

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG

NGUYỄN THỊ THÙY DUNG

**NGHIÊN CỨU ỨNG DỤNG MÃ NGUỒN MỞ
MICROSOFT SDK SPEECH 5.1
ĐỂ XÂY DỰNG
PHẦN MỀM LUYỆN PHÁT ÂM TIẾNG ANH**

**Chuyên ngành: Khoa học máy tính
Mã số: 60.48.01**

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

ĐÀ NẴNG – Năm 2011

**Công trình được hoàn thành tại
ĐẠI HỌC ĐÀ NẴNG**

Người hướng dẫn khoa học: **PGS.TS. Võ Trung Hùng**

Phản biện 1: TS. Nguyễn Thanh Bình

Phản biện 2: PGS.TS. Đoàn Văn Ban

Luận văn được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp thạc sĩ Khoa học máy tính họp tại Đại học Đà Nẵng vào ngày 18 tháng 06 năm 2011.

Có thể tìm hiểu luận văn tại:

- Trung tâm Thông tin – Học liệu, Đại học Đà Nẵng
- Trung tâm Học liệu, Đại học Đà Nẵng

MỞ ĐẦU

1. Lý do chọn đề tài

Theo thống kê của vụ giáo dục đại học sau khi đánh giá 59 trường đại học lớn tại Việt Nam không chuyên ngữ, thì có 51.7% sinh viên tốt nghiệp không đáp ứng được yêu cầu về kỹ năng tiếng Anh. Trước tình hình này cho thấy, vấn đề tiếng Anh đang là mối quan tâm hàng đầu của các doanh nghiệp khi thực hiện phỏng vấn tuyển dụng. Vậy có thể nói rằng, Anh ngữ là tiêu chuẩn đánh giá đầu tiên và quan trọng nhất cho một nhân viên muốn vào làm tại doanh nghiệp, không những là các công ty nước ngoài, mà nhiều doanh nghiệp trong nước hiện nay cũng đánh giá cao yêu cầu này của nhân viên.

Thực tế, chương trình học tiếng Anh trên lớp có rất ít thời gian để luyện đọc. Việc luyện đọc tại nhà lại không ai chỉ bảo cho người học rằng họ đã phát âm đúng hay chưa.

Yêu cầu đặt ra với mỗi người học môn Anh Văn là phải tự học, bên cạnh ngữ pháp cần tự luyện phát âm chuẩn, ngoài việc viết đúng chính tả một từ vựng mới chúng ta còn có thể đọc đúng được nó. Nắm được vai trò quan trọng của việc rèn luyện phát âm Tiếng Anh, cũng như nhu cầu học tập nâng cao khả năng Tiếng Anh của các bạn học sinh sinh viên, tôi đã ứng dụng và triển khai đề tài: “Nghiên cứu xây dựng phần mềm mã nguồn mở Microsoft SDK Speech 5.1 để xây dựng phần mềm luyện phát âm Tiếng Anh”.

2. Mục tiêu và nhiệm vụ nghiên cứu

- Xây dựng một phần mềm thân thiện với người học Anh Văn, đóng vai trò như một người bạn hỗ trợ việc luyện phát âm
- Tạo tài liệu đầy đủ và chính xác về mã nguồn mở Microsoft Speech SDK 5.1 làm cơ sở ứng dụng và phát triển sản phẩm phần mềm
- Rèn luyện thêm kỹ năng đọc hiểu tài liệu Tiếng Anh chuyên ngành và kỹ năng tổng hợp các kiến thức để độc lập tự chủ tạo ra một bản thiết kế, một phần mềm hoàn chỉnh

3. Đối tượng và phạm vi nghiên cứu

- Đối tượng
 - Công nghệ xử lý tiếng nói
 - Mã nguồn mở xử lý tiếng nói Microsoft Speech SDK 5.1
- Phạm vi
 - Chỉ xử lý và nhận dạng giọng nói Tiếng Anh

4. Giả thiết nghiên cứu

- Dựa vào thư viện giọng chuẩn của mã nguồn mở Microsoft Speech SDK 5.1 để phát âm cũng như làm cơ sở để nhận dạng từ Tiếng Anh chuẩn
- Dựa vào các phương thức nhận dạng, các hàm nhận dạng do bộ sản phẩm mã nguồn mở này cung cấp.

5. Phương pháp nghiên cứu

- Phương pháp tài liệu: nghiên cứu các tài liệu liên quan đến công nghệ nhận dạng giọng nói
- Phương pháp thực nghiệm: nghiên cứu ứng dụng mã nguồn mở Microsoft Speech SDK 5.1

6. Ý nghĩa khoa học và thực tiễn của đề tài

- Ý nghĩa khoa học:
 - Nâng cao hiệu quả đánh giá phát âm
 - Tạo tài liệu Tiếng Việt hướng dẫn phát triển ứng dụng Microsoft Speech SDK 5.1
- Ý nghĩa thực tiễn:
 - Người học có thể tự luyện phát âm, chủ động được thời gian cũng như không gian học tập
 - Nâng cao hứng thú cho người học nhờ vào khả năng nhận dạng giọng nói một cách tự động mới mẻ này

CHƯƠNG 1: NGHIÊN CỨU TỔNG QUAN

Trong chương này, chúng tôi trình bày cơ sở lý thuyết liên quan đến đề tài. Nội dung của chương này giới thiệu về: Công cụ xử lý tiếng nói nhằm nghiên cứu sơ lược lý thuyết nhận dạng; một số phần mềm nhận dạng tiếng nói hiện nay nhằm tìm hiểu thực tế các chức năng và yêu cầu cần có cho một sản phẩm phần mềm của mình; giới thiệu về Microsoft Speech SDK 5.1 nhằm tạo một tài liệu lý thuyết để phát triển mã nguồn mở nhận dạng giọng nói Microsoft Speech SDK 5.1.

