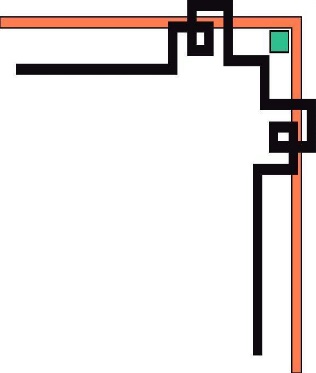
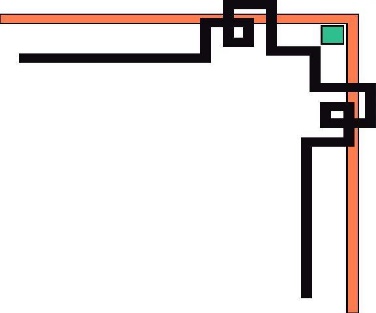
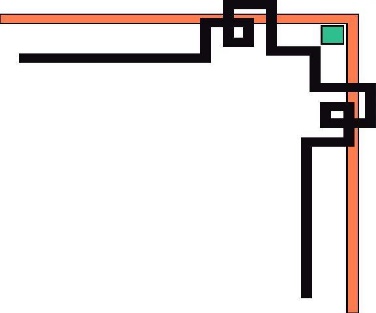
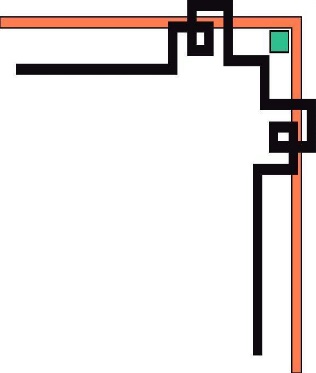
**BỘ GIÁO DỤC VÀ ĐÀO TẠO**



**TRƯỜNG ĐẠI HỌC** **SƯ PHẠM TP. HCM**

**KHOA** **CÔNG NGHỆ THÔNG TIN**

****

**BÁO CÁO ĐỒ ÁN  
CÁC HỆ CƠ SỞ TRI THỨC**

***ĐỀ TÀI***: **Xây dựng chương trình chuyển đề bài toán hình học phẳng theo ngôn ngữ tự nhiên sang ngôn ngữ đặc tả**

Tên Nhóm: Nhóm 4

**Giáo viên hướng dẫn: \_ TS. Nguyễn Đình Hiển**

**Năm 2021 - 2022**

**Lời Nói Đầu**

Hệ thống dựa trên tri thức (Knowledge-Based System - KBS) là một dạng trí tuệ nhân tạo (Artificial intelligence - AI) nhằm mục đích nắm bắt kiến ​​thức của các chuyên gia về con người để hỗ trợ việc ra quyết định. Ví dụ về hệ thống dựa trên tri thức bao gồm hệ thống chuyên gia, được gọi như vậy vì sự phụ thuộc vào chuyên môn của con người.

Trong bài báo cáo này, chúng em xin trình bày về cách chúng em xây dựng chương trình chuyển đề bài toán hình học phẳng theo ngôn ngữ tự nhiên sang ngôn ngữ đặc tả trên nền tảng web thông qua Regex (Regular Expression) và Django của Python.

Qua đây, chúng em xin được gửi lời cảm ơn đến TS. Nguyễn Đình Hiển, người đã tận tâm truyền đạt những kiến thức nền tảng cơ bản về môn học “Các hệ cơ sở tri thức", để chùn em có cơ sở kiến thức để có thể làm bài báo cáo này. Do kiến thức còn hạn hẹp, bài thu hoạch có thể có những sai sót nhất định, mong thầy góp ý để bài báo cáo ngày càng hoàn thiện hơn. Chúng em rất cảm ơn thầy!.

