Supplementary Material

Schema-Guided Tree-Structured Reasoning for Whole-Slide Pathology Workflows

Table S1. Coverage of the decision schema

The schema is derived programmatically from the chain-of-thought annotations of the training corpus (Sec. 2.5); the counts below describe the resulting artefact. A question is counted once per distinct text, and a branch is one (question, answer, follow-up question) triple. Per-case figures are measured over the 14,414 annotated cases.

Schema component	Count
Organ/procedure schemas	27
Organs covered	10
Procedures covered	21
Unique questions	101
Unique (question, answer) pairs	265
Distinct answer categories	158
Branches (question → answer → follow-up)	480
Questions per schema (mean / min / max)	20.5 / 10 / 37
Questions activated per case (mean / median / max)	15.8 / 15 / 33
Dialogue turns per case (mean / max)	12.4 / 25
Cases exceeding the 25-turn inference cap	0

The largest schemas are breast core needle biopsy (37 questions), breast mammotome biopsy and stomach endoscopic biopsy (34 each); the smallest is stomach endoscopic mucosal resection (10).

S2. Representative successful case

Slide PIT_01_03122_01.tiff — Uterine cervix, colposcopic biopsy. Workflow Ranking Score 1.000 (Binary Path Validity 1.00, Edge F1 1.000, MESS 1.000, Final Report Score 1.000). The model reproduces the reference path exactly: every question, every intermediate answer and the final report are identical, so a single listing is given. Note the convergent re-entry into the neoplasm question at steps 3–4, which the schema engine defers until both parent branches are resolved.

Reasoning path (12 steps; reference and model output identical)

0.What is the organ? → **Uterine cervix**

1.What is the procedure? → **Colposcopic biopsy**

2.Is there any abnormality present? → **Yes, there is an abnormality.**

3.Is there any neoplasm present? → **Yes, there is a neoplasm.**

4.Is there any neoplasm present? → **Yes, there is a neoplasm.**

5.What is the behavior of neoplasm? → **Malignant**

6.What is the primary of neoplasm? → **Cervical**

7.What is the histologic type of neoplasm? → **Squamous cell carcinoma**

8.Is there any additional finding present? → **No, there is no additional finding.**

9.What is the number of diagnoses to includes? → **1**

10.What is the #1 diagnosis? → **Invasive squamous cell carcinoma**

11.What is the final pathology report? → **Uterine cervix, colposcopic biopsy; Invasive squamous cell carcinoma**

Final report (reference and model output identical)

Uterine cervix, colposcopic biopsy;
Invasive squamous cell carcinoma

S3. Representative failed case

Slide PIT_01_08593_01.tiff — Stomach, endoscopic biopsy. Workflow Ranking Score 0.183 (Binary Path Validity 0.00, Edge F1 0.308, MESS 0.210, Final Report Score 0.094). The failure is characteristic of the design. The two paths share their first 3 steps and then fork at step3: the model answers *Is there any neoplasm present?* positively where the reference answers negatively. Because the schema engine expands whichever branch the answer selects, every later question is drawn from a subtree the reference never visits, and the model ends up grading a gastric adenocarcinoma while the reference describes chronic active gastritis with ulceration, intestinal metaplasia and CMV-infected cells. The trace remains structurally valid — the engine only ever emits schema-legal questions — but it is diagnostically wrong from step3 onwards, which is why Edge F1 and MESS collapse together rather than independently.

Shared prefix: steps 0–2, identical in both

0.What is the organ? → **Stomach**
1.What is the procedure? → **Endoscopic biopsy**
2.Is there any abnormality present? → **Yes, there is an abnormality.**

Divergence: step 3

3.*Reference:* Is there any neoplasm present? → **No, there is no neoplasm.**
3.*Model:* Is there any neoplasm present? → **Yes, there is a neoplasm.**

Reference path continues: steps 3–19

3.Is there any neoplasm present? → **No, there is no neoplasm.**
4.Is there any polyp present? → **No, there is no polyp.**
5.Is there any active inflammation present? → **Yes, there is an active inflammation.**
6.Is there any active inflammation present? → **Yes, there is an active inflammation.**
7.Is there any active inflammation present? → **Yes, there is an active inflammation.**
8.Is there any active inflammation present? → **Yes, there is an active inflammation.**
9.Is there any lymphoid aggregate present? → **Yes, there is a lymphoid aggregate.**
10.Is there any lymphoid follicle present? → **No, there is no lymphoid follicle.**
11.Is there any ulceration present? → **Yes, there is an ulceration.**
12.What is the depth of ulceration? → **Confined to the mucosa**
13.Is there any intestinal metaplasia present? → **Yes, there is an intestinal metaplasia.**
14.Is there any foveolar epithelial hyperplasia present? → **No, there is no foveolar epithelial hyperplasia.**
15.Is there any additional finding present? → **Yes, there is an additional finding.**
16.What is the additional finding? → **CMV-infected cells**
17.What is the number of diagnoses to includes? → **1**
18.What is the #1 diagnosis? → **Chronic active gastritis**
19.What is the final pathology report? → **Stomach, endoscopic biopsy; Chronic active gastritis 1) erosion 2) intestinal metaplasia 3) lymphoid aggregate 4) CMV-infected cells**

Model path continues: steps 3–18

3.Is there any neoplasm present? → **Yes, there is a neoplasm.**
4.Is there any neoplasm present? → **Yes, there is a neoplasm.**
5.What is the behavior of neoplasm? → **Malignant**
6.What is the primary of neoplasm? → **Gastric**
7.What is the histologic type of neoplasm? → **Adenocarcinoma**
8.What is the histologic subtype of neoplasm? → **Mixed adenocarcinoma**
9.What is the number of distinct histological components? → **2**
10.What is the histological subtype of #1 component? → **Tubular adenocarcinoma**
11.What is the grading system of neoplasm? → **3-tier grading system**

12.What is the grading system of neoplasm? → **3-tier grading system**

13.What is the grade of neoplasm? → **Poorly differentiated**

14.What is the grade of neoplasm? → **Poorly differentiated**

15.What is the histological subtype of #2 component? → **Poorly cohesive carcinoma, not otherwise specified**

16.What is the number of diagnoses to includes? → **1**

17.What is the #1 diagnosis? → **Adenocarcinoma, poorly differentiated with poorly cohesive carcinoma component**

18.What is the final pathology report? → **Stomach, endoscopic biopsy;**
Adenocarcinoma, poorly differentiated with poorly cohesive carcinoma component

Final report: reference

Stomach, endoscopic biopsy; Chronic active gastritis 1) erosion 2) intestinal metaplasia 3) lymphoid aggregate 4) CMV-infected cells

Final report: model output

Stomach, endoscopic biopsy; Adenocarcinoma, poorly differentiated with poorly cohesive carcinoma component