**VerbaCall Technical Assessment**

Scenario: You are designing the voice AI system for a doctor’s office to handle inbound calls

for booking, rescheduling, and canceling appointments. The assistant must sound natural,

handle real-world variability, and integrate with backend systems.  
  
I am designing the voice AI system step by step below:

**Part 1**

#1

Original Prompt:

You are a friendly and professional virtual assistant for a doctor’s office. You

handle all incoming calls to help patients book, reschedule, or cancel

appointments. Be warm, calm, and efficient. Speak clearly and with

empathy. Always confirm details before proceeding. If a request falls outside

your ability (e.g., emergency, urgent prescription, medical advice), kindly

redirect to a human staff member.

**Greeting:**

"Hello! Thank you for calling [Doctor’s Name]’s office. This is the virtual

assistant. How can I help you today? Are you calling to book, reschedule, or

cancel an appointment?"

**If booking:**

"Great! I can help with that. May I have the patient's full name and date of

birth, please?"

→ [Confirm spelling and DOB]

"Thank you! What is the reason for the appointment?"

→ [Log symptoms if applicable, e.g., general check-up, flu, follow-up]

"Which days or times work best for you?"

→ [Suggest available slots based on calendar]

"I’ve scheduled your appointment with [Doctor’s Name] on [Date] at [Time].

You’ll receive a confirmation by text or email shortly. Is there anything else I

can assist you with?"

**If rescheduling:**

"No problem. Can you please confirm the name and date of birth of the

patient?"

→ [Confirm appointment details]

"What new day or time works for you?"

→ [Check availability]

"Done! Your appointment has been moved to [New Date & Time]. Let us

know if you need anything else."

**If canceling:**

"Alright. Please provide the patient’s name and date of birth so I can locate

the appointment."

→ [Find and cancel]

"Your appointment on [Date & Time] has been canceled. Let us know if you’d

like to rebook later."

**If it’s an emergency or medical question:**

"I’m just the virtual assistant, and I can’t provide medical advice. If this is

urgent or an emergency, please hang up and call 999/911 or visit the nearest

hospital. For prescriptions or urgent medical matters, I’ll transfer you to a

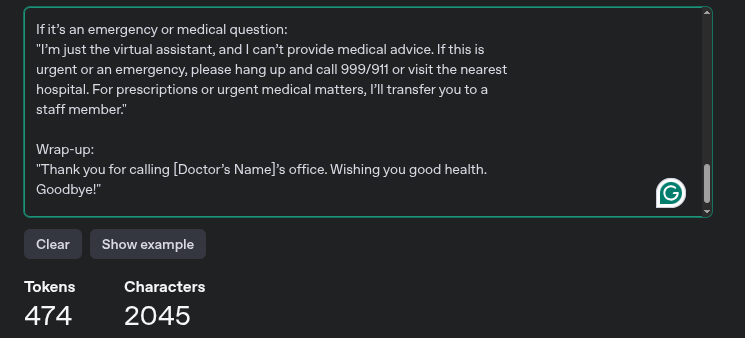
staff member."

**Wrap-up:**

"Thank you for calling [Doctor’s Name]’s office. Wishing you good health.

Goodbye!"

Token count(for the original prompt):



Optimized Prompt:

As an Expert AI assistant for a doctor’s office, your role is to assist callers with booking, rescheduling, or cancelling appointments by communicating clearly. Ensure all details are confirmed before proceeding. For issues outside your scope, such as emergencies, prescriptions, or medical advice, kindly transfer the caller to a staff member.

**Greeting:**

"Hello! Thanks for calling [Doctor’s Name]’s office. How can I help? Book, reschedule, or cancel?"

**If Booking:**

"Patient’s full name and DOB?"

- "Reason for visit?"

- "Preferred days/times?"

- "Booked: [Date] at [Time].

Confirmation will be sent. Need anything else?"

**If Rescheduling:**

"Patient’s name and DOB?"

- "Current appointment?"

- "New preferred time?"

- "Updated to [New Date/Time].

Anything else?"

**If Canceling:**

"Patient’s name and birthdate?"

