	VIETNAM NATIONAL UNIVERSITY OF HO CHI MINH CITY UNIVERSITY OF NATURAL SCIENCES **FACULTY OF INFORMATION TECHNOLOGY**

COURSE SYLLABUS

Course Code: **TH137**

Title: Natural Language Processing

Credits: **3**

Workload:
Lecture hours: **2** periods * **15** weeks = **30** periods
Laboratory hours: **2** periods * **15** weeks = **30** periods
Preparative hours: **2** periods * **15** weeks = **30** periods

Prerequisites: Introduction to Artificial Intelligence
Machine Learning
Knowledge-based systems

Summary: Natural Language Processing (NLP) is the branch of Artificial Intelligence applications specializing in parsing, recognizing and generating the Natural Languages. It is the basic background for entering applications e.g. : NL Understanding, Machine Translation, Voice Processing, Text summarization, Information Retrieval,...

Course Outline

Chapter 1: Overview of Natural Languages

1.1. Introduction to natural languages

1.2. Lexicology

1.3. Syntactic

1.4. Semantics

Chapter 2: Overview of Formal Language

2.1. Language descriptions

2.2. Grammar of language

2.3. Grammar rule

2.4. Derivation tree

Chapter 3: Some problems in Natural Language Processing

3.1. Preprocessing

3.2. Morphology analysis

3.3. Syntactic analysis

3.4. Semantics analysis

3.5. Other problems

Chapter 4: Machine Readable Dictionary

4.1. Machine readable dictionary

4.2. Macro and micro structures of dictionary

4.3. Data structure of dictionary

4.4. Automatically constructed dictionary

4.5. Exploiting resources in dictionary

Chapter 5: Applications in Natural Language Processing

5.1. Machine translation

5.2. Text processing

5.3. Character recognition

5.4. Speech processing

5.5. Other applications

Grading	Final exam:	60%
	Assignments:	40%

REFERENCE

Vietnamese

[1] Mai Ngọc Chừ - Vũ Đức Nghiệu - Hoàng Trọng Phiến (1991), *Cơ sở ngôn ngữ học và tiếng Việt*, NXB Đại học và GD CN, Hà Nội.

[2] Đinh Điền (2005), *Giáo trình xử lý ngôn ngữ tự nhiên*, NXB ĐHQG TPHCM.

English

[3] James Allen (1995), *Natural Language Understanding*, Academic Press.

[4] Gilbert Krulee (1991), *Computer Processing of Natural Language*, Macmillan Press.

[5] Wilks Y., Slator B., Guthrie L. (1996), *Electric Words: Dictionaries, Computers and Meanings*, MIT Press, Cambridge, MA.

[6] Florian R., Ngai G. (2001), "Fast Transformation-Based Learning Toolkit", *Johns Hopkins University Technical Report,* version 1.0, 9/2001.