1.1. CÔNG CỤ XỬ LÝ TIẾNG NÓI

1.1.1. Nhận dạng tiếng nói

1.1.2. Tổng hợp tiếng nói

1.2. MỘT SỐ PHẦN MỀM NHẬN DẠNG TIẾNG NÓI HIỆN NAY

+ Vspeech: phần mềm nhận dạng giọng nói tiếng Việt trên máy tính: do nhóm BK02 của Đại học Bách khoa TP.HCM chế tạo thành công vào cuối năm 2004.

+ Phần mềm luyện phát âm Tiếng Anh Power Pronunciation: Chương trình được thiết kế để luyện phát âm các chữ cái, từng từ một, từng câu một. Với giao diện đồ họa sinh động thể hiện vòm miệng khi phát âm các từ, câu người học theo đó quan sát mà học theo sao cho đúng.

1.3. GIỚI THIỆU VỀ MICROSOFT SDK 5.1

1.3.1. Qui định khi dùng mã nguồn mở Microsoft Speech SDK

+ Yêu cầu về hệ thống:

- Windows XP Professional or Home editions;
- Windows.NET Server editions;
- Microsoft Windows 2000 Professional Workstation;
- Microsoft Windows Millennium edition.
- Microsoft Windows 98.
- Microsoft Windows ® NT Workstation hoặc phiên bản 4.0,
- Windows 95 hay Win trước đó thì không được hỗ trợ

+ Yêu cầu về phần cứng:

- A PentiumII\PentiumII-equivalent trở lên với 233 MHz; 128 MB RAM.
- Một micro có chất lượng cao với bộ lọc tiếng ồn phục vụ cho việc nhận dạng.

Các sản phẩm phần mềm do người dùng xây dựng được bảo vệ bởi luật bản quyền và các hiệp ước bản quyền quốc tế, cũng như luật sở hữu trí tuệ và các điều ước, đồng thời được cấp phép, nhưng không bán được.

1.3.2. Phương thức xử lý trong Microsoft Speech SDK 5.1

Các giao diện chương trình ứng dụng (Application Programming Interface _API) SAPI làm giảm đáng kể chi phí để viết mã cho một ứng dụng bằng việc sử dụng “nhận dạng giọng nói”

(Speech recognition_SR) và từ văn bản đến giọng nói “text-to-speech” (TTS), giúp cho công nghệ giọng nói trở nên dễ tiếp cận hơn và mạnh mẽ cho một loạt các ứng dụng.

SAPI API cung cấp một giao diện cấp cao giữa ứng dụng và các công cụ giọng nói. SAPI thực hiện tất cả các chi tiết cấp thấp cần thiết để kiểm soát và quản lý hệ thống thời gian thực của các công cụ khác nhau.

Hai loại cơ bản của công cụ SAPI là TTS và SR. Hệ thống TTS tổng hợp các chuỗi văn bản và các tập tin âm thanh bằng cách sử dụng giọng nói được tổng hợp. SR chuyển đổi giọng nói của con người vào chuỗi văn bản và những tập tin có thể đọc được.

+ TTS của API:

Để tạo các ứng dụng ta điều khiển TTS bằng việc sử dụng giao diện đối tượng thành phần (Component Object Model) ISpVoice. Khi tạo ra một đối tượng ISpVoice, chỉ cần gọi ISpVoice::Speak để đọc một dữ liệu văn bản. Ngoài ra, ISpVoice cũng cung cấp một số phương pháp để thay đổi giọng nói và các thuộc tính khác như tốc độ ISpVoice::SetRate, Âm thanh ISpVoice::SetVolume và giọng nói hiện tại ISpVoice::Setvoice.

+ SR của API

Cũng như ISpVoice giao diện chính để tổng hợp giọng nói, ISpRecoContext là giao diện chính cho nhận dạng giọng nói.

Một ứng dụng có thể chọn một trong hai kiểu nhận dạng giọng nói. (Speech recognition engines_ISpRecognizer). Nhận dạng chia sẻ với những ứng dụng nhận dạng giọng nói khác. Để tạo một ISpRecoContext cho một ISpRecognizer chia sẻ, chỉ cần gọi CoCreateInstance của COM, trên thành phần CLSID_SpSharedRecoContext. Để tạo một ISpRecoContext cho một

ISpRecognizer InProc cho một InProc ISpRecognizer, đầu tiên phải gọi CoCreateInstance trên CLSID_SplnprocRecoInstance để tạo ra InProc ISpRecognizer riêng của nó. Sau đó phải thực hiện một cuộc gọi đến IspRecohnizer::Setinput để thiết lập các đầu vào âm thanh. Cuối cùng, có thể gọi ISpRecognizer::CreateRecoContext để có được một IspRecoContext.