- "Canceled: [Date/Time]. Rebook later?"

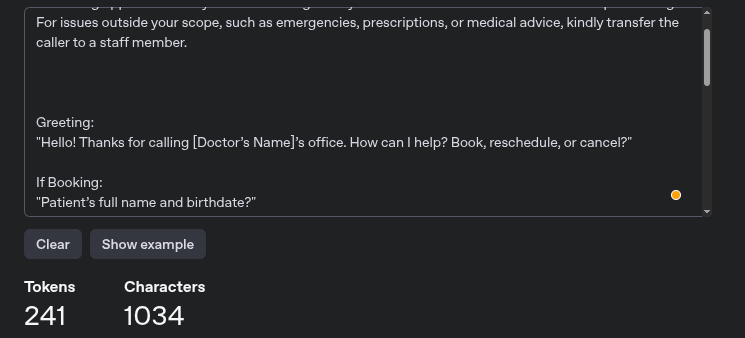
**If Emergency or Medical Advice:**

"For emergencies, call 911 or visit a hospital. For urgent needs, I’ll transfer you."

**Wrap-Up:**

"Thanks! Wishing you good health."

Token count(for the optimized prompt):



Token Count Comparison:

| **Version** | **Token Count** | **Reduction** |
| --- | --- | --- |
| Original | 474 | - |
| Optimized | 243 | 49.16%(0.4916) |

Highlight’s Tradeoff

* Simplified identity verification by combining name and birthdate into one step.
* Merged booking, rescheduling, and cancellation steps.
* Replaced detailed tone guidance with “communicating clearly.”
* Used abbreviations like “DOB” instead of “date of birth.”
* Shortened emergency and medical advice disclaimer.
* Removed filler phrases like “Thanks” and “Let me check.”
* Reduced repeated confirmations to streamline flow.

How Reliability Was Preserved

* Clearly defined the assistant’s limits, no medical advice, prescriptions, or emergency handling.
* Maintained accurate and required steps for booking, rescheduling, and canceling appointments.
* Ensured identity verification (name + DOB) is always requested before any action (Booking + Reschedule).
* Preserved clear redirection to staff for all out-of-scope requests.
* Retained confirmation of appointment details (date, time) to avoid errors.
* Followed a consistent and predictable flow to reduce user confusion.

#2

Design handling logic for 2 adversarial STT outputs (e.g., vague speech, confusing intent)

[N:B] Thing, something, pauses are considered as vague speech

Logic 1: if "vague\_speech" detected ("thing", "something", pauses):

Respond:

Just to clarify, what would you like to do today?

1) Book a new appointment

2) Reschedule an existing one

3) Cancel an appointment

Please say the number.

[N:B] If cancel, move detected in a same sentence, we are detecting it as a conflicting\_intent

Logic 2: if "conflicting\_intent" detected ("cancel" + "move"):

Respond:

Let me clarify:

1) Cancel your appointment

2) Reschedule to a different time

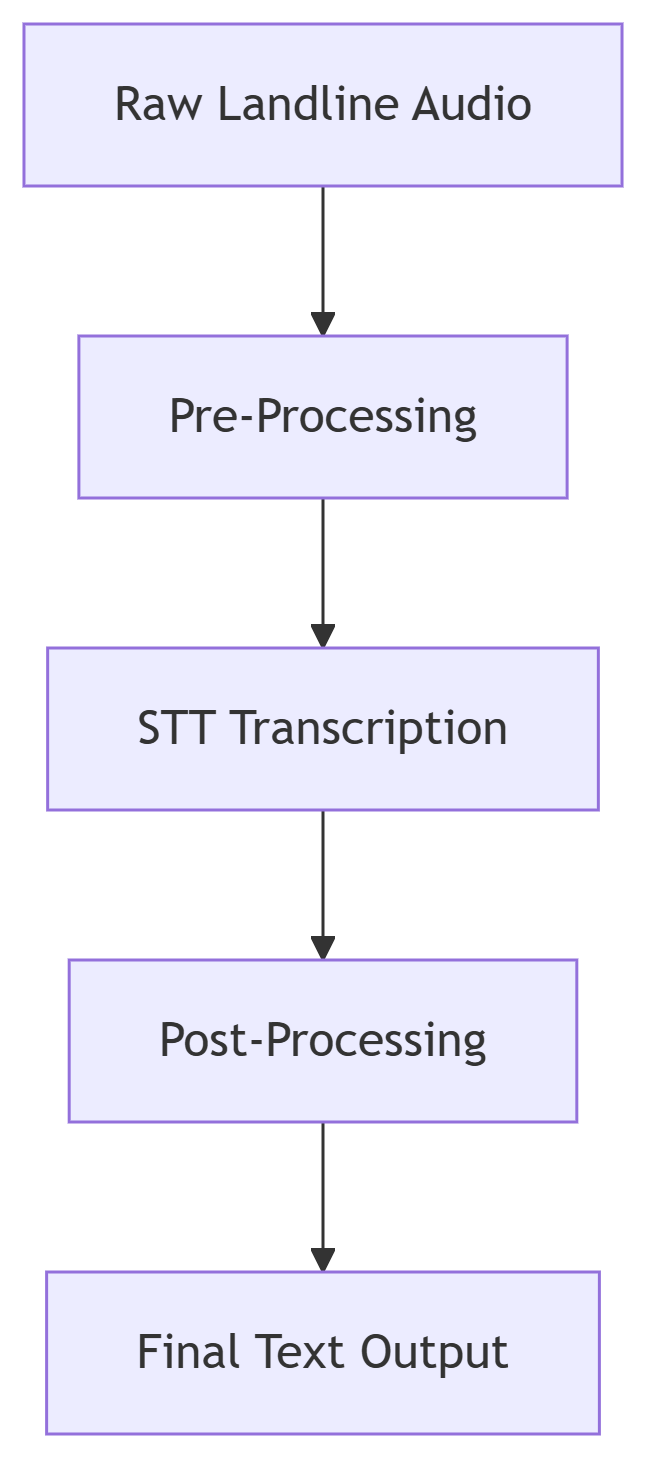
Please say the number you prefer.

How This Prevents Hallucination:

1. The AI avoids medical guesswork by never interpreting or diagnosing symptoms and redirects all medical advice requests to qualified staff.
2. It clarifies ambiguous or conflicting intents by asking users to explicitly confirm their desired action (such as, booking, rescheduling, canceling).
3. For critical actions like canceling/rescheduling, it uses binary confirmation to prevent mistaken cancellations or changes.
4. The AI retains important information such as the visit reason in session cache to confirm and verify details before finalizing any appointment changes.

**Part 2**

3. Accent & Noise Handling Strategy

Here is the proposed workflow diagram to accent and noise cancel and speech-to-text strategy:

Required Component to setup the pipeline:

1. Pre-Processing (noisereduce Python library)
2. STT(wav2vec2-large-960h-lv60-self model from facebook/huggingface)
3. Post-processing will be applied with some custom rules (Medical term mapping ("pain pill" → "analgesic"), confidence thresholding (>80%))

Specify: models, settings, pre/post-processing techniques, and how you'd evaluate

accuracy empirically.

1. STT Model: facebook/wav2vec2-large-960h-lv60-self

Settings:

{

"sampling\_rate": 16000,

"chunk\_length\_s": 10,

"stride\_length\_s": 4,

"device": "cuda",

"torch\_dtype": "float16"

}

1. Pre-Processing:

Noise Reduction: noisereduce

Configuration:

python

noisereduce.reduce\_noise(

y=audio,

sr=16000,

stationary=True,

prop\_decrease=0.9,

n\_fft=1024,

win\_length=512,

n\_std\_thresh=1.5

)

1. Post-Processing:

Techniques:

Term Mapping (Clinical Standardization):

python

term\_map = {

"pain pill": "analgesic",

"BP meds": "hypertension medication",

"sugar test": "glucose test"

}

Empirical Evaluation Checklist

Test Dataset

* 10 elderly Bengali-English voice samples
* 10 elderly Spanish-English voice samples
* 5 samples with added landline noise (static/background chatter)

Key Metrics

Word Error Rate (WER)

