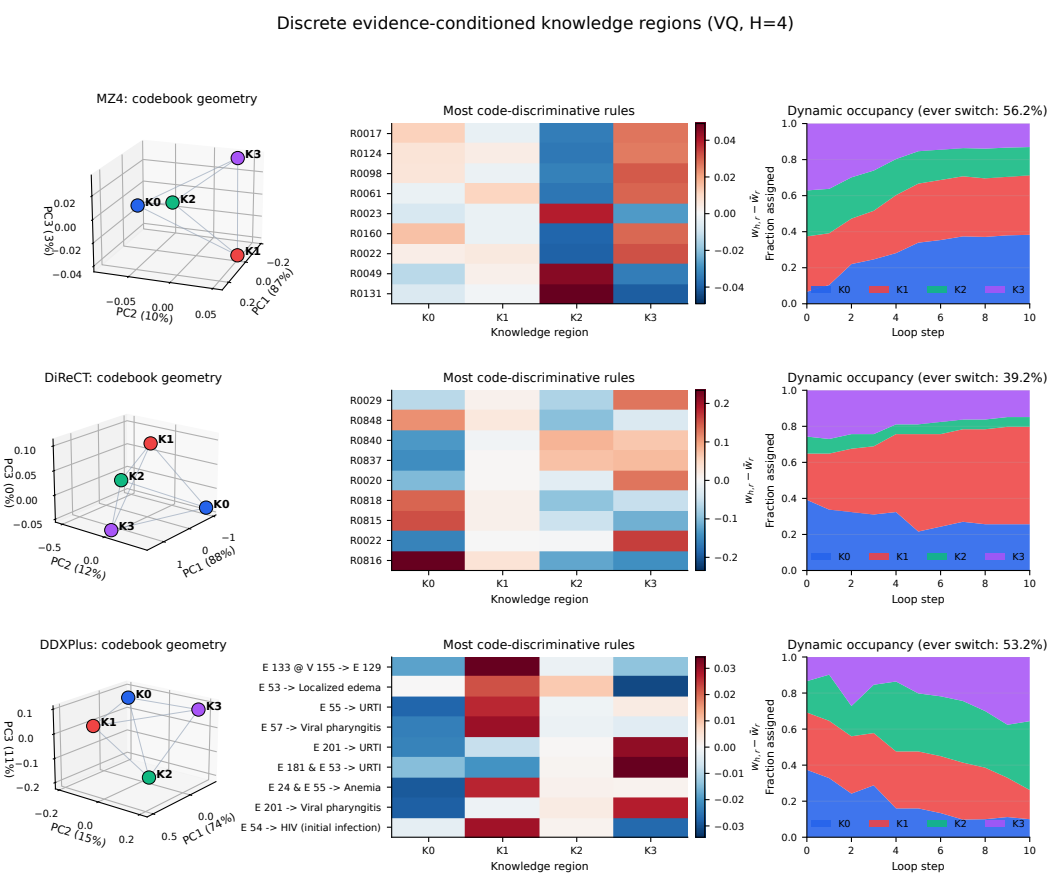


Figure 7: Discrete evidence-conditioned VQ knowledge regions on all three datasets ($H = 4$). Each row uses the same three panels as Figure 3: codebook PCA of $\boldsymbol{W}_K$, the nine most code-discriminative rules, and per-loop occupancy of the frozen test trajectories.

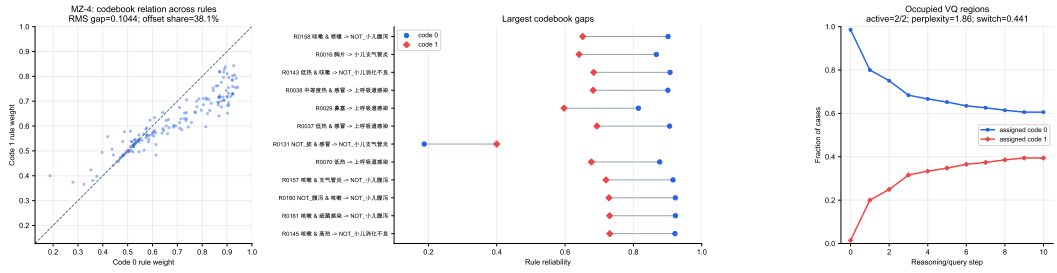


Figure 8: MZ-4 knowledge-region geometry, the largest rule-reliability gaps, and occupancy over loop time.

one inquiry per step, repeated for three seeds. The checkpoint and the confidence threshold $\tau_\star$ are selected on development data. The test set is evaluated once after that choice. Entries there are mean $\pm$ sample SD, and bold values are that default.

H sets the number of knowledge-region codes. λ_{proc} balances process supervision against diagnosis, λ_{VQ} scales the quantization objective, and β is the commitment coefficient in Equation 19. Table 7 reports that three-seed sweep. State F1 is the C versus $C^\star$ overlap in Equation 21.

