CMPE 510 MACHINE TRANSLATION
SYLLABUS

Instructor: Tunga Güngör (E-mail: gungort@boun.edu.tr, Room: ETA 34)

Course Description:

This course covers Machine Translation (MT), i.e. the use of computers to translate (or help humans to translate) between natural languages. First, a brief background, definitions related to the field, and history of MT will be given. This will be followed by the study of the rule-based approaches, a practical MT scenario, and the explanation of an example translation. Then, the translation approach that dominated the field, which is statistical MT, will be covered in detail. Another data-driven approach, which is example-based MT, will be studied. This will be followed by the most recent approach, neural MT. The final topic is the evaluation of automatic translations. There will be paper presentations throughout the semester.

Prerequisites:

- Background in Natural Language Processing

Text Book:

- Statistical Machine Translation, P. Koehn, Cambridge Unv. Press, 2010

Supplementary Books:

- An Introduction to Machine Translation, W.J.Hutchins, H.L.Somers, Academic Press, 1992
- Translation Engines: Techniques for Machine Translation, A.Trujillo, Springer Verlag, 1999
- Machine Translation: An Introductory Guide, D.Arnold, L.Balkan, S.Meijer, R.L.Humphreys, L.Sadler, Blackwell Pub., 2001

Reference Books:

- Hybrid Approaches to Machine Translation, M.R.Costa-jussa, et al. (eds.), Springer, 2016
- Machine Translation, Pushpak Bhattacharyya, CRC Press, 2015
- Handbook of Natural Language Processing and Machine Translation, J.Olive, C.Christianson, J.McCary (eds.), Springer, 2011
- Improving Reordering and Modeling in Statistical Machine Translation, Marta Ruiz Costa-jussa, VDM Pub., 2010
- New Statistical and Syntactic Models for Machine Translation, M.Khalilov, J.A.R.Fonollosa, VDM Pub., 2010
- Machine Translation: Its Scope and Limits, Yorick Wilks, Springer, 2009
- Learning Machine Translation, C.Goutte, N.Cancedda, M.Dymetman, G.Foster (eds.), MIT, 2009
- Language Modeling For Machine Translation, M.Raab, VDM Verlag, 2007
- Recent Advances in Example-Based Machine Translation, M.Carl, A.Way (eds.), Kluwer Academic Pub., 2003
- Readings in Machine Translation, S.Nirenburg, H.Somers, Y.Wilks (eds.), MIT, 2003

Lecture Hours and Rooms:

Tuesday 14:00-17:00 BM A3

Course Schedule:

Overview, History of MT, Basic concepts
(*Arnold, Chp.1; Koehn, Chp.1)*
Basic definitions, Rule-based approaches (RBMT), An MT scenario
(Hutchins, Chp.1.2; Hutchins, Chp.4.1-4.3; Arnold, Chp.2)
Example translations (Direct, Transfer, Interlingua)
(Hutchins, Chp.6.4-6.7)
Paper presentations (RBMT and related topics)
Statistical MT
Word-based translation
(Koehn, Chp.4)
IBM models
(Koehn, Chp.4)
Phrase-based translation
(Koehn, Chp.5)
Decoding
(Koehn, Chp.6)
Language models
(Koehn, Chp.7)
Paper presentations (SMT and related topics)

Example-based MT
(Trujillo, Chp.8.1)
Neural MT
(Koehn, draft Chp.13)
Paper presentations (NMT and related topics)
MT evaluation and quality
(Trujillo, Chp.10; Arnold, Chp.9; Koehn, Chp.8)
Research project presentation

Evaluation (subject to change):

Midterm	: % 15
Presentations	: % 20
Research Project	: % 20
Application Project	: % 20
Final	: % 25

Notes:

- Midterm will be held on 31.03.2020 Tuesday during lecture hours.
- The midterm and final exams will be "closed books and notes".
- Attendance for the exams, submitting the projects, and attendance for the presentations are obligatory. Otherwise, you will fail the course, regardless of the grades obtained in other parts of the course.
- You can follow the announcements via the university's Moodle system (https://moodle.boun.edu.tr) (Students registered to the course were automatically subscribed to the course. If not, send me an e-mail using your "boun.edu.tr account".)
- You can obtain some of the course materials from the instructor.
- Please read the section "graduate courses" in the web page General Information for Students. This page explains the course policy, the grading system, and information about the assignments and projects.