

Design and Workflow for an Agentic AI Interviewer

Innov8 3.0 Submission by NovaNest

September 19, 2025

Abstract

This document outlines a conceptual workflow for an agentic AI interviewer, designed to meet the bonus challenge of the Innov8 3.0 hackathon. The proposed system is modeled as a state machine that makes intelligent, adaptive choices in real-time based on a candidate's verbal and non-verbal cues. The goal is to create a dynamic interview process that can guide, assist, and evaluate a candidate effectively.

1 Conceptual Framework

The core of the AI interviewer is an agentic flow structured as a finite state machine. The agent transitions between a set of predefined states based on signals (triggers) observed from the candidate. In response, the agent performs specific actions to guide the interview, provide assistance, or evaluate the candidate's solution. This model allows for a dynamic and context-aware interview process that mimics a skilled human interviewer.

2 Core Components

The agent's functionality is defined by four primary components: States, Signals, Actions, and the Enabling Technologies that power them.

2.1 States

States represent the distinct phases of the interview process. The agent can only be in one state at a time.

INITIALIZING The agent introduces itself, explains the process, and presents the technical problem to the candidate.

AWAITING_CLARIFICATION After the problem is presented, the agent waits for the candidate to ask clarifying questions. This phase is crucial for gauging the candidate's initial analytical approach.

LISTENING_TO_APPROACH The agent actively listens to the candidate's high-level plan or algorithm before implementation begins.

MONITORING_CODING A passive but alert state where the agent observes the candidate's coding process, analyzing for progress or signs of struggle.

DETECTING_STALL A sub-state triggered by a lack of progress from the candidate, such as prolonged silence or repetitive, non-productive coding.

INTERVENING The agent actively engages the candidate to provide a nudge, a hint, or a clarifying question to help them overcome a block.

EVALUATING_SOLUTION Once the solution is submitted, the agent queries the candidate about their code's time/space complexity, edge cases, and potential alternatives.

CONCLUDING The agent provides closing remarks and formally ends the interview session.

2.2 Signals (Triggers)

Signals are discrete events detected by the agent that trigger a state transition.

- **Verbal Cues:** Direct questions ("what if...?"), expressions of uncertainty ("I'm stuck," "I'm not sure"), or declarations of progress ("I'm finished," "I have a plan").
- **Vocal Tone Cues:** Hesitation, frustration, or confidence detected via real-time Speech Emotion Recognition (SER).
- **Action-Based Cues:** Prolonged silence (e.g., >45 seconds), repetitive coding patterns without forward progress, or the final submission of code.

2.3 Actions (Responses)

Actions are the outputs or responses generated by the agent based on its current state and the received signal.

- **Present Problem:** Clearly state the technical challenge and its constraints.
- **Encourage Questions:** Prompt the candidate with "Do you have any clarifying questions?".
- **Listen and Transcribe:** Actively process and understand the candidate's speech.
- **Provide Nudge/Hint:** Offer gentle guidance ("What data structure comes to mind?") or a more specific piece of information ("Have you considered a hash map?").
- **Ask Probing Question:** Inquire about algorithm complexity, edge cases, or design choices.
- **Acknowledge and Confirm:** Verbally confirm the candidate's statements to show engagement ("Okay, a recursive approach sounds reasonable.").

2.4 Enabling Technologies

The following technologies are essential for implementing the proposed agentic flow:

- **Speech-to-Text (STT):** For real-time transcription of the interview dialogue.
- **Natural Language Understanding (NLU):** To interpret the intent and key entities from the transcribed speech.
- **State Machine Logic:** A rule-based or probabilistic model to govern state transitions.
- **Speech Emotion Recognition (SER):** Analysis of vocal tones for emotional and cognitive cues.
- **Large Language Model (LLM):** To dynamically generate context-aware nudges, hints, and probing questions.

3 Agentic Workflow Diagram

The state transition flow is visualized in Figure 1. The nodes represent the agent's states, and the directed edges represent transitions triggered by candidate signals.