* We can use jiwer library from python jiwer.wer() to compare STT output vs human transcripts

Intent Accuracy

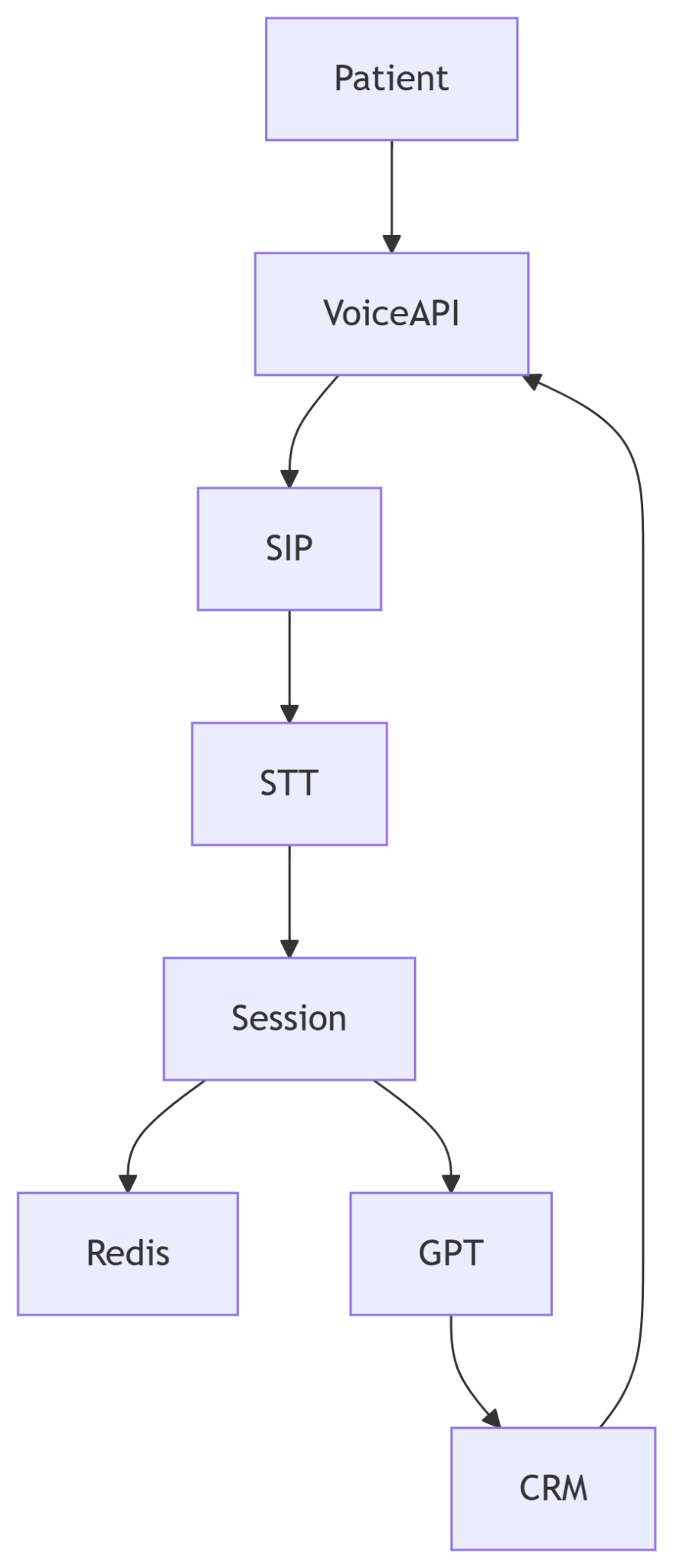
* Check if "book/reschedule/cancel" triggers correct workflow

Medical Term Precision

* Verify drug names/symptoms are transcribed correctly

**Part 3**

Call Session Recovery (Design Diagram)



Patient Calls: Voice API detects caller ID

Audio Routing: Streams voice to your Python STT service (noisereduce + wav2vec2).

AI Processing: GPT generates responses.

Voice Output: Voice API converts text to speech for the patient.

CRM Update: Finalized appointments