

Fall Semester 2017  
Iowa State University

## LING 120 - Language and Computers

### Course Handbook

**Instructor:** Dr. Sowmya Vajjala

- *Office:* 331 Ross Hall
- *Email:* sowmya@iastate.edu

**Course Objectives:** Computers have become a preferred medium to read and write texts in several walks of daily life. Some examples of our interactions with computers involving text include: reading news, posting on social media, reading/writing email, google search etc., So, modern computers spend a lot of time working with human language(s) and we spend a lot of working with computing machines. Considering this background, this course offers an overview of various real-world challenges and applications that involve computers working with human language. The primary objective of this course is to provide you an overview of how language technology works by focusing on a few applications and giving the necessary conceptual background to understand these technologies.

**Pre-requisites:** No pre-requisites other than some curiosity about human language and how computers see it.

### Course Details:

- *Class timings:* Monday, Wednesday, Friday - 2:10-3 pm
- *Classroom:* E HALL 0111 on Mondays and Fridays. ROSS H 0137 on Wednesdays.
- *Office hours:* Monday and Wednesday, 1-2 pm.

### Credits:

- Credit Points: 3

**Nature of the course and expectations:** This is a 3 credit, undergraduate course. Primary mode of instruction is by lectures followed by discussion. We will have regular assignments, a mid-term and a final exam. Readings for each topic are specified in the syllabus and it is expected that the students read them **before** coming to the class. There may be a few additional (mostly optional) readings or videos from other sources for some of the topics. Students enrolled in the course are expected to

1. regularly and actively participate in class, and submit the assignments on time (60% of the grade)
2. attend a mid-term and a final exam (40% of the grade)

**Grading Policy:** There are 6 assignments with 10 marks each, one mid-term for 20 marks and one final exam for 20 marks. The assignments will come about once in 2 weeks you usually will have 2 weeks of time for submission. For the mid-term, you have to give a classroom presentation on a topic related to language technologies as a team of 3-4 people. The topics should be decided by the end of September from a list of given topics. You are allowed to come up with your own idea, but should get my approval before starting to work on that. The final exam is a take-home exam which involves writing a detailed report (optionally: writing a small piece of software instead of a report) on a topic related to language and computers. Topics can be either chosen from a list of announced topics or you can choose your own. Detailed instructions will be provided in due course of time. Plus/minus grading will be used (93% = A, 90% = A-, 87% = B+, 83% = B, 80% = B-, etc.).

**Attendance Policy:** You require a minimum of 80% attendance to pass the course with full credits. If you drop below that, your grade will reduce by one grade point for each 5% reduction (i.e., If you got A before taking attendance into account, 75-80% attendance will give you a A-, 70-75% attendance will give you a B+, and so on). I don't take attendance by roll-call, but will give an exercise in the class for each class day. Your completion of these counts as attendance for that day.

**Class etiquette:** Please do not read or work on materials for other classes in this class. Come to class on time and do not pack up early. Electronic devices like mobile phones, tablets etc should not be used in the class. Laptops should be used only for activities related to the classwork. If for some reason, you must leave early or you have an important call coming in, or you have to miss class for an important reason, please let me know (via email) and get it approved *before* the class. Being absent from the class does not allow you to skip submitting any assignments that were assigned in that class. Do not ask questions that can easily be answered by looking at the textbook or in online discussion forums. Value my time, your time, and everyone else's time. Asking such questions will not result in any answers from me **in the class** even if you write negative feedback about it.

**Academic Conduct:** Generally, you are encouraged to work in groups, discuss, and exchange ideas. At the same time, you are expected to do your assignments by yourself and with honesty. For the text you write, you always have to provide explicit references for any ideas or passages you reuse from somewhere else. Note that this includes text taken from the web. You should cite the url of the web site in case no more official publication is available. It is common to search in websites such as stackoverflow.com for solutions to some problems or to fix the bugs in your program. However, copying full code samples from somewhere else (including your colleague's program) is considered academic dishonesty. Generally speaking, the class will follow Iowa State University's policy on academic dishonesty. Anyone suspected of academic dishonesty will be reported to the Dean of Students Office.

