CS5002 Assignment #1 - Test Plan

reMIND: powered by Remi

Hannah Laman, Austin Spencer, Spencer Will, Dylan Francis

Part I. Description of Test Plan

Our testing strategy for reMIND is organized into three layers: module/unit testing, subsystem integration testing, and end-to-end scenario testing. We will begin by testing individual modules (speech-to-text, memory retrieval, text-to-speech, caregiver alerting, and authentication) using repeatable simulated inputs and controlled test fixtures. These tests will cover normal, abnormal, and boundary conditions to validate reliability and make sure that the system behaves safely under unexpected conditions (no network, missing memory entries, invalid caregiver tokens). This approach aligns with the course guidance that tests should be repeatable/verifiable and should include abnormal/boundary cases.

Next, we will perform integration tests across interfaces: wearable ↔ backend, backend ↔ caregiver app, and the emergency notification pipeline. Finally, we will validate real-world usage flows with scenario-based blackbox tests (“user is confused and asks where they are,” “device detects a potential emergency and notifies caregiver”), along with performance tests for latency, battery impact, and notification delivery time. We will use a test matrix summarizing coverage across the required classifications as specified in the assignment.

Part II. Test Case Descriptions

RM-01: Wake Word / Activation

- **Purpose:**
  - Ensure the device activates reliably when addressed.
- **Description:**
  - Speak the wake phrase (“Hey Remi”) in a quiet environment and observe device behavior.
- **Inputs:**
  - Audio wake phrase spoken from 0.5–1.5 meters away.
- **Expected Output:**
  - The device enters listening mode within 1 second and provides confirmation (tone or light).
- **Case Type:**
  - Normal
- **Test Method:**
  - Blackbox
- **Test Category:**
  - Functional

- **Test Level:**
    - Integration

RM-02: Speech-to-Text Accuracy (Clear Speech)

- **Purpose:**
    - Validate baseline speech-to-text accuracy.
- **Description:**
    - The user asks standard identity and orientation questions using clear speech.
- **Inputs:**
    - 10 scripted spoken phrases.
- **Expected Output:**
    - >= 90% transcription accuracy and correct intent detection.
- **Case Type:**
    - Normal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Performance
- **Test Level:**
    - Unit

RM-03: Speech-to-Text Robustness (Background Noise)

- **Purpose:**
    - Evaluate STT performance in noisy conditions.
- **Description:**
    - Repeat RM-02 with background TV or crowd noise present.
- **Inputs:**
    - 10 phrases with 60–70 dB background noise.
- **Expected Output:**
    - >= 75% transcription accuracy or clarification prompt when confidence is low.
- **Case Type:**
    - Boundary
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Performance
- **Test Level:**
    - Unit

RM-04: Intent Recognition: Identity Queries

- **Purpose**
  - Ensure correct intent routing for identity-related questions.
- **Description:**
  - Ask variations of “Who am I?” and “What’s my name?”.
- **Inputs:**
  - 6 paraphrased identity questions.
- **Expected Output:**
  - Correct intent classification and appropriate response generation.
- **Case Type:**
  - Normal
- **Test Method:**
  - Blackbox
- **Test Category:**
  - Functional
- **Test Level:**
  - Unit

RM-05: Memory Retrieval: Existing Profile Data

- **Purpose:**
  - Verify correct retrieval of stored user information.
- **Description:**
  - Query stored profile data such as name and daily routine.
- **Inputs:**
  - Preloaded user profile in database.
- **Expected Output:**
  - Accurate responses matching stored data with no hallucinated information.
- **Case Type:**
  - Normal
- **Test Method:**
  - Whitebox
- **Test Category:**
  - Functional
- **Test Level:**
  - Unit

RM-06: Missing Memory Entry Handling

- **Purpose:**
  - Ensure safe behavior when requested data is unavailable.
- **Description:**

    - Ask for personal information that has not been stored in the system.
- **Inputs:**
    - Query for a missing memory key.
- **Expected Output:**
    - System states information is unavailable and optionally prompts caregiver update.
- **Case Type:**
    - Abnormal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Functional
- **Test Level:**
    - Unit

