Input

Question:

Why did the skier turn around after his last jump?

A.   He spots a wolf
B.   He accidentally dropped
C.   …
D.   …

Frame_xxxx: subjects, actions, location/scene details
Frame_0001: A skier, clad in a vibrant….
Frame_0002: A skier
.
.
Frame_3598: A bearded male skier, wearing a bright blue jacket, a black helmet with a white Oakley goggle strap, and reflective red-orange mirrored goggles, intently holds black and red ski poles while skiing down a snowy mountain slope with distant peaks under a clear sky.
.
.

Character/Event/Scene Tracking Log:

Main Characters:
- Bearded male skier in blue, with black and red poles…
- Old, skinny man in brown…

Global Summary:

The documentary about Matt – a dedicated and skilled…

Video Logs

Captions Database

1-Hour Video

Extracted Frames

LLM:
`<think> . . . </think>`
… I'm prioritizing the core action ("skier turn around")...

```
{
    "tool":
"CAPTION_SEARCH",
    "prompt": "skier… "
}
```

Captions:
Frame_0120… Score: 0.6971
Frame_0125… Score: 0.6723
.
Frame_2891… Score: 0.5110

LLM Response:
`<think> . . . </think>`
```
{
    "tool": "QUERY_VLM",
    "prompt": Is the skier turned around after his jump? Describe the skier's…
    "frames": 2885 -> 2950
}
```

VLM Response:

Individual frame responses: …

Condensed response:

There is no wolf in any of the frames. The skier is always holding his poles. There …

Critic response:

Confidence: 30,
Possible_errors:
        "VLM reports that in frame 0120… ",
        "The answer states that all…"
Suggestion":
        "Re-evaluate the answer by considering the evidence from frame 0120…"

VLM response:

In frame 0120, a skier wears an orange jacket instead of a blue…

Critic response:

```
{
    "tool":
"QUERY_VLM",
    "prompt": Is there evidence of the skier looking…        "frames": …,
}
```

LLM Response:

Based on the VLM's response, the most plausible answer…
```
{
    "tool": "ANSWER",
    "answer": "D"
    "evidence_frames": …,
    "reasoning": …,
}
```