**Bảng phân công việc**

|  |  |  |  |
| --- | --- | --- | --- |
| ***STT*** | ***Họ và tên*** | ***Mssv*** | ***Công việc*** |
| ***1*** | **Nguyễn Văn Phong** | **4501104175** | Xây dựng thuật toán, thiết kế API |
| ***2*** | **Huỳnh Thanh Phong** | **4501104172** | Thu thập phân tích xây dựng quan hệ giữa các khái niệm, thiết kế giao diện,  tìm hiểu Django. |
| ***3*** | **Lê Tấn Lộc** | **4501104135** | Thiết kế Regex, thiết kế word, tìm hiểu về Regex |
| ***4*** | **Lê Văn Trung** | **4501104259** | Thiết kế Regex, thiết kế powerpoint, thử nghiệm thuật toán. |

**Mục lục**

[NỘI DUNG 8](#_Toc90447770)

[1 Mục đích và yêu cầu 8](#_Toc90447771)

[1.1 Mục đích 8](#_Toc90447772)

[1.2 Yêu cầu 8](#_Toc90447773)

[2 Giới thiệu 8](#_Toc90447774)

[2.1 Regex 8](#_Toc90447775)

[2.2 Django 8](#_Toc90447776)

[3 Bài toán 8](#_Toc90447777)

[4 Quy ước và cú pháp 9](#_Toc90447778)

[4.1 Qui ước 9](#_Toc90447779)

[4.2 Cú pháp 10](#_Toc90447780)

[4.2.1 Cú pháp câu viết cho giả thiết là một khái niệm 10](#_Toc90447781)

[10](#_Toc90447782)

[4.2.2 Cú pháp câu viết cho một số giả thiết là sự liên hệ giữa nhiều khái niệm 11](#_Toc90447783)

[4.2.2.1 Điểm - điểm 11](#_Toc90447784)

[4.2.2.2 Điểm - đường thẳng 11](#_Toc90447785)

[4.2.2.3 Điểm - đoạn 11](#_Toc90447786)

[4.2.2.4 Điểm – tia 11](#_Toc90447787)

[4.2.2.5 Điểm - đường tròn 12](#_Toc90447788)

[4.2.2.6 Đoạn - đường tròn 12](#_Toc90447789)

[4.2.2.7 Điểm - đoạn - đường tròn 12](#_Toc90447790)

[4.2.2.8 Đoạn - Đoạn 12](#_Toc90447791)

[4.2.2.9 Tia – Tia 13](#_Toc90447792)

[4.2.2.10 Đoạn – Tia 13](#_Toc90447793)

[4.2.2.11 Điểm – Tia – Đoạn 13](#_Toc90447794)

[4.2.2.12 Đoạn - tam giác 13](#_Toc90447795)

[4.2.2.13 Đoạn - tứ giác 13](#_Toc90447796)

[4.2.2.14 Góc – Góc 13](#_Toc90447797)

[4.2.2.15 Góc - Đoạn, Góc – Tia 14](#_Toc90447798)

[4.2.2.16 Tam giác - Tam giác 14](#_Toc90447799)

[4.2.3 Cú pháp câu viết cho một số giả thiết về khai báo giá trị 14](#_Toc90447800)

[4.2.4 Cú pháp câu viết cho một số khẳng định trong yêu cầu chứng minh 15](#_Toc90447801)

[5 Mô hình xử lý 15](#_Toc90447802)

[6 Chương trình 16](#_Toc90447803)

[6.1 Giao diện ban đầu khi vào web 16](#_Toc90447804)

[6.2 Giao diện khi nhập bài toán 17](#_Toc90447805)

[6.3 Giao diện khi hiển thị kết quả 18](#_Toc90447806)

[7 Kết luận 18](#_Toc90447807)

[7.1 Ưu điêm 18](#_Toc90447808)

[7.2 Nhược điểm. 19](#_Toc90447809)

[8 Tài liệu tham khảo 19](#_Toc90447810)

**DANH MỤC HÌNH ẢNH**

[Hình 2‑1 Chuổi tổng quát kiểm tra có phải email hay không thông qua Regex 7](#_Toc90447669)

[Hình 4‑1 Cú pháp câu viết cho giả thiết 9](#_Toc90447670)

[Hình 4‑2 Cú pháp quan hệ điểm - điểm 10](#_Toc90447671)