Bước tiếp theo thiết lập các thông báo cho các sự kiện được quan tâm như: ISpRecognizer cũng như IspEventSource lần lượt là IspNotifySource, có thể gọi một trong những phương thức của IspNotifySource từ ISpRecoContext để chỉ ra nơi các sự kiện IspRecoContext nên được báo cáo.

Cuối cùng, một ứng dụng giọng nói phải tạo, tải và kích hoạt IspRecoGrammar, để chỉ ra loại phát âm sẽ nhận dạng, ví dụ như: chính tả hay một câu lệnh và cú pháp điều khiển. Đầu tiên ứng dụng sẽ tải ngữ pháp thích hợp bằng cách gọi IspRecoGrammar::LoadDictation cho chính tả hoặc một trong các phương thức IspRecoGrammar::LoadCmdxxx cho lệnh và điều khiển. Cuối cùng, để kích hoạt các văn phạm cho việc nhận dạng có thể bắt đầu, thì sẽ gọi IspRecoGrammar::SetDictationStat cho chính tả hay IspRecoGrammar::SetRuleState hay IspRecoGrammar::SetRuledState cho lệnh và điều khiển.

Khi sự nhận dạng quay trở lại ứng dụng bằng cơ chế thông báo yêu cầu thì sẽ là một IspRecoResult qua đó xác định được những gì nhận dạng được. Mỗi một ISpRecognizer có thể có nhiều IspRecoContexts liên kết với nó, và một số có thể được thông báo theo cách riêng của mình về các sự kiện liên quan đến nó.

1.3.3. Thư viện trong Microsoft Speech SDK 5.1

+ Các giao diện được sử dụng trong luận văn

GIAO DIỆN	CHỨC NĂNG
1. ISpeechGrammarRule: - Thuộc tính: + Attributes + Id + IntialState + Name - Phương thức: + AddResource + AddState + Clear	Định nghĩa những thuộc tính và phương thức của một qui tắc ngữ pháp tiếng nói - Trả về thông tin những thuộc tính của mỗi qui tắc ngữ pháp - Chỉ rõ ID của SpeechGrammarRule - Chỉ rõ trạng thái ban đầu - Chỉ rõ tên - Thêm chuỗi vào qui tắc nhận dạng - Thêm trạng thái vào qui tắc nhận dạng - Xóa bỏ những qui tắc và rời khỏi trạng thái khởi tạo
2. ISpeechRecoGrammar - Thuộc tính: + Id + RecoContext + Rules + State - Phương thức: + Reset + CmdStRuleState	Cho phép ứng dụng quản lý từ và mệnh đề - Trả về ID của ngữ pháp khi ngữ pháp được tạo ra - Trả về đối tượng tạo ra ngữ pháp này - Trả về tập hợp những qui tắc ngữ pháp - Thiết lập và trả về trạng thái hệ thống của ngữ pháp - Xóa bỏ mọi qui tắc ngữ pháp và cài đặt lại - Kích hoạt hoặc khử kích hoạt của qui tắc bởi tên qui tắc đó.

3. ISpeechRecoResult	Trả về kết quả nhận dạng. Một kết quả nhận dạng được trả về bởi một văn cảnh nhận dạng trong 3 trường hợp: nhận dạng thành công, nhận dạng trung gian và nhận dạng không thành công.
- Nhận dạng thành công	- Là một từ hay một cụm từ được xem là đủ chính xác để quay trở lại người sử dụng như văn bản đã được nói trên thực tế.
- Nhận dạng trung gian (giả thuyết nhận dạng)	- Là bước trung gian để đi đến 1 nhận dạng cuối cùng, giả thuyết nhận dạng được sản xuất suốt trong quá trình nhận dạng. Một giả thuyết nhận dạng có thể không phản chiếu nhận dạng cuối cùng.
- Nhận dạng sai	- Do một trong các lý do sau: tạp nhiễu nền, phát âm không chính xác, từ khác thường và chuỗi khác thường của từ.
- Thuộc tính: + PhraseInfo	- Trả về một cấu trúc kết quả nhận dạng chứa đựng thông tin chi tiết về mệnh đề nhận dạng cuối cùng
- Phương thức: + Audio	- Tạo ra luồng âm thanh từ dữ liệu kết quả

+ Các lớp đối tượng được sử dụng trong luận văn

ĐỐI TƯỢNG	CHỨC NĂNG
1. SpSharedRecoContext	<p>Định nghĩa một văn cảnh nhận dạng. Văn cảnh nhận dạng là một đối tượng cho phép một ứng dụng bắt đầu hay dừng nhận dạng, nhận được những kết quả nhận dạng và những sự kiện khác. Nó cũng điều khiển và kiểm soát những từ và cụm từ sẵn có mà người sẽ nói. Những từ không chứa trong tập hợp hay văn cảnh, sẽ không được sử dụng cho việc nhận dạng. Bằng việc thiết đặt những văn cảnh nhận dạng, ứng dụng sẽ giới hạn hay mở rộng phạm vi của những từ cần cho nhận dạng. Vì vậy, để cải thiện chất lượng nhận dạng ta nên loại bỏ những từ không cần và thêm vào những từ cần trong một khoảng khắc nhất định cho việc nhận dạng.</p> <p>Tài nguyên duy nhất là Grammar, để tạo Grammar ta sử dụng phương thức <code>CreatGrammar()</code>. Grammar định nghĩa tập hợp những từ cho văn cảnh nhận dạng, những từ đó có trong danh sách đã thiết kế.</p>
2. SpVoice	<p>Mang công cụ Text to Speech _ đọc văn bản. Đặc trưng cơ bản của giọng nói là thuộc tính Voice (giọng đọc nam hay nữ), thuộc tính Rate (tốc độ) và thuộc tính Volume (âm lượng). Sử dụng phương thức <code>GetVoice()</code> để xác định những giọng nói khác nhau sẵn có, những thuộc tính này có thể được thay đổi với một giao diện người dùng.</p>

