Ontology Design and Development for University Admission in Sri Lankan Universities

T. S. D. H. S. Wickramarathna[1], C. Anutariya[2]
[1]Uva Wellassa University, Badulla, Sri Lanka.
[2]Asian Institute of Technology, Pathumthani, Thailand.

Abstract

University Admission process is a complex and time-consuming task which requires experts' advice in all the stages. Since the different stages belong to different organizations, it is rare to find experts having an overall knowledge of the process, from whom anyone can get a complete guide. Local universities in Sri Lanka are fully government funded but able to admit only around 10% of the qualified students. This leads to a high competition in admission but proper guidance in the admission process is still not established. A research has been conducted to develop a recommendation system to encapsulate the complexity of university admission process and help students to take the most suitable decision. In this paper, design and development of the Ontology, which is the knowledge base for the recommendation system is discussed. This Ontology model the undergraduate admission scenario in Sri Lanka, together with the career path mapping, so that students can far see their future opportunities before taking decisions on advanced level subject selection and university application. Ontology was developed using Protégé and was tested by running SPARQL queries. SPARQL query results confirmed that the ontology design is comprised with the real world.

Keywords: *University Admission, Ontology with Protégé, Career Path Mapping, SPARQL Queries.*

1. Introduction

University admission is a complex decision process that goes beyond simply matching test scores and admission requirements. Whether a student chooses a right university and right degree program, and conversely whether a university chooses a right student, decides the success of both institute and the students in carrying through the education. Admission and enrollment studies are generally focused on the university or collage perspective, and studies based on secondary school subject selection perspective is limited.

Higher education at local universities in Sri Lanka is fully government funded. However, due to lack of budget allocation and limited facilities, local universities are able to admit only around 10% of the qualified students [2]. This leads to a high competition in the admission, but proper guidance in admission process is still not established. Therefore, it is crucial to detect the influential factors in order to be able to reduce the failures of qualified students in admission. Data is available located in different data stores, with different names for identical entities. For this reason, the entities of the university domain and the relations between them must be defined.

There are around 100 distinct degree streams offered by Universities in Sri Lanka. Each stream has minimum subject requirements from Advanced level subjects while some degrees may have an additional requirement such as Ordinary Level results and aptitude tests. Each degree stream has planned intake for the particular academic year. Universities offer degrees under degree streams through different faculties and departments. Same time these degrees mapped to certain career paths according to the standards set by organizations and companies.

In summary, major problems and difficulties in Sri Lankan University admission system can broadly be identified as follows [3].

- Limited facilities, high competitiveness
- Limited Guidance in A-Level subject selection
- Limited awareness about specific entry requirements in some universities
- Mismatch between labor market supply and demand

- Limited awareness about other territory education options

Therefore, a recommendation system would be much helpful for students as it will encapsulate this complexity and help students to take the most suitable decision. In this study, ontology will be developed, as it will be the knowledge base for the future recommendation system. Ontology was developed using Protégé as Protégé is the most popular tool for ontology editing and developing [8]. This ontology model the undergraduate admission scenario in Sri Lanka together with the career path mapping such that the data can be published by University Grants Commission (UGC), universities and companies to be referred by students to far see their future opportunities. Though this study is limited only to Sri Lankan National university admission system, it can simply be expanded and applied to other countries with minor changes.

2. Related Work

There are several possible approaches to solve the mentioned problem. Considering the limitations and known problems in other approaches, Linked Data approach is selected and Ontology was developed as the knowledge base.

Ontology is the statement of a logical theory. The term Ontology comes from philosophy. It is an explicit specification of a conceptualization. In AI field, it believes that whatever exists can be represented. When the knowledge of a domain is represented in a declarative form, the represented set of objects is called the universe of discourse. These sets of objects along with the relationships described among them are reflected in the representational vocabulary. The knowledge-based program represents knowledge using this vocabulary [12].