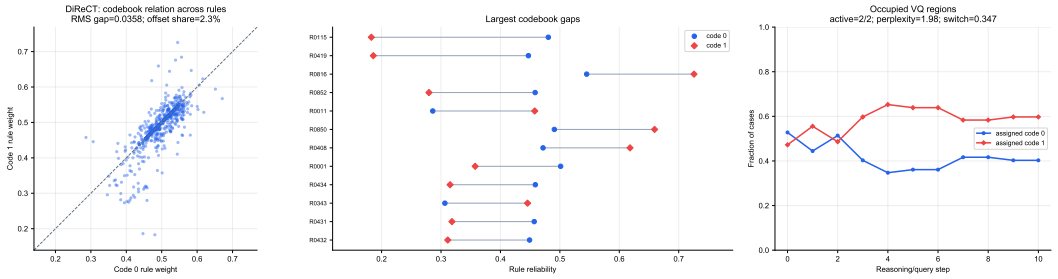


Figure 9: DiReCT knowledge-region diagnostics. DiReCT has the largest individual rule gaps.

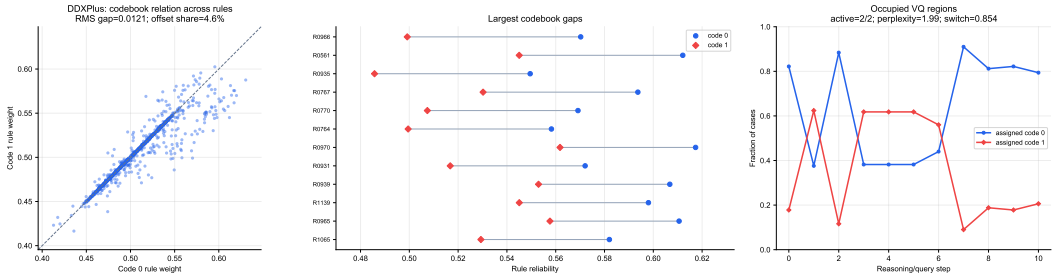


Figure 10: DDXPlus knowledge-region diagnostics. Routing changes substantially over the loop.

Table 7: MZ-4 training sensitivity over three seeds. Accuracy, macro-F1, and state F1 are percentages. Entries are mean ± sample SD. Repeated default rows are the Figure 2 configuration.

Factor	Value	Dev Acc.	Dev F1	Test Acc.	Test F1	State F1
H	1	78.78±0.36	77.21±0.36	71.69±0.44	68.80±0.46	90.73±0.47
	2	**79.02±0.21**	**77.41±0.28**	**72.46±0.29**	**69.45±0.37**	**90.78±0.48**
	4	78.54±0.75	76.51±1.13	71.11±0.93	66.64±2.23	84.26±6.18
λ_{proc}	0	70.02±0.21	66.39±0.48	63.48±0.77	57.14±0.95	43.62±1.23
	0.3	**79.02±0.21**	**77.41±0.28**	**72.46±0.29**	**69.45±0.37**	**90.78±0.48**
	1	77.82±0.55	75.98±1.00	71.01±0.29	67.11±1.93	87.89±4.62
λ_{VQ}	0	78.66±0.42	77.10±0.35	72.46±0.50	69.38±0.55	90.82±0.44
	1	**79.02±0.21**	**77.41±0.28**	**72.46±0.29**	**69.45±0.37**	**90.78±0.48**
β	**0.25**	**79.02±0.21**	**77.41±0.28**	**72.46±0.29**	**69.45±0.37**	**90.78±0.48**
	1	78.54±0.75	76.69±1.26	71.69±0.73	67.65±2.13	87.97±5.60

The $H = 2$ default reaches test accuracy 72.46 ± 0.29 and state F1 90.78 ± 0.48. $H = 1$ stays close, at 71.69 ± 0.44 and 90.73 ± 0.47, while $H = 4$ is 71.11 ± 0.93 with state F1 84.26 ± 6.18. All codes are occupied. Training-assignment perplexity is 1.00 at $H = 1$, 1.97 ± 0.02 at $H = 2$, and 3.59 ± 0.02 at $H = 4$. Raising λ_{proc} from 0.3 to 1 lowers macro-F1 by 2.34 points. The $\lambda_{\text{proc}} = 0$ retraining is the no-process-supervision column of Table 3. Setting $\lambda_{\text{VQ}} = 0$ leaves test accuracy at 72.46 ± 0.50. Raising β to 1 lowers test accuracy by 0.77 points.