CHƯƠNG 2

PHÂN TÍCH THIẾT KẾ HỆ THỐNG

Trong chương này, chúng tôi bắt đầu đưa ra các chức năng cần có cho sản phẩm phần mềm của mình và tiến hành phân tích thiết kế hệ thống theo hướng đối tượng. Nội dung của chương này gồm: phân tích yêu cầu hệ thống, đặc tả chức năng chính của hệ thống và phân tích hệ thống theo hướng đối tượng.

2.1. PHÂN TÍCH YÊU CẦU HỆ THỐNG

Người sẽ sử dụng hệ thống mà chương trình phát triển là người luyện đọc, luyện phát âm chuẩn Tiếng Anh.

2.1.1. Các yêu cầu chức năng

- Phân tích tổng hợp giọng nói nhanh ổn định;
- Nhận dạng, phân biệt được các từ có phát âm gần giống nhau; cho kết quả nhận dạng tương đối chính xác;
- Nhận dạng giọng nói một cách tự động;
- Cho giọng đọc có điểm nhấn, âm gió, từ nối chuẩn. Giọng đọc phải tương thích với việc chọn giọng, chọn tốc độ, âm thanh; .
- Các thông báo cho người dùng về câu đối thoại, người đọc, điểm số phải chính xác; hỗ trợ đặc lực cho người học trong việc luyện đối thoại.

2.1.2. Các yêu cầu phi chức năng

2.1.2.1. Nhằm lợi ích của đối tượng người học

- Tính tiện dụng: giao diện thân thiện bằng Tiếng Việt;
- Tính sáng tạo: Hệ thống có cơ chế nhận dạng giọng nói một cách tự động là một khả năng mới mà các sản phẩm luyện Tiếng Anh hiện nay chưa có; gây hứng thú cho người học tạo kết quả học tập tốt hơn;
- Tính đối xứng: Sản phẩm cung cấp cho người học các chức năng có các cặp loại trừ nhau như: Thêm-Bớt, Tăng-Giảm, chỉnh sửa Cao-Thấp; làm tăng thêm tính tiện dụng cho người học.

2.1.2.2. Nhằm lợi ích việc phát triển dự án

- Tính dễ phát triển và hoàn thiện: có thể mở rộng, tăng cường thêm chức năng một cách dễ dàng;
- Tính mô-đun: có sự độc lập chức năng của các thành phần trong chương trình.

2.2. ĐẶC TẢ CHỨC NĂNG CHÍNH

2.2.1. Chức năng luyện đọc từ hay câu Tiếng Anh

Chức năng này hỗ trợ người học luyện nghe, đọc những câu Tiếng Anh khó đã được bỏ qua bên bài đối thoại đã luyện từ trước, nhằm giúp người học đạt được kết quả cao hơn khi thực hiện lại bài đối thoại đó. Đồng thời, người học cũng có thể luyện những từ hay câu bất kỳ nào thấy cần thiết. Quy trình Luyện từ hay câu của người học được thực hiện như sau:

- Hệ thống khởi tạo đối tượng nhận dạng và phát âm.
- Người học thiết lập một danh sách các từ và câu để luyện phát âm bằng cách mở file để lấy danh sách các từ và câu đã lưu; hoặc thêm từ, thêm câu vào danh sách; hoặc xóa từ, xóa câu khỏi danh sách. Hệ thống cũng đồng thời kích hoạt danh sách các từ và câu dùng cho nhận dạng; tiến hành kiểm tra sự tồn tại của các đối tượng nhận dạng và phát âm.
- Sau khi đã hoàn thành việc thiết lập danh sách từ và câu để luyện phát âm, người học có thể chọn một trong hai chức năng sau để thực hiện: nghe đọc từ hay câu với giọng phát âm chuẩn; đọc để hệ thống nhận dạng và đánh giá.

Ngoài ra với chức năng nghe đọc từ hay câu, người học còn có thể chọn giọng đọc; chọn tốc độ, độ to nhỏ cho giọng đọc đó. Hệ thống còn hỗ trợ người học có thể lưu lại danh sách các từ và câu chưa luyện thành công để lần sau luyện lại.