RM-07: Reminder Creation via Caregiver App

- **Purpose:**
    - Validate caregiver-created reminders sync correctly to the wearable.
- **Description:**
    - Caregiver creates a reminder in the app and syncs it to the device.
- **Inputs:**
    - Reminder payload and user ID.
- **Expected Output:**
    - A reminder appears on the device and the backend confirms synchronization.
- **Case Type:**
    - Normal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Functional
- **Test Level:**
    - Integration

RM-08: Reminder Delivery Timing

- **Purpose:**
    - Ensure reminders trigger at the scheduled time.
- **Description:**
    - Set a reminder for two minutes in the future and observe device behavior.
- **Inputs:**

    - Scheduled reminder event.
- **Expected Output:**
    - Device announces reminder within ±30 seconds of target time.
- **Case Type:**
    - Boundary
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Performance
- **Test Level:**
    - Integration

RM-09: Orientation Query: “Where Am I?”

- **Purpose:**
    - Validate orientation guidance responses.
- **Description:**
    - User asks “Where am I?” while location context is available.
- **Inputs:**
    - Orientation query with location context.
- **Expected Output:**
    - Calm response stating location and current activity context.
- **Case Type:**
    - Normal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Functional
- **Test Level:**
    - Integration

RM-10: Emergency Trigger: Manual SOS

- **Purpose:**
    - Ensure emergency alert system functions correctly.
- **Description:**
    - User activates SOS via button or voice command.
- **Inputs:**
    - SOS activation event.
- **Expected Output:**
    - Caregiver receives alert within 30 seconds and event is logged.
- **Case Type:**

    - Normal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Performance
- **Test Level:**
    - Integration

RM-11: Emergency False Positive Prevention

- **Purpose:**
    - Prevent unintended emergency alerts.
- **Description:**
    - Speak non-emergency phrases containing the word “help.”
- **Inputs:**
    - 10 non-emergency speech samples.
- **Expected Output:**
    - No emergency alert is sent; clarification requested if needed.
- **Case Type:**
    - Abnormal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Functional
- **Test Level:**
    - Unit

RM-12: Caregiver App Authentication

- **Purpose:**
    - Verify secure caregiver login.
- **Description:**
    - Log in using valid caregiver credentials.
- **Inputs:**
    - Valid username and password.
- **Expected Output:**
    - Successful login and access to authorized user data.
- **Case Type:**
    - Normal
- **Test Method:**
    - Blackbox
- **Test Category:**

    - Functional
- **Test Level:**
    - Integration

RM-13: Unauthorized Access Attempt

- **Purpose:**
    - Ensure the system blocks invalid access attempts.
- **Description:**
    - Attempt login with incorrect credentials.
- **Inputs:**
    - Invalid password or expired token.
- **Expected Output:**
    - Access denied and security event logged.
- **Case Type:**
    - Abnormal
- **Test Method:**
    - Blackbox
- **Test Category:**
    - Functional
- **Test Level:**
    - Integration

RM-14: Data Encryption in Transit

- **Purpose:**
    - Validate secure data transmission between system components.
- **Description:**
    - Inspect network traffic during device–backend communication.
- **Inputs:**
    - API requests.
- **Expected Output:**
    - Only encrypted TLS connections allowed; plaintext rejected.
- **Case Type:**
    - Normal
- **Test Method:**
    - Whitebox
- **Test Category:**
    - Functional
- **Test Level:**
    - Integration

RM-15: Offline Mode Operation

- **Purpose:**
  - Ensure limited functionality when internet connectivity is unavailable.
- **Description:**
  - Disable network and issue identity and reminder queries.
- **Inputs:**
  - Offline device state.
- **Expected Output:**
  - Cached responses are provided and sync is queued for later.
- **Case Type:**
  - Abnormal
- **Test Method:**
  - Blackbox
- **Test Category:**
  - Functional
- **Test Level:**
  - Integration

RM-16: Sync Conflict Resolution

- **Purpose:**
  - Validate conflict handling between device and caregiver edits.
- **Description:**
  - Modify the same memory entry from device and caregiver app simultaneously.
- **Inputs:**
  - Conflicting updates with timestamps.
- **Expected Output:**
  - Deterministic conflict resolution with no data loss.
- **Case Type:**
  - Boundary
- **Test Method:**
  - Whitebox
- **Test Category:**
  - Functional
- **Test Level:**
  - Integration

RM-17: Voice Response Latency

- **Purpose:**
  - Ensure responsive voice interaction experience.