The temperatures τ_{att}, $\tau_{\wedge}$, and $\mu_{\vee}$ follow the same protocol: each listed value is a separate training run, repeated for three seeds. The checkpoint and $\tau_{\star}$ are selected on development data, and the test set is evaluated once. Entries in Tables 8–10 are mean ± sample SD. τ_{att} scales named query scores before entmax, $\tau_{\wedge}$ sets the SoftMin in Equation 10, and $\mu_{\vee}$ sets the soft-OR in Equation 11. Bold values are the shared defaults.

Table 8: Test accuracy (%) after retraining τ_{att}. Entries are mean ± sample SD over three seeds. The bold row is the shared default.

τ_{att}	MZ-4	DDXPlus	DiReCT
0.25	72.37±0.73	98.87±0.83	43.06±3.67
0.5	72.37±0.44	98.87±0.83	48.15±2.12
1	**72.27±0.17**	**99.00±0.80**	**50.46±2.89**
2	71.79±0.67	99.00±0.80	50.93±3.21
4	71.79±0.67	99.00±0.80	50.93±3.21

Table 9: Test accuracy (%) after retraining $\tau_{\wedge}$. Entries are mean ± sample SD over three seeds. The bold row is the shared default.

$\tau_{\wedge}$	MZ-4	DDXPlus	DiReCT
0.05	72.27±0.17	99.00±0.80	43.52±6.85
0.1	**72.27±0.17**	**99.00±0.80**	**50.46±2.89**
0.2	72.37±0.17	99.00±0.80	52.78±6.05

Table 10: Test accuracy (%) after retraining $\mu_{\vee}$. Entries are mean ± sample SD over three seeds. The bold row is the shared default.

$\mu_{\vee}$	MZ-4	DDXPlus	DiReCT
5	70.63±0.33	99.00±0.80	49.07±5.26
10	**72.27±0.17**	**99.00±0.80**	**50.46±2.89**
20	71.98±0.17	99.00±0.80	44.91±2.89

The effect of these temperatures is small, and accuracy is stable across the grid. We follow Ren et al. (2026) and set $\tau_{\wedge} = 0.1$ and $\mu_{\vee} = 10$.

D WORKING SETS

The sets below appear throughout.

Symbol	Meaning
$C_0 = O_0$	Chief complaint already in hand: definite facts recorded at initialization, before any attention step. IMCS21: `explicit_info`; MZ-4: `exp_sxs`; DDxPlus: `INITIAL_EVIDENCE`; DiReCT: note-grounded `input1`, intersected with U.
U	*Patient-record facts* available for the case, with recorded signs. IMCS21 and MZ-4: every status-0/1 fact, including later answers and text-rich predicates (status-2 excluded). DDxPlus: closed-world signed sheet (positives and negatives). DiReCT: note-grounded `P::`. U is the fixed set of available replies in the complete case record. A predicate is revealed only after the executed inquiry selects it.
A_t	*Inquiry* at step t: the single named predicate $j_{(1)} = \arg\max_{j \in \mathcal{L}_t} m_j^{(t)}$ (Equation 6). Scores $m_j^{(t)}$ are the column-wise maximum of row-wise $\mathrm{entmax}_{1.5}$ over occupied query rows and unread, not-yet-queried keys. Hence $\lvert A_t \rvert = 1$.
O_t	*Recorded observations* after the executed inquiry of step t, $O_t = O_{t-1} \cup (A_t \cap U)$. On the training process path, a gold-constrained set action is installed only after the current prediction. Only definite replies are recorded. Once in O_t, a fact stays at its recorded sign.
I_t	*Inferred* slots: $\lvert s_i^{(t)} \rvert > \varepsilon$ and $o_i^{(t)} = 0$. They never enter U or O_t.
Q_t	*Query rows* of self-attention: occupied slots $O_{t-1} \cup I_{t-1}$ (observed or already inferred). If this set is empty the round falls back to all N rows.
C_t	*Predicted cognitive path.* $C_0 = O_0$; for $t \geq 1$, $C_t = \{P_i \in \mathcal{P} : \lvert s_i^{(t)} \rvert > \varepsilon\} = O_t \cup I_t$. Asking a slot does not enter the path unless the write or an implication hop changes $\boldsymbol{E}_i$. Path metrics score this occupied set.
$C_t^\star$	*Reference cognitive path* and process target. $C_0^\star = C_0$. For $t \geq 1$, a gold-constrained rollout reveals only members of the dataset-specific reference set, *not* all available patient facts U: ordered doctor requests on IMCS21; the stored implicit-symptom insertion order on MZ-4; remaining positive `EVIDENCES` on DDxPlus (not the negative remainder of U); physician-tree premises and diagnosis nodes on DiReCT. The same implication operator is then applied, so $C_t^\star = O_t^\star \cup I_t^\star$ includes both revealed and inferred slots. For set-valued supervision, any still-unread reference member is a valid next action.

