

LING 120:
Language and Computers
Semester: Fall '17

Instructor: Sowmya Vajjala

Iowa State University, USA

13 November 2017

Outline

- ▶ General Announcements
- ▶ Discussion about previous exercise
- ▶ Plan for this week:
 - ▶ Machine Translation: introduction (today and wednesday)
 - ▶ Friday: General discussion about different applications, impact of these on our life etc.
 - ▶ How does Machine Translation work? (after the break)
- ▶ Exercise about Machine Translation

General Announcements

General Stuff

- ▶ **Reminder:** Assignment 6 due this week!
- ▶ Revision topics: no one posted anything
- ▶ But I was reading about representing Sign Language and Braille on computer over the weekend - may be I will try to know more and talk about this after the break.

Grade improvement: oral exam timings

People wanting grade improvement, choose one of the following timings and get back:

- ▶ 4th December: 1-2 pm, 5-5:30 pm. 6th December: 1-2pm, 5-5:30 pm. 8th December: 1-2 pm.
- ▶ Choose a time (actual oral exam will only take 15 min), choose any 3 topics from the course, and send me an email.
- ▶ Show up at the time you promised to come, and I will pick one (or two) of the three topics and ask questions.
- ▶ Maximum improvement: 5% (I am flexible if you do extraordinarily well)

Final Exam: Recap of description

- ▶ Carries: 20% of your grade, and involves writing 2 short essays.
- ▶ Has to be submitted in three parts:
 1. Write a part of the assignment (one question) and submit by 2nd December - 5%
 2. Do an in-class peer review of one of your classmate's work (6th December) - 5% - if you don't come to class, you don't get graded for this part.
 3. Final submission (of both the questions) - 10%
- ▶ Exact details about word limits, how to submit and list of topics for questions are on Canvas.

Final Exam- Questions?

- ▶ I am hoping you read through the requirements and expectations
- ▶ I am also hoping you now know the deadlines
- ▶ Any questions on finals?

Discussion of Last Class Exercise

Attendance Exercise From Last Class

Running on MT -

<http://nacloweb.org/resources/problems/2011/A.pdf>

Solution:

- ▶ Looking for words that seem out of place
- ▶ Thinking about synonyms of those words that fit in the context.
- ▶ Full list: http://nacloweb.org/resources/problems/2011/naclo11r1_sol.pdf

Nellie on Speech-Speech Translation in Skype

Intro to Machine Translation

Questions, questions, questions!

- ▶ Is translation easy for humans?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?
 - ▶ What is possibly difficult to translate?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?
 - ▶ What is possibly difficult to translate?
 - ▶ What do you think is impossible to translate?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?
 - ▶ What is possibly difficult to translate?
 - ▶ What do you think is impossible to translate?
- ▶ What do you think is the case for machines?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?
 - ▶ What is possibly difficult to translate?
 - ▶ What do you think is impossible to translate?
- ▶ What do you think is the case for machines?
- ▶ What are the applications of machine translation?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?
 - ▶ What is possibly difficult to translate?
 - ▶ What do you think is impossible to translate?
- ▶ What do you think is the case for machines?
- ▶ What are the applications of machine translation?
- ▶ How do we evaluate human translation?

Questions, questions, questions!

- ▶ Is translation easy for humans?
- ▶ What in your idea should machine translation do?
- ▶ For humans:
 - ▶ What is possibly easy to translate?
 - ▶ What is possibly difficult to translate?
 - ▶ What do you think is impossible to translate?
- ▶ What do you think is the case for machines?
- ▶ What are the applications of machine translation?
- ▶ How do we evaluate human translation?
- ▶ How do we evaluate machine translation?

Let us talk about Human Translation first

Can 2–3 people who can speak/read/write in another language come forward and translate the following sentence:


Why do we need machine translation?

Machine Translation today - Example 1


I want to use this example to illustrate the issues in translating from English to Telugu automatically.

ఇంగ్లీష్ నుండి తెలుగు భాషకు అనువాదములోని సమస్యలను ఉదాహరించడానికి నేను ఈ ఉదాహరణను ఉపయోగించాలనుకుంటున్నాను.


Machine Translation today- Example 2




Translate

Turn off instant translation 

Hindi Finnish **Telugu** Detect language ▾





 English Greek Telugu ▾ **Translate**


నా చెల్లెలు మీ ఇంటికి వచ్చిందా?
నా అక్క మీ ఇంటికి వచ్చిందా?

 ▾

61/5000

Did my sister come to your house?
Did my sister come to your house?


   

 Suggest an edit

Nā celledu mi iṅṭiki vaccindā?

Nā akka mi iṅṭiki vaccindā?

Machine Translation today- Example 3



Translate

Turn off instant translation

Hindi Finnish Telugu Detect language

English Greek Telugu

Translate


చెప్పాలనిపించడంలేదు
నాకు నిజం చెప్పాలనిపించడంలేదు
నాకు నిజం చెప్పాలని అనిపించడం లేదు

95/5000

×

Ceppalanipincadanledu
I do not want to tell you the truth
I do not want to tell you the truth

☆ 📄 🔊 ↻

 Suggest an edit

Ceppālanipinḥaḍanlēdu
nāku nijarḥ ceppālanipinḥaḍanlēdu
nāku nilku nijarḥ ceppālani anipinḥaḍarḥ lēdu

So, what are the issues?

- ▶ From last class, we also know: Each word can have multiple meanings, so mapping words one-one between languages is not possible.
- ▶ From today's examples, we know there are several other issues such as:
 - ▶ How words are formed between languages
 - ▶ What is the order of words
 - ▶ Words that don't seem to translate
 - ▶ Words that seem to come up in translation that do not exist in original
 - ▶ Issues such as: articles don't exist in Telugu, but they do in English. Kinship words are more fine grained in Telugu than English etc.

Other Issues

- ▶ Stylistic and cultural differences - which can make even manual translation difficult
- ▶ Metaphor is even difficult - same metaphors need not exist in both languages
- ▶ Poetry with rhyme - what if those words don't rhyme in Target language?

Different Scenarios for Machine Translation

- ▶ Tasks where rough translation is adequate (e.g., doing a web search)
- ▶ Tasks where a human post-editor is used (e.g., computer aided human translation, to reduce effort for human translators, in scenarios like creating multi-lingual interfaces to Windows, Facebook etc)
- ▶ Tasks within a restricted domain, which need high quality translation (e.g., Weather forecasting)

Today's Attendance Exercise

In this course so far, where is the idea of n-grams useful among the topics we discussed? Do you think they are useful in Machine Translation too?

Just an FYI: our main topics were - text and speech input, spelling and grammar correction, tutoring systems, search, text classification, dialog systems, speech recognition and synthesis.