In AI ontology is described by defining a set of representational terms. The definitions associate the names of entities in the universe of discourse with human- readable text by describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms [13]. Gruber [12] defines an ontology as explicit formal specifications of the terms in the domain and relations among them. An ontology is consists of definitions of concepts in a domain and relations among these concepts. These definitions allow developing artificially intelligent applications as they are expressed in a machine-interpretable way. Ontologies are considered as the cornerstone of the semantic web. Ontologies will capture the domain knowledge in a generic way in order to provide a clear understanding of a domain. They are referred as the shared conceptualization of a domain since they include the representation of these conceptualizations [7]. Ontologies do not depend on the application that uses them. This independence makes the software and knowledge maintenance easier and enhances the semantic interoperability between applications [8]. Therefore, it can be claimed that ontologies constitute the backbone of the Semantic Web. Ontologies can be used very effectively in different application fields and they provide a specification of a formal common representation of an application field trying to reduce the distance between the way people and machines manage information.

Several software tools are required in the ontology development process [8]. There are number of open-source as well as commercial tools are available for ontology development and they referred as Ontology editors. Some tools can be used throughout the entire ontology development process whereas some tools can be used in several stages of the ontology lifecycle like designing, implementation and maintenance of ontologies. Protégé is a free, open-source ontology and knowledge base editor and framework produced by Stanford University for building intelligent systems. Protégé is used by diverse of users for build knowledge-based solutions in several areas. Protégé provides facility to constructing domain ontologies as well as customized data entry forms to enter data. Classes, variables, class hierarchies, the relationship between classes, variable value restrictions and properties of relationships can be defined using Protégé.

Even though there is no any ontology developed to cover up the entire university admission process, there are few initiatives to model the university by creating ontology for that. Naveen et al. [10] have created a university ontology using Protégé as the tool. They

performed their task in seven steps starting with defining the classes and class hierarchy followed by defining object properties and data properties, identifying the relationships between properties, axioms of ontology and finally the reasoning of ontology. There are some other similar initiatives in a related topic of the university in education domain such as developing university ontology, by Sanjay Malik in 2010 [4]. In their ontology, they have the emphasis on the university employee detail only based on the name of the employee, date of joining , address etc. and ontology based on course, Ling Zeng, et al. 2009 [11] which focus on a particular course to reuse of course for teaching purpose. Ayesha Ameen, et al in [5] developed a university ontology mainly focused on the properties and their restrictions. They also used Protégé as the developing tool.

3. Ontology Design and Development

Protégé is chosen for implementation because it enables the construction of domain ontologies, and customized data entry forms to enter data. Protégé allows the definition of classes, class hierarchies, variables, variable-value restrictions, and the relationships between classes and properties of these relationships [4]. Moreover according to some surveys about most widely used ontology editors and most widely used domain for ontology development.

There are several design strategies introduced for ontology design [7], [9]. However design strategies presented in "Ontology Development 101: A Guide to Creating Your First Ontology" [1] was followed to develop the ontology. That strategy consists of following steps:

- Determine domain and scope of ontology
- Consider reusing existing ontologies
- List important terms in the ontology
- Define the classes and the class hierarchy
- Define properties of classes
- Define facets of the properties
- Create instances

When determining the domain and scope of the ontology, identifying the competency questions is crucial. Competency questions summarize the types of questions that the ontology should be able to provide answers and in here competency questions were designed by referring the information provided in [2],[6] and also from the official websites of University Grants Commission, Examination Department, Ministry of Higher Education.

A. Define the classes

A class definition can be done using different approaches. A top-down development process: In this approach, class definition is started by defining the most general concept in the domain. A bottom-up development process: Specific classes of the domain will start defining first in this approach. Apart from that, a combination development process, which is a combination of both top-down and bottom-up approaches, can also be used.

Thirteen classes; University, ALStream, Faculty, Subject, Department, Course Of Study, Higher Educational Institute, Score Restriction, Eligible Requirements, District, Student, Degree, and Career are identified as the top level classes in the ontology class hierarchy. Figure 1 visualizes the top-level classes of the ontology.