E PREDICATE AND RULE IDENTIFIERS

Case-study figures name predicates by P_i and Horn clauses by Ik (implication) or Dk (diagnostic), using the frozen-bank order. The tables below list every identifier that appears in Figures 4, 5, and 6. DiReCT clinical findings are the knowledge-graph premises; `DX::` nodes are written as category followed by the node name.

MZ-4 (Figure 4, case `test-226`).

ID	Predicate
P12	Moderate fever
P17	Low-grade fever
P26	Laboratory test
P30	Fever
P35	Auscultation
P44	Cough
P55	Wheezing
P79	Dry cough
P93	Common cold
P130	Sputum
P132	Rhonchi
P173	Chest radiograph
P195	Complete blood count

ID	Implication
I01	Low-grade fever → Fever
I05	Dry cough → Cough

ID	Diagnostic rule
D011	Cough AND Sputum → Pediatric bronchitis
D028	Low-grade fever AND Common cold → Upper respiratory infection
D061	Low-grade fever → Upper respiratory infection
D099	NOT Chest radiograph AND Cough → Pediatric bronchitis
D101	NOT Chest radiograph AND Sputum → Pediatric bronchitis
D124	NOT Chest radiograph AND Cough AND Sputum → Pediatric bronchitis
D132	NOT Chest radiograph AND Cough → NOT Pediatric dyspepsia
D134	Low-grade fever AND Cough → NOT Pediatric dyspepsia

DiReCT (Figure 5, note `19441468-DS-19`). P616 is the inquiry at $t = 3$; the state panel retains the twelve highest-importance predicates, so that row is not drawn.

ID	Predicate
P0	Acute Coronary Syndrome: Suspected ACS
P1	Acute Coronary Syndrome: Strongly Suspected ACS
P2	Acute Coronary Syndrome: NSTE-ACS
P5	Acute Coronary Syndrome: STEMI-ACS
P7	Hypertension
P126	Aortic Dissection: Suspected Aortic Dissection
P381	Heart Failure: Suspected Heart Failure
P382	Heart Failure: Strongly Suspected Heart Failure
P383	Heart Failure: Heart Failure
P616	Pulmonary Embolism: Pulmonary Embolism
P654	Stroke: Suspected Stroke
P655	Stroke: Ischemic Stroke
P656	Stroke: Hemorrhagic Stroke

ID	Implication
I358	Hypertension → Heart Failure: Suspected Heart Failure
I624	Hypertension → Stroke: Suspected Stroke
I653	Arm weakness → Stroke: Suspected Stroke

ID	Diagnostic rule
D00	Acute Coronary Syndrome: NSTE-ACS → NSTE-ACS
D02	Acute Coronary Syndrome: STEMI-ACS → STEMI-ACS
D08	Aortic Dissection: Type A Aortic Dissection → Type A Aortic Dissection
D09	Aortic Dissection: Type B Aortic Dissection → Type B Aortic Dissection
D35	Heart Failure: HFpEF → HFpEF
D53	Stroke: Hemorrhagic Stroke → Hemorrhagic Stroke
D54	Stroke: Ischemic Stroke → Ischemic Stroke

DDXPlus (Figure 6, case `test:14071`). Valued answers are separate predicates. No implication identifier is among the active rules in the figure.

ID	Predicate
P35	What color is the rash? = NA
P42	Do your lesions peel off? = no
P45	Is the rash swollen? = 0
P57	Where is the affected region located? = nowhere
P223	How intense is the pain caused by the rash? = 0
P235	Is the lesion (or are the lesions) larger than 1 cm? = no
P238	How severe is the itching? = 0
P267	Where is the swelling located? = nowhere
P463	Do you have nasal congestion or a clear runny nose?
P486	Do you have a cough?
P490	Have you traveled out of the country in the last 4 weeks? = no
P559	Do you have pain somewhere, related to your reason for consulting?

ID	Diagnostic rule
D005	Do you have pain somewhere, related to your reason for consulting? → HIV (initial infection)
D036	Do you have pain somewhere, related to your reason for consulting? → Anemia
D054	Do you have pain somewhere, related to your reason for consulting? → Viral pharyngitis
D060	Do you have nasal congestion or a clear runny nose? → Viral pharyngitis
D077	Do you have pain somewhere, related to your reason for consulting? → Localized edema
D096	Do you have pain somewhere, related to your reason for consulting? → URTI
D104	Do you have nasal congestion or a clear runny nose? → URTI
D244	Do you have swelling in one or more areas of your body? AND Do you have pain somewhere, related to your reason for consulting? → Localized edema
D328	Do you have nasal congestion or a clear runny nose? AND Do you have pain somewhere, related to your reason for consulting? → URTI
D354	Do you have nasal congestion or a clear runny nose? AND Do you have pain somewhere, related to your reason for consulting? AND Do you have a sore throat? → URTI