- **Description:**
  - Measure time from end of user speech to Remi's response.
- **Inputs:**
  - 20 representative user queries.
- **Expected Output:**
  - Median latency ≤2 seconds; 95th percentile ≤4 seconds.
- **Case Type:**
  - Normal
- **Test Method:**
  - Blackbox
- **Test Category:**
  - Performance
- **Test Level:**
  - Integration

RM-18: Wearable Battery Life

- **Purpose:**
  - Validate acceptable wearable battery performance.
- **Description:**
  - Simulate a typical day of use with reminders and interactions.
- **Inputs:**
  - 30 interactions, reminders, and periodic syncs.
- **Expected Output:**
  - Device meets target battery life with no overheating.
- **Case Type:**
  - Normal
- **Test Method:**
  - Blackbox
- **Test Category:**
  - Performance
- **Test Level:**
  - Integration

Part III. Test Case Matrix

ID	Purpose	Description	Inputs	Expected Output / Results	Case Type	Test Method	Test Category	Test Level
RM-01	Ensure device activates reliably	Speak wake phrase to activate Remi	Wake phrase audio	Device enters listening mode within 1s	Normal	Blackbox	Functional	Integration

ID	Purpose	Description	Inputs	Expected Output / Results	Case Type	Test Method	Test Category	Test Level
RM-02	Validate STT accuracy	Ask standard identity/orientation questions	10 scripted phrases	≥90% transcription accuracy	Normal	Blackbox	Performance	Unit
RM-03	Test STT under noise	Repeat STT test with background noise	Phrases + noise	≥75% accuracy or clarification	Boundary	Blackbox	Performance	Unit
RM-04	Validate intent recognition	Ask identity-related questions	Paraphrased identity queries	Correct intent classification	Normal	Blackbox	Functional	Unit
RM-05	Verify memory retrieval	Query stored user profile data	Preloaded profile	Accurate responses, no hallucination	Normal	Whitebox	Functional	Unit
RM-06	Handle missing memory safely	Ask for unstored personal info	Missing memory query	“Information unavailable” response	Abnormal	Blackbox	Functional	Unit
RM-07	Validate reminder creation	Create reminder via caregiver app	Reminder payload	Reminder synced to device	Normal	Blackbox	Functional	Integration
RM-08	Validate reminder timing	Deliver scheduled reminder	Timed reminder	Reminder delivered ±30s	Boundary	Blackbox	Performance	Integration
RM-09	Validate orientation guidance	Ask “Where am I?”	Orientation query	Calm, contextual response	Normal	Blackbox	Functional	Integration
RM-10	Validate emergency alerts	Trigger SOS manually	SOS event	Caregiver notified within 30s	Normal	Blackbox	Performance	Integration
RM-11	Prevent false emergencies	Say non-emergency phrases	Speech samples	No emergency triggered	Abnormal	Blackbox	Functional	Unit
RM-12	Validate caregiver login	Log in with valid credentials	Username/password	Successful authentication	Normal	Blackbox	Functional	Integration
RM-13	Block unauthorized access	Attempt invalid login	Invalid credentials	Access denied and logged	Abnormal	Blackbox	Functional	Integration
RM-14	Ensure encrypted transport	Inspect API traffic	Network requests	TLS-only communication	Normal	Whitebox	Functional	Integration
RM-15	Validate offline behavior	Disable network and query	Offline device	Cached response provided	Abnormal	Blackbox	Functional	Integration
RM-16	Resolve sync conflicts	Concurrent edits from device/app	Conflicting updates	Deterministic resolution	Boundary	Whitebox	Functional	Integration
RM-17	Validate response	Measure voice response time	User queries	≤2s median latency	Normal	Blackbox	Performance	Integration