2.2.2. Chức năng luyện đọc một đoạn đối thoại

Chức năng này tạo thêm hứng thú cho người học trong việc luyện phát âm. Khi luyện phát âm với một danh sách các từ và câu đã nhằm chán, người học có thể soạn ra một bài đối thoại liên quan đến những từ và câu mới đã được luyện bên phần Luyện từ và câu để đối thoại trực tiếp với máy và nhận được điểm tổng kết từ hệ thống. Hệ thống thực hiện các chức năng như sau:

- Hệ thống cho phép người học chọn một bài đối thoại trong các bài đối thoại đã soạn và hiển thị nội dung bài đối thoại vừa chọn.
- Người học nhấn nút **Bắt Đầu** để khởi động chương trình đối thoại. Hệ thống sẽ xác định vai người học cần đóng, phân loại và xử lý lần lượt từng câu đối thoại một cho đến hết đoạn đối thoại.
- Hệ thống yêu cầu người học phát âm và đánh giá phát âm của người học nếu đó là câu đối thoại thuộc vai người học đóng, các câu đối thoại còn lại sẽ được đọc với một giọng phát âm chuẩn.
- Kết thúc phát âm mỗi câu đối thoại của máy, hệ thống tự động chuyển sang câu đối thoại tiếp theo.
- Kết thúc phát âm mỗi câu đối thoại của người học hệ thống sẽ tự động chuyển sang câu đối thoại tiếp theo nếu điểm số đó đạt đủ điểm trung bình, tức lớn hơn hoặc bằng 50/100; nếu không hệ thống sẽ yêu cầu người học đọc lại câu đối thoại này. Lúc này, người học có thể đọc lại hay cũng có thể nhấn nút **Bỏ Qua** để chuyển sang câu đối thoại tiếp theo đồng thời sẽ lưu từ chưa đọc được vào File để luyện lại sau bên chức năng Luyện từ và câu.

- Mỗi khi người học nhấn nút Bỏ Qua, hệ thống sẽ tính điểm cho câu đối thoại đó đạt 0/100 và cộng vào tổng điểm. Kết thúc bài đối thoại, hệ thống sẽ hiển thị thông báo tổng điểm cho người học.
- Hệ thống còn có hỗ trợ các thông báo về: câu đang đối thoại, tên người sẽ đọc câu đối thoại đó, thông báo tới máy đọc hay người học đọc để thuận tiện cho người học thực hiện bài đối thoại. Hệ thống còn cập nhật trạng thái Form: nút Bỏ Qua sẽ vô hiệu hóa khi máy đọc; khi người học đọc, hệ thống có hỗ trợ thanh chạy chữ nhằm chạy ra những từ mà người sử dụng đã đọc đúng để dễ theo dõi.

2.2.3. Chức năng cập nhật đối thoại

2.3. PHÂN TÍCH HỆ THỐNG THEO HƯỚNG ĐỐI TƯỢNG

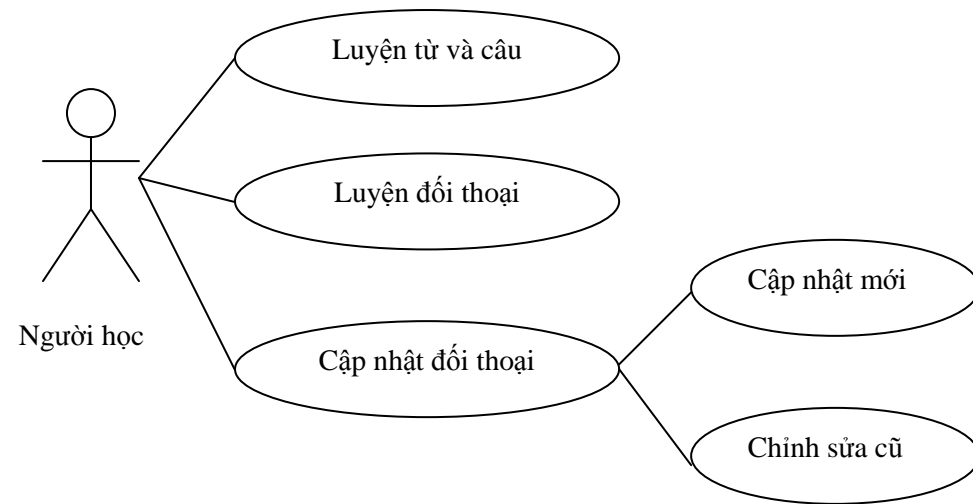
Qua khảo sát, nghiên cứu xây dựng phần mềm luyện phát âm Tiếng Anh cũng như phân tích các yêu cầu của hệ thống, để xây dựng chương trình tốt nên cần xây dựng một mô hình tốt. Đề tài này tôi đề cập đến việc ứng dụng mô hình UML (Unified Modeling Language) trong việc xây dựng mô hình cho phần mềm luyện phát âm Tiếng Anh với chức năng nhận dạng giọng nói tự động ứng dụng thư viện các hàm nhận dạng giọng nói trong bộ Microsoft SDK 5.1 do Microsoft hỗ trợ.

2.3.1. Phân tích Use case

Tác nhân (actor) là một thực thể đóng vai trò tương tác với hệ thống, tác nhân có thể là người sử dụng hệ thống hoặc một hệ thống khác. Một tác nhân có vai trò nhất định đối với Use case tương

tác. Tuy nhiên, đối với hệ thống này: Chức năng Luyện từ và câu; Luyện đối thoại; Cập nhật đối thoại đều là người học sử dụng. Người học có thể là Học Sinh, Sinh Viên, Giáo Viên,... nhưng đều có vai trò như nhau, có thể tương tác với hệ thống. Vì vậy hệ thống chỉ có một tác nhân duy nhất là Người học.