[Hình 4‑3 Cú pháp quan hệ Điểm - Đường thẳng 10](#_Toc90447672)

[Hình 4‑4 Cú pháp quan hệ Điểm – Đoạn 10](#_Toc90447673)

[Hình 4‑5 Cú pháp quan hệ Điểm – Tia 10](#_Toc90447674)

[Hình 4‑6 Cú pháp quan hệ Điểm – Đường tròn 11](#_Toc90447675)

[Hình 4‑7 Cú pháp quan hệ Đoạn – Đường tròn 11](#_Toc90447676)

[Hình 4‑8 Cú pháp quan hệ Điểm - Đoạn – Đường tròn 11](#_Toc90447677)

[Hình 4‑9 Cú pháp quan hệ Đoạn – Đoạn 11](#_Toc90447678)

[Hình 4‑10 Cú pháp quan hệ Tia - Tia 12](#_Toc90447679)

[Hình 4‑11 Cú pháp quan hệ Đoạn - Tia 12](#_Toc90447680)

[Hình 4‑12 Cú pháp quan hệ Điểm – Tia – Đoạn 12](#_Toc90447681)

[Hình 4‑13 Cú pháp quan hệ Đoạn – Tam giác 12](#_Toc90447682)

[Hình 4‑14 Cú pháp quan hệ Đoạn – Tứ giác 12](#_Toc90447683)

[Hình 4‑15 Cú pháp quan hệ Góc - Góc 12](#_Toc90447684)

[Hình 4‑16 Cú pháp quan hệ Góc – Đoạn, Góc - Tia 13](#_Toc90447685)

[Hình 4‑17 Cú pháp quan hệ Tam giác – Tam giác 13](#_Toc90447686)

[Hình 4‑18 Cú pháp câu viết cho một số giả thiết về khai báo giá trị 13](#_Toc90447687)

[Hình 4‑19 Cú pháp câu viết cho một số khẳng định trong yêu cầu chứng minh 14](#_Toc90447688)

[Hình 5‑1 Giao diện khi hiển thị kết quả 14](#_Toc90447689)

[Hình 6‑1 Giao diện ban đầu khi vào web 15](#_Toc90447690)

[Hình 6‑2 Giao diện khi nhập bài toán 16](#_Toc90447691)

[Hình 6‑3 Giao diện khi hiển thị kết quả 17](#_Toc90447692)

# NỘI DUNG

# Mục đích và yêu cầu

## Mục đích

Xây dựng một trang web có thể chuyển đổi bài toán hình học phẳng từ ngôn ngữ tự nhiên sang ngôn ngữ đặc tả.

## Yêu cầu

* Giao diện thân thiện, dễ sử dụng
* Chuyển đổi bài toán từ ngôn ngữ tự nhiên qua các quy ước và cú pháp đã định nghĩa trước
* Hiển thị rõ hai thông tin là giả thuyết và kết luận

+ Giả thiết: là các phát biểu, thông tin, hay gọi là những giả thiết mà đề bài cho trước, hoặc đề cập đến.

+ Kết luận: là những yêu cầu tìm ra một phát biểu (thông tin) mới dựa vào các thông tin giả thiết cho trước (bao gồm việc tính hay chứng minh một giả thiết mới nào đó bất kì).

# Giới thiệu

## Regex

Regex (Regular Expression) là thư viện hỗ trợ cho việc đối sánh mẫu với chuỗi hoặc đối sánh chuỗi, giống như thao tác “tìm” và “tìm và thay thế” thông thường. Regex cho phép đưa ra một chuỗi tổng quát để so khớp với các chuỗi. Thư viên này có thể hoạt động trên nhiều ngôn ngữ như C++, Java, Python, ...

Ví dụ: Để kiểm tra một chuỗi kí tự có phải là email hay không, ta có chuỗi tổng quát như sau:



Hình 2‑1 Chuổi tổng quát kiểm tra có phải email hay không thông qua Regex

## Django

Django là một web framework Python cấp cao, khuyến khích sự phát triển nhanh chóng và thiết kế sạch sẽ, thực dụng. Được xây dựng bởi các nhà phát triển có kinh nghiệm, nó xử lý phần lớn sự phức tạp của việc phát triển web. Django là mã nguồn mở và miễn phí.