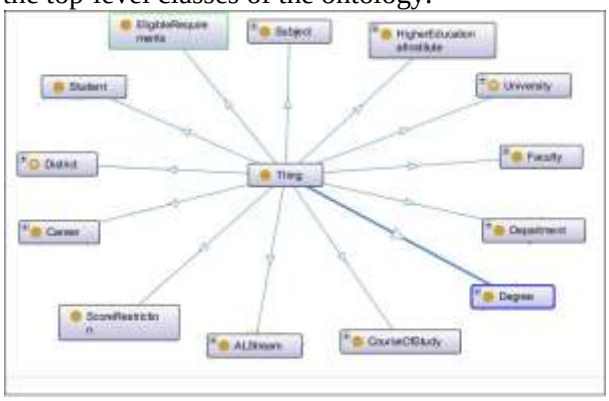

Figure 1: Top Level classes of Ontology

After identifying the top-level classes of the ontology, ontology hierarchy was developed by creating the subclasses. When defining the classes and subclasses, rules and restrictions, which needed to be embedded in the ontology, are also modeled. In the top-level class Subject, it has been identified that there are around 50 A/L subjects offered in Sri Lankan schools and they are categorized into 7 categories. In Sri Lanka, there are five major A/L streams available. The student will fall into one of the A/L Streams based on his subject selection. Not all three-subject combinations are valid. A/L streams are having restrictions in subjects, and student should select the subjects accordingly. The courses of Studies in

the Universities are categorized based on the A/L streams. However, there are some courses of studies that are not restricted only to one A/L stream.

Degree is another top-level class of the ontology. Degrees are grouped based on the type of the degree (eg: BSc, MBBS) offered by the university. There are restrictions about a valid course of studies to obtain a specific degree. For example to obtain a medical degree student must have followed biological science course of study. Career, Student, District, and ScoreRestrictions are other top-level classes. To enter into a certain career, a student must have done a specific degree. For example, if a student wanted to become a medical doctor then he must have followed medical degree. However, there are some career opportunities for students holding any type of degree.

Student is one of the important classes, which will visualize the current user of the system. In order to be eligible to apply for the university, admission student must fulfill some eligibility requirements mentioned by the UGC. Since the district of the student applying for the university take into the consideration districts in Sri Lanka and their minimum scores restrictions for the course of studies are modeled.

Higher Educational Institutes in Sri Lanka can be basically divided into two categories; Degree offering institutes and non-degree offering institutes. Universities are degree-offering institutes. Teacher training colleges, Technical Training colleges etc. categorized under non-degree offering institutes as they are not offering degrees but only diplomas. However, a student cannot enroll into any of the two higher educational institutes simultaneously, regardless of whether it is a degree offering or non-degree offering institute. Major academic division of a university is called Faculty and every university must have at least one faculty. A faculty is composed of several departments and in one faculty, there must be at least one department.

B. Define properties of classes

To construct the relationship between classes data and object property restrictions are added. Table 2 shows the object properties defined in the ontology. Properties link individuals from the domain to individuals from the range. Inverse property of the defined property and also some restrictions added into the properties are demonstrated in the figure.

Data properties are defined in order to visualize the relationship between individuals of the classes and the literal values. Table 1 shows the data properties used in the ontology.

Property restrictions are added to make sure the ontology is representing the correct existing knowledge. Following are some sample existential and cardinality restrictions.

Table1. Data Properties of the Ontology

Property	Domain	Range
hasALAttempts	EligibleRequirements	integer
hasAptitudeTest	Degree	boolean
hasDistrictRank	Student	positiveInteger
hasIslandRank	Student	positiveInteger
hasIntake	CourseOfStudy	positiveInteger
hasDuration	Degree	positiveInteger
hasMinScore	ScoreRestriction	string
hasUniversityRank	University	positiveInteger

- A University should have at least one Faculty
- Faculty should have at least one Department
- Department should offer at least one Degree
- A Degree should belong to at least one-degree stream.
- Medium of instruction of a degree should be at least one language and should have exactly one duration in years, which should be greater than 3.
- Advanced Level Stream is composed of exactly 3 Advanced Level subjects.