2.3.2. Lược đồ chính của mô hình Use case



Hình 2.1. Lược đồ chính của mô hình Use case

2.3.3. Mô tả các Use case

Use case luyện đọc từ và câu

* Tóm tắt

Use case này cho phép người học tự luyện phát âm hay nghe đọc từ hoặc câu Tiếng Anh.

* Dòng sự kiện chính

Use case này bắt đầu khi một actor muốn luyện nghe và luyện đọc từ hoặc câu Tiếng Anh.

1. Hệ thống khởi tạo đối tượng nhận dạng và phát âm.
2. Hệ thống đòi hỏi người học chọn chức năng muốn thực hiện để tạo danh sách từ hay câu dùng cho nhận dạng (Thêm từ, xóa từ, mở file). Một trong các luồng phụ sau được thực hiện:

+ Thêm từ: Actor nhập từ hoặc câu cần thêm vào danh sách và bấm nút **Thêm** thì từ hay câu sẽ được thêm vào danh sách.

+ Xóa từ: Actor chọn từ cần xóa trong danh sách và bấm nút **Xóa** thì từ hay câu sẽ được xóa bỏ khỏi danh sách.

+ Mở file: Actor nhấn nút **Mở file** cũng có thể tạo được một danh sách từ và câu để luyện. File này đã được lưu trước đó

3. Hệ thống kích hoạt danh sách từ dùng cho nhận dạng và kiểm tra sự tồn tại của các đối tượng nhận dạng và phát âm.
4. Actor lựa chọn một trong 3 chức năng: nghe đọc, đọc để kiểm tra, lưu file. Một trong các luồng phụ sau được thực hiện:

+ Nghe đọc: Actor chọn từ cần nghe đọc có trong danh sách rồi nhấn nút **Đọc** để nghe được giọng phát âm chuẩn. Giọng phát âm này tương thích với các tùy chọn: giọng đọc, âm thanh, tốc độ mà Actor đã chọn trước đó.

+ Đọc để kiểm tra: Actor phát âm một từ hay câu có trong danh sách để hệ thống kiểm tra nhận dạng. Hệ thống sẽ hiển thị từ vừa nhận dạng được và chọn từ này trong danh sách. Actor theo dõi nếu từ hiển thị cũng như từ được chọn trùng với từ mà mình đọc thì tự biết mình đã đọc đúng, ngược lại là đã đọc sai, cần luyện lại.

+ Lưu file: Actor nhấn nút **Lưu file** để lưu danh sách này vào file. Hệ thống mở file danh sách, xóa danh sách cũ, lưu danh sách mới vào file.

Use case luyện đối thoại

* Tóm tắt

Use case này cho phép người học đối thoại trực tiếp với máy dựa trên nội dung một bài đối thoại soạn sẵn. Trong quá trình thực hiện đối thoại, người học có thể theo dõi được kết quả phát âm của mình, bỏ qua và lưu câu khó khi cần.

* Dòng sự kiện chính

Use case này bắt đầu khi người học muốn luyện đọc một đoạn đối thoại xác định đã chọn và nhấn nút **Bắt Đầu**

1. Dựa trên bài đối thoại, hệ thống xác định tên vai người sử dụng đóng trong đoạn đối thoại.

2. Hệ thống kiểm tra câu kết thúc bằng dấu “:” để xác định tên người sẽ nói câu đối thoại.

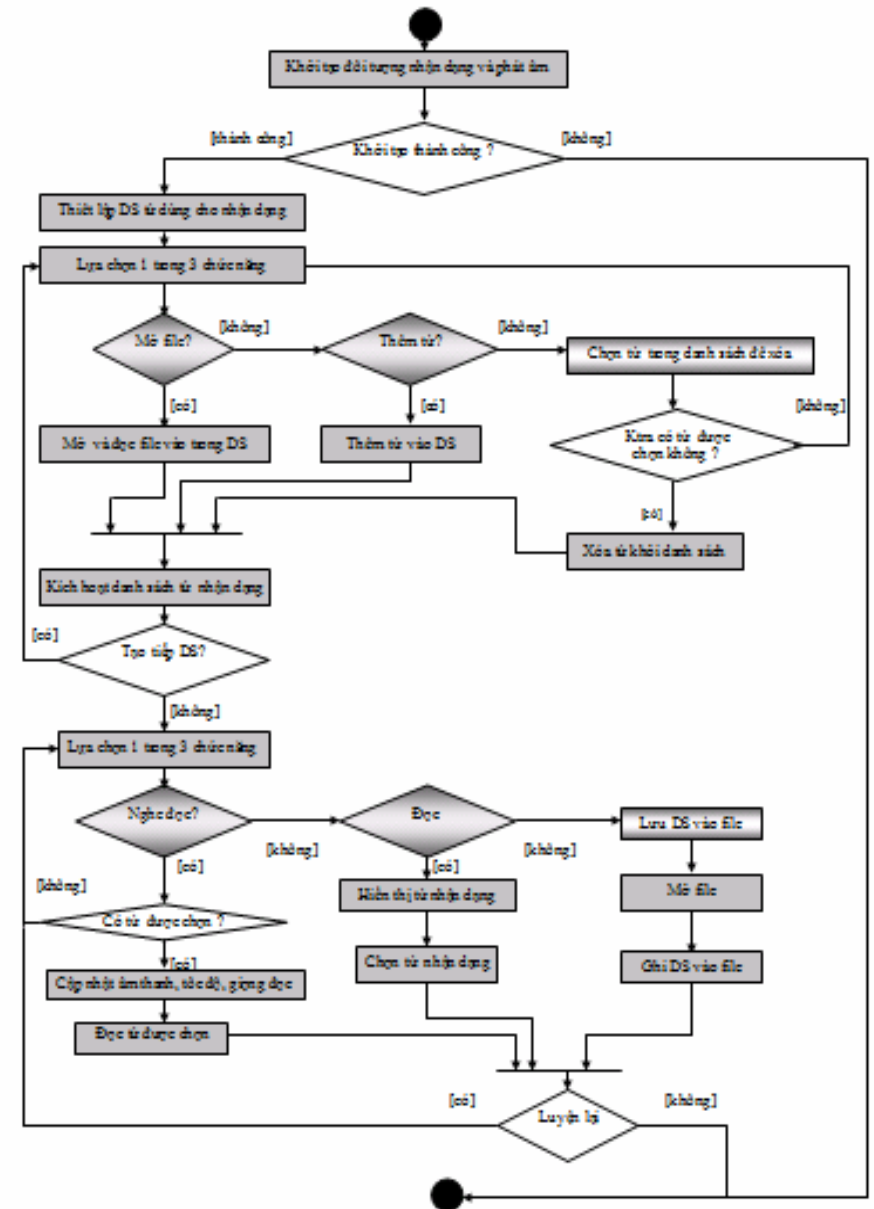
3. Hệ thống so sánh tên người nói có trùng tên vai người sử dụng đóng vai hay không và đưa ra xử lý. Một trong các luồng phụ sau được thực hiện:

- Nếu tên người nói trùng với tên vai: Hệ thống kích hoạt nhận dạng câu đối thoại. Kết thúc phát âm của Actor hệ thống kiểm tra điểm số để chuyển sang câu đối tiếp theo

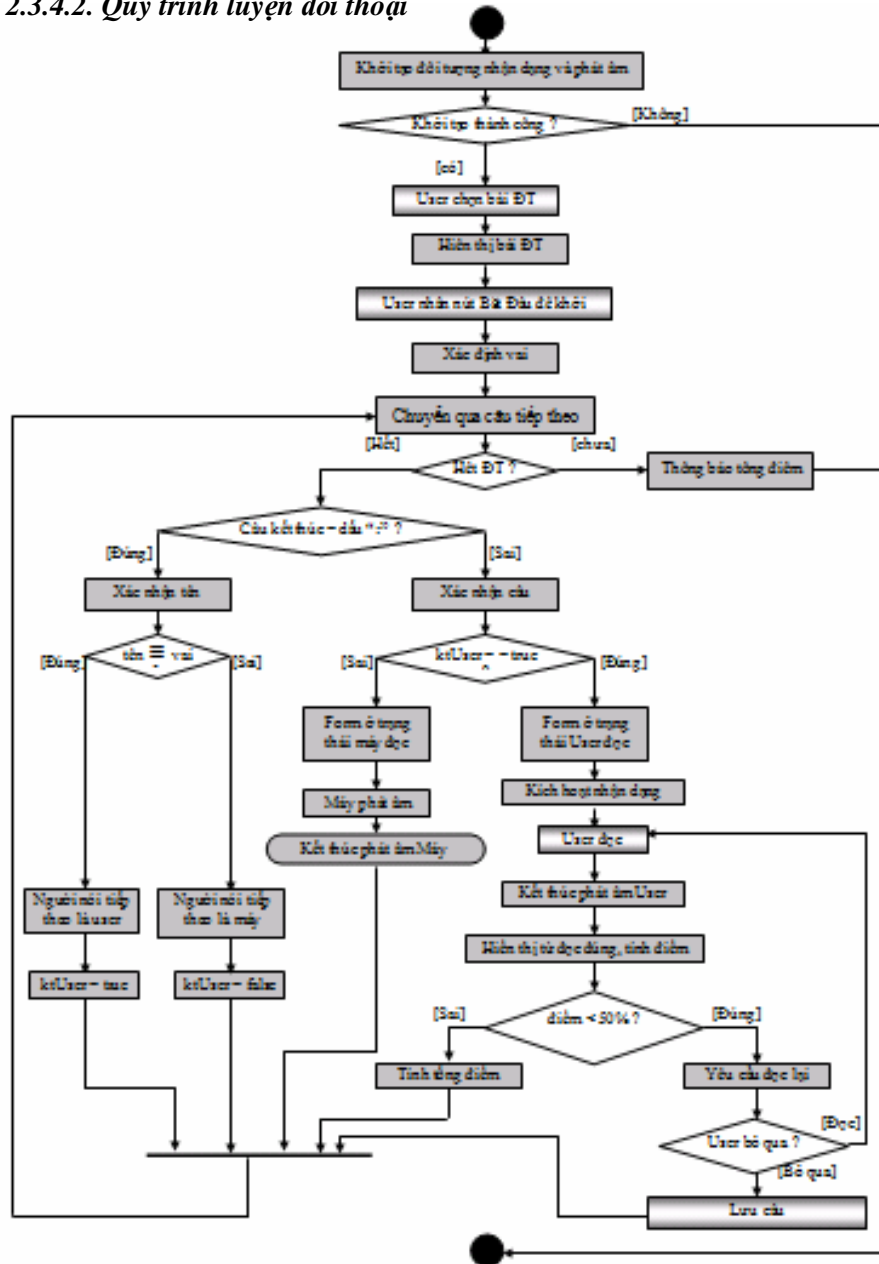
- Nếu tên người nói không trùng với tên vai: có nghĩa là câu đối thoại dưới thuộc máy đọc, thì sẽ cho phép máy đọc câu đối thoại với giọng phát âm chuẩn, sau khi kết thúc phát âm của máy tự động chuyển sang câu tiếp theo.
4. Kết thúc đối thoại, hệ thống thông báo điểm số tổng cộng cho cả bài đối thoại.