# Bài toán

***Ta có ví dụ như sau:*** Cho tam giác vuông ABC vuông tại A.

Cho điểm H với AH là đường cao của tam giác ABC.

Ta có BH=1, HC=4.

Tính: AB, AC

***Ta có hai thông tin như sau:***

*Giải thuyết:*

Tam giác vuông ABC (vuông tại A);

AB là đường cao tam giác vuông ABC;

BH=1;

HC=4;

*Kết luận:*

Tính: AB, AC;

***Trong đó:*** Giả thiết (kết luận) là một số phát biểu được cho trước, và mỗi phát biểu được viết thành từng câu tương ứng theo một quy ước cụ thể.

Mỗi câu cho phép nhập một đặc tả tự nhiên theo quy ước và mỗi câu được phân biệt nhau bằng dấu kết thúc câu là dấu chấm phẩy ( ; )

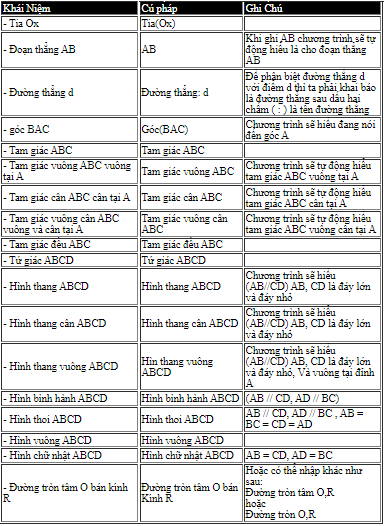
# Quy ước và cú pháp

## Qui ước

* Quy ước nhập theo từng câu: mỗi câu là một phát biểu cho một giả thiết có trong đề bài toán.
* Cách nhập một câu được đặc tả theo một cú pháp được quy ước bên dưới.
* Quy ước kết thúc một câu: kết thúc mỗi câu được kết thúc bằng dấu chấm phẩy ( ; )
* Một câu chỉ có nghĩa khi nhập đúng quy ước bao gồm việc nhập đúng cú pháp của một câu và có dấu kết thúc câu

## Cú pháp

### Cú pháp câu viết cho giả thiết là một khái niệm



Hình 4‑1 Cú pháp câu viết cho giả thiết

### Cú pháp câu viết cho một số giả thiết là sự liên hệ giữa nhiều khái niệm

#### Điểm - điểm



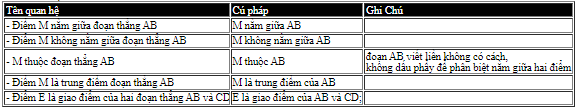
Hình 4‑2 Cú pháp quan hệ điểm - điểm

#### Điểm - đường thẳng



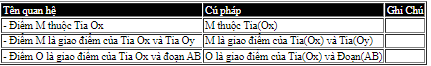
Hình 4‑3 Cú pháp quan hệ Điểm - Đường thẳng

#### Điểm - đoạn



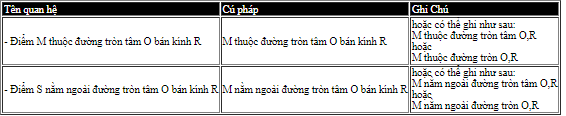
Hình 4‑4 Cú pháp quan hệ Điểm – Đoạn

#### Điểm – tia



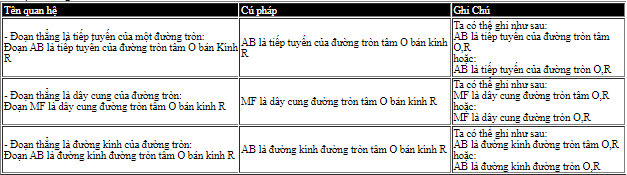
Hình 4‑5 Cú pháp quan hệ Điểm – Tia

#### Điểm - đường tròn



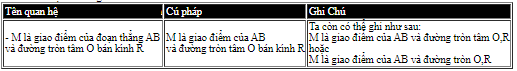
Hình 4‑6 Cú pháp quan hệ Điểm – Đường tròn