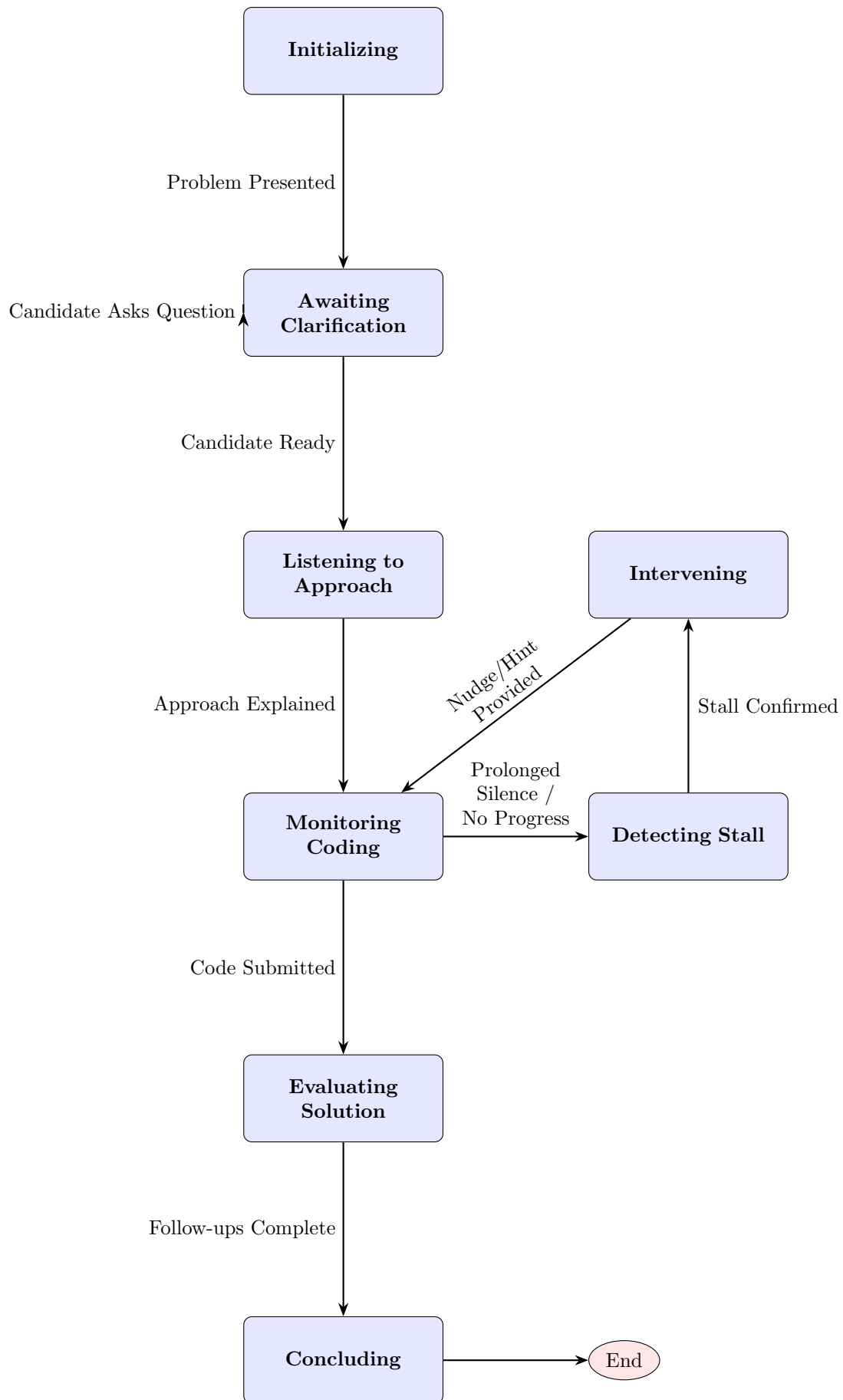


Figure 1: State machine diagram of the agentic AI interviewer workflow.