Table2. Object Properties of the Ontology

Property	Domain	Range
hasAdmissionRequirement	CourseOfStudy	ALStream , ScoreRestriction
hasALRequirement	CourseOfStudy	ALStream
hasScoreRequirement	CourseOfStudy	ScoreRestriction
hasDegree	Department	Degree
hasDistrict	Student	District
hasDivision	University	Faculty, Department
hasDepartment	Faculty	Department
hasFaculty	University	Faculty
hasFulfill	Student	EligibleRequirement
hasOLRequirement	Degraa	OLSubject
hasQualification	Career	Degree
hasSubject	ALStream	ALSubject
hasSubjectCategory	ALStream	ALSubject
offerDegree	University	Degree

C. *Define Instances*

In order to check the correctness of relationships and confirm the strength of ontology to planned Competency Questions, set of individuals under each class are defined. After creating, the instances reasoner were synchronized and tested to check whether it correctly performed according to the defined rules. All the rules were correctly reasoned. For example Chemistry individual were created under Science subject category and once the reasoner synchronized, it automatically identified that Chemistry is a subject of Physical Science, Biological Science Agri Science and Technology Science A/L streams. Moreover, it identified the degrees, which Chemistry is a subject requirement.

4. Experimental Results

Correctness of class hierarchies and the restrictions were tested by adding test classes to the ontology. All the tests were successful and inferred hierarchy correctly listed the classes. For example, to check the correctness of the A/L streams, few test classes were added and tested whether the reasoner inferred it correctly. Figure 2 shows that the reasoner correctly identified the Test class should fall in to Arts Stream (Highlighted in light orange color).

Figure 2: Inferred hierarchy test of A/L Stream

The ontology or the knowledge base can then be used to retrieve answers for several intelligent questions. Reasoners and SPARQL query engines can be used to query and grab the desired answers from the system. Among the question pool created by the questions sent by the A/L and O/L students, 50 questions were selected and tested. All the queries returned the expected answer. Some answers were not complete due to lack of data embedded to the ontology, but the answer was correct according to the available data. Following are some sample scenarios that demonstrate the knowledge retrieval from the developed knowledge base.

Scenario 1: Student wants to enroll into any course of study under Biological Science course of study. He wants to know the possible list of subjects he can follow.

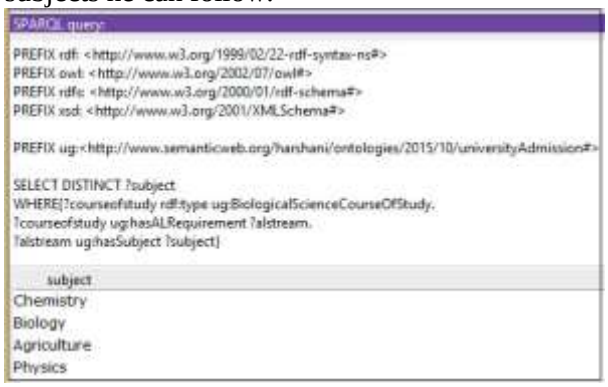

Figure 3: SPARQL Query and Output: Subject requirement for A/L streams

Scenario 2: Student wants to know the possible degrees he can follow in order to become a Software Engineer.

Figure 4: SPARQL Query and Output: Degree qualification for a career

Scenario 3: Student wants to know the degrees which are having Aptitude test as additional entry requirements.

Figure 5: SPARQL Query and Output: Degrees having Aptitude tests

Scenario 4: If a student has already done Chemistry, Physics and Combined Mathematics for his A/Ls what are the course of studies he can enroll?

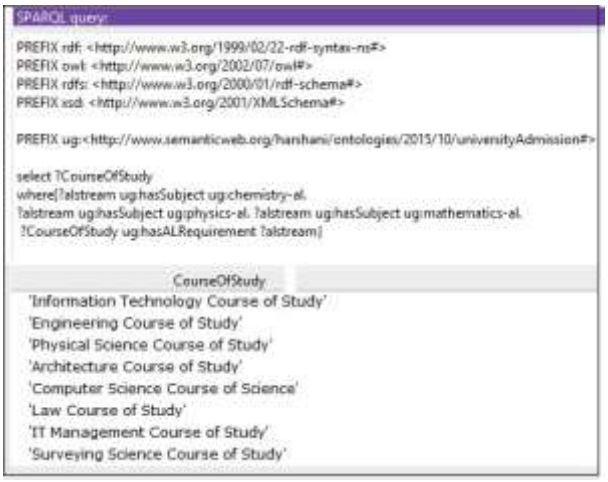

Figure 6: SPARQL Query and Output: Course of studies for A/L subjects

Scenario 5: Student wants to know what the possible career opportunities are if he obtains a particular Degree.