#### Đoạn - đường tròn



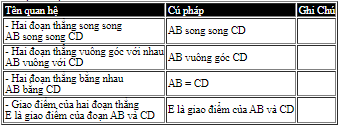
Hình 4‑7 Cú pháp quan hệ Đoạn – Đường tròn

#### Điểm - đoạn - đường tròn



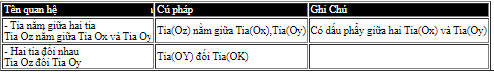
Hình 4‑8 Cú pháp quan hệ Điểm - Đoạn – Đường tròn

#### Đoạn - Đoạn



Hình 4‑9 Cú pháp quan hệ Đoạn – Đoạn

#### Tia – Tia



Hình 4‑10 Cú pháp quan hệ Tia - Tia

#### Đoạn – Tia



Hình 4‑11 Cú pháp quan hệ Đoạn - Tia

#### **Điểm – Tia – Đoạn**



Hình 4‑12 Cú pháp quan hệ Điểm – Tia – Đoạn

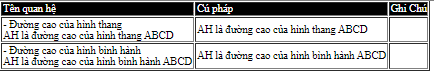
#### Đoạn - tam giác

A screenshot of a computer

Description automatically generated

Hình 4‑13 Cú pháp quan hệ Đoạn – Tam giác

#### Đoạn - tứ giác



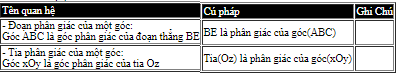
Hình 4‑14 Cú pháp quan hệ Đoạn – Tứ giác

#### Góc – Góc



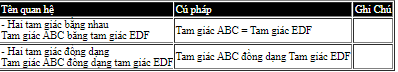
Hình 4‑15 Cú pháp quan hệ Góc - Góc

#### Góc - Đoạn, Góc – Tia



Hình 4‑16 Cú pháp quan hệ Góc – Đoạn, Góc - Tia

#### Tam giác - Tam giác



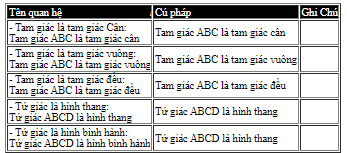
Hình 4‑17 Cú pháp quan hệ Tam giác – Tam giác

### Cú pháp câu viết cho một số giả thiết về khai báo giá trị



Hình 4‑18 Cú pháp câu viết cho một số giả thiết về khai báo giá trị

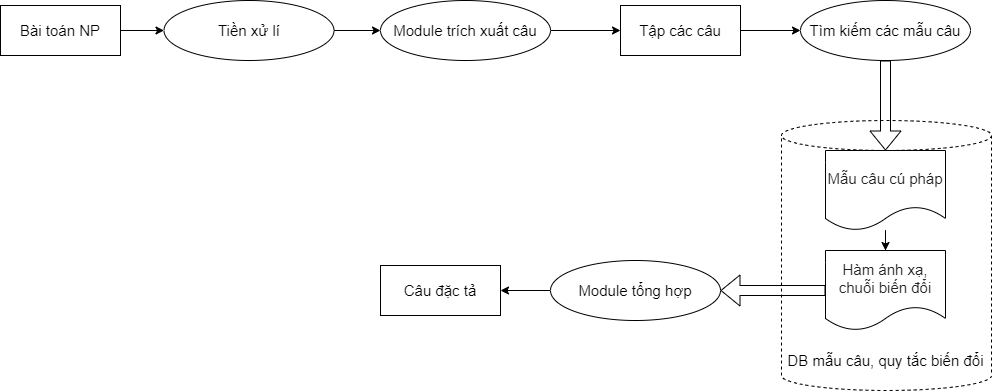
### Cú pháp câu viết cho một số khẳng định trong yêu cầu chứng minh



Hình 4‑19 Cú pháp câu viết cho một số khẳng định trong yêu cầu chứng minh

# Mô hình xử lý

Bằng cách phân tích và nghiên cứu về ngôn ngữ đặc tả và mô hình COKB giải quyết bài toán hình học 2D, nhóm chúng em đưa ra phương mô hình để xử lí.