2.3.4. Các quy trình hỗ trợ

2.3.4.1. Quy trình Luyện từ và câu



2.3.4.2. Quy trình luyện đối thoại



2.3.4.3. Quy trình cập nhật đối thoại mới

2.3.4.4. Quy trình chỉnh sửa đối thoại cũ

CHƯƠNG 3: PHÁT TRIỂN ỨNG DỤNG

Trên cơ sở phân tích thiết kế hệ thống ở chương trước, trong chương này, chúng tôi tiến hành thiết kế giao diện, viết mã cho ứng dụng. Chương này sẽ trình bày các nội dung như sau: thiết kế chức năng, thiết kế giao diện và demo chương trình.

3.1. THIẾT KẾ CHỨC NĂNG

3.1.1. Luyện từ và câu

3.1.1.1. Thủ tục thêm từ vào danh sách

3.1.1.2. Thủ tục xóa từ khỏi danh sách

3.1.1.3. Thủ tục đọc từ có trong danh sách

3.1.1.4. Khởi động lại danh sách từ dùng cho nhận dạng

3.1.1.5. Khai báo ủy quyền cho sự kiện nhận dạng giọng nói

3.1.2. Luyện đối thoại

3.1.2.1. Thủ tục chọn bài

3.1.2.2. Thủ tục bỏ qua câu đối thoại

3.1.2.3. Hàm xử lý câu đối thoại

3.1.2.4. Hàm xử lý khi kết thúc giọng đọc của User

3.1.3. Cập nhật đối thoại

3.1.3.1. Cập nhật mới

3.1.3.2. Chỉnh sửa cũ

2.3.2.3. Đổi vai

3.2. THIẾT KẾ GIAO DIỆN

- 3.2.1. Form chính của chương trình
- 3.2.2. Form luyện đọc từ hay câu Tiếng Anh
- 3.2.3. Form luyện đọc một đoạn đối thoại
- 3.2.4. Form cập nhật mới đối thoại
- 3.2.5. Form chỉnh sửa đối thoại
- 3.2.6. Form hướng dẫn

3.3. DEMO CHƯƠNG TRÌNH

- 3.3.1. Chức năng Luyện từ và câu
- 3.3.2. Chức năng Luyện đối thoại
- 3.3.3. Chức năng cập nhật đối thoại
 - 3.3.3.1. *Cập nhật mới*
 - 3.3.3.2. *Chỉnh sửa cũ*

KẾT LUẬN

Phần mềm đã đáp ứng được chức năng đánh giá phát âm Tiếng Anh cho người học tương đối chính xác; giao diện thân thiện, dễ sử dụng; với tính năng nhận dạng phát âm mới mẻ sẽ mang đến hứng thú cho hoạt động luyện phát âm của người học tiếng Anh.

Chúng tôi đã cài đặt phần mềm này cho máy tính ở toàn bộ phòng LAP ở trường Cao đẳng Kinh tế - Kế hoạch Đà Nẵng và đưa vào cho sinh viên tại trường thực hành luyện phát âm trong một số giờ học Anh văn. Theo dõi kết quả học tập, cũng như thăm dò ý kiến của các bạn sinh viên và giáo viên bộ môn, kết quả cho thấy đã mang lại một kết quả nhất định; sinh viên trở nên hứng thú trong việc luyện, thực hành đối thoại với máy; giúp sinh viên chủ động, có ý thức và tiến bộ hơn hẳn trong việc luyện phát âm của mình; đồng thời sinh viên tự phân biệt cũng như phát âm được các từ có phiên âm quốc tế gần giống nhau.

Bên cạnh những kết quả đạt được như trên, phần mềm cũng không tránh khỏi những thiếu sót như: khi người học chưa đọc đúng thì mới chỉ hỗ trợ giọng đọc chuẩn chứ không có phiên âm quốc tế để kiểm tra, luyện đọc; hệ thống không kiểm tra chính tả của từ hay ngữ pháp của câu nhập vào.

Chúng tôi dự kiến hướng phát triển thêm cho đề tài này bằng việc mở rộng thêm chức năng tra cứu từ điển, người học có thể xem được phiên âm quốc tế, xem nghĩa của từ; hỗ trợ thêm những bài nghe theo chủ đề soạn sẵn; những bài nghe có kèm bài test kiểm tra trình độ nghe hiểu cho người học nhằm mang lại hiệu quả luyện phát âm tốt hơn nữa trong thời gian sắp tới.