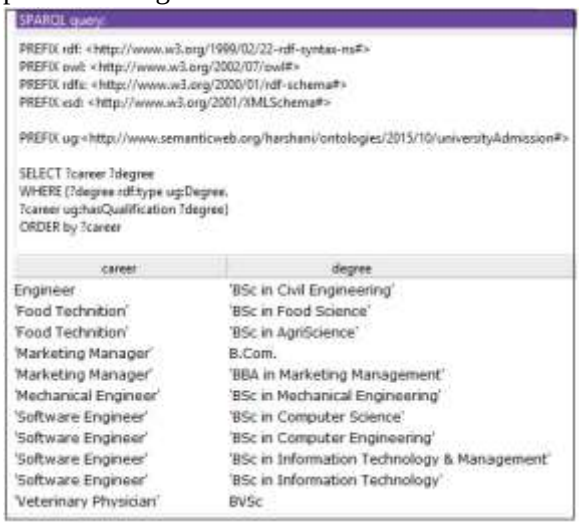

Figure 7: SPARQL Query and Output: Career opportunities for a Degree

5. Conclusion

University Admission process is a complex and time-consuming task which requires experts' advice in all the stages. All the required information are available but scattered in the web. There is no proper linking in this available information hence reduce the value of it. Not only the lack of interconnectedness but also this information is presented in a format which only human users can consume. Computer agents or applications cannot understand this information so that data is isolated.

Developing a recommendation system to guide the students in this complex process was the main objective of this research and Ontology was designed and developed as the knowledge base of the system. It was designed and implemented using the ontology editor Protégé. Correctness of the Ontology was tested by running several SPARQL queries. Around 50 queries were tested and all the queries returned the expected results and it confirmed that the ontology design is comprised with the real world.

References

[1] Noy, N. F., & Mcguinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology.

[2] M.Amilajith, N. L. (2012). Advance level, sri lankan high school curriculum and university admission. Hanyang University, Institute of Educational Technology.

[3] Tharmaseelan, N. (2012). Tertiary education in sri lanka: Issues and challenges. Bulgarian Journal of Science and Education Policy, 1(1), 173–190.

[4] Malik, S. K. (2010). Developing an university ontology in education domain using protégé for semantic web. International Journal of Science and Technology, 2(9), 4673–4681

[5] Ayesha ameen, P. R., Khalel khan. (2012). Construction of university ontology. World Congress on Information and Communication Technologies.

[6] Ghanashyam Rout, S. M., Sreedhar Madichetty, & Pattnaik, S. (2012). An efficient wavelet based automation process for e-admission in universities of odisha. International Journal of Computer Applications, 47(22).

[7] Gruber, T. (1993). Towards principles for the design of ontologies used for knowledge sharing. International journal human-computer studies, 43, 907–928.

[8] Kapoor, B., & Sharma, S. (2010). A comparative study ontology building tools for semantic web applications. International journal of Web Semantic Technology (IJWesT), 1(3).

[9] K. OHara, C. B. (2007). Knowledge representation with ontologies: present challenges future possibilities. International journal human-computer studie, 65, 907–928.

[10] Naveen Malviya, S. S., Nishchol Mishra. (2011). Developing university ontology using protégé owl tool: Process and reasoning. International Journal of Scientific Engineering Research, 2(9).

[11] Zeng, L. (2009). Study on construction of university course ontology: Content, method and process. International Journal of Science and Technology, 2(9), 4673–4681.

[12] Gruber, T. (2009). Ontology. Encyclopedia of Database Systems.

[13] T. Segaran, J. T., C. Evans. (n.d.). Programming the semantic web. International journal human computer studies.