Hình 5‑1 Giao diện khi hiển thị kết quả

Từ dữ liệu nhập vào là bài toán dạng ngôn ngữ tự nhiên sẽ được tiền xử lý, loại bỏ các kí thừa. Sau đó, module trích xuất câu thực hiện tách bài toán thành nhiều câu dựa trên các tiêu chí khác nhau. Từ tập dữ liệu các câu đã có ở bước trước, tiến hành tìm kiếm các mẫu câu (pattern) mà nhóm đã nghiên cứu và xây dựng từ trước. Ở bước này, nhóm sử dụng Regular Expression để thiết kế các mẫu câu sau đó xây dựng các hàm ánh xạ chuyển đổi thành cú pháp tương ứng với quan hệ. Các câu ban đầu sẽ qua hàm ánh xạ, chuỗi biến đổi quy tác cho trước rồi đưa vào module tổng hợp để xuất ra được bài toán được mô tả dưới dạng ngôn ngữ đặc tả. Các bước này phức tạp nhất là quy trình thiết kế các mẫu câu để tránh bị xung đột làm mất dữ liệu, ảnh hưởng đến mẫu câu khác.

# Chương trình

Trang web được xây dựng dựa trên ngôn ngữ Python và sử dụng frameword Django với giao diện chính cho phép nhập vào bài toán ngôn ngữ tự nhiên và ouput là ngôn ngữ đặc tả để có thể đưa vào hệ thống COKB.

## Giao diện ban đầu khi vào web

A picture containing diagram

Description automatically generated

Hình 6‑1 Giao diện ban đầu khi vào web

## Giao diện khi nhập bài toán

A picture containing diagram

Description automatically generated

Hình 6‑2 Giao diện khi nhập bài toán

## Giao diện khi hiển thị kết quả

Text

Description automatically generated with low confidence

Hình 6‑3 Giao diện khi hiển thị kết quả

# Kết luận

## Ưu điêm

* Giao diện thân thiện, đẹp mắt, dễ sử dụng
* Hiển thị rõ hai thông tin là giả thuyết và kết luận
* Danh sách các bài toán mẫu
* Chuyển đổi bài toán nhanh chóng, giúp các bài toán dạng ngôn ngữ tự nhiên có thể gián tiếp đưa vào hệ thống COKB để giải bài toán

## Nhược điểm.

* Chưa có hệ thống gợi ý các bài toán tương tự, mà chỉ có danh sách bài toán mẫu cố định
* Chưa hỗ trợ các bài toán phổ thông như vector, hình học không gian.

# Tài liệu tham khảo

[1] CUPHAPNHAPBAITOANHINHHOC – Nguyễn Đình Hiển.

[2] Python RegEx – W3Shools, <https://www.w3schools.com/python/python_regex.asp>

[3] regex101: build, test, and debug regex, <https://regex101.com/>

[4] Do, Nhon, Hoai Phan Truong, and Tuyen Trong Tran. "An approach for translating mathematics problems in natural language to specification language COKB of intelligent education software." *2010 International Conference on Artificial Intelligence and Education (ICAIE).* IEEE, 2010.

**LỜI CẢM ƠN**

Chúng em xin chân thành cảm ơn đến thầy TS. Nguyễn Đình Hiển đã giúp chúng em về tài liệu kham khảo và kiến thức của môn học. Nhóm em đã hoàn thiện đồ án “Xây dựng chương trình chuyển đề bài toán hình học phẳng theo ngôn ngữ tự nhiên sang ngôn ngữ đặc tả” mặc dù cố gắng nhưng không thể tránh khỏi những thiếu sót. Chương trình chưa được hoàn thiện do nhóm chưa có kinh nghiệm nhiều trong việc xây dựng một chương trình. Rất mong thầy đóng góp ý kiến để chương trình chúng em ngay càng được hoàn thiện trong tương lai.