**Disability Accommodation:** Iowa State University complies with the Americans with Disabilities Act and Sect 504 of the Rehabilitation Act. If you have a disability and anticipate needing accommodations in this course, please contact (instructor name) to set

up a meeting within the first two weeks of the semester or as soon as you become aware of your need. Before meeting with (instructor name), you will need to obtain a SAAR form with recommendations for accommodations from the Student Disability Resources, located in Room 1076 on the main floor of the Student Services Building. Their telephone number is 515-294-7220 or email [disabilityresources@iastate.edu](mailto:disabilityresources@iastate.edu) . Retroactive requests for accommodations will not be honored.

**Harassment and Discrimination:** Iowa State University strives to maintain our campus as a place of work and study for faculty, staff, and students that is free of all forms of prohibited discrimination and harassment based upon race, ethnicity, sex (including sexual assault), pregnancy, color, religion, national origin, physical or mental disability, age, marital status, sexual orientation, gender identity, genetic information, or status as a U.S. veteran. Any student who has concerns about such behavior should contact his/her instructor, Student Assistance ([dso-sas@iastate.edu](mailto:dso-sas@iastate.edu)), or the Office of Equal Opportunity and Compliance at 515-294-7612.

**Dead Week Policy:** This class follows the Iowa State University Dead Week policy as noted in section 10.6.4 of the Faculty Handbook: <http://www.provost.iastate.edu/resources/faculty-handbook>

## Textbook

- Language and Computers (Dickinson, Brew and Meurers)

(note: I will occasionally ask you to read other stuff or work with web-based tools)

## Syllabus - topics covered

1. Introduction (course overview, impact of language technology etc)
2. Encoding human language on computers (topic for Assignment 1)
3. Writers aids on computers (topic for Assignment 2)
4. Intelligent Tutoring Systems
5. Search and Information Retrieval (topics for Assignment 3)
6. Natural Language Processing - an overview of challenges, and applications (topic for Assignment 4)
7. Text Classification (topic for Assignment 5)
8. Interactive Dialogue Systems
9. Automatic Speech Recognition (topics for Assignment 6)
10. Machine Translation

**Important Deadlines** (note: all deadlines end at 11:59 pm on that day, and are typically on Saturdays for assignments and Wednesdays for other exams)

1. Assignment 1: 8 September 2017
2. Assignment 2: 23 September 2017
3. Assignment 3: 7 October 2017
4. Mid-term: 11th October 2017 (midterms presentations on Monday and Wednesday that week)
5. Assignment 4: 21st October 2017
6. Assignment 5: 4th November 2017
7. Assignment 6: 18th November 2017
8. Final exam: 13th December 2017

**Detailed Scheduling and Deadlines (tentative)** Note that the following session plan is subject to change; it only constitutes the current state of our planning as the semester unfolds.

1. **Week 1:** Monday, August 21: Course orientation
2. Wednesday, August 23: What is the big deal about computers working with human languages?
3. Friday, August 25: Encoding written and spoken language  
(Assignment 1 assigned. Deadline: 9th September, 11:59pm )
4. **Week 2:** Monday, August 28: Conclusion of the topic: Encoding human language
5. Wednesday, August 30: Writers aids - spelling and grammar checking tools - an introduction
6. Friday, September 1: Spelling correctors
7. **Week 3:** Monday, September 4: University Holiday, no classes (Yay!)
8. Wednesday, September 6: Grammar correctors
9. Friday, September 8: Conclusion of the topic writers aids.  
(Assignment 2 assigned. Deadline: 23rd September)
10. **Week 4:** Monday, September 11: Language tutoring systems - introduction
11. Wednesday, September 13: Role of language in tutoring systems
12. Friday, September 15: Conclusion of the topic tutoring systems
13. **Week 5:** Monday, September 18: Search: introduction  
(Midterm presentation topics are put up. Teams are assigned.).
14. Wednesday, September 20: Regular expressions
15. Friday, September 22: Using regular expressions to search for patterns. (Assignment 3 assigned. Deadline: 7th October)

16. **Week 6:** Monday, September 25: Search: conclusion
17. Wednesday, September 27:
18. Friday, September 29:
19. **Week 7:** Monday, October 2: Common Tasks in language processing (overview)
20. Wednesday, October 4:
21. Friday, October 6: The idea of "machine learning" for language processing (Assignment 4 assigned. Deadline: 21st October 2017)
22. **Week 8:** Monday, October 9: Mid-term presentations
23. Wednesday, October 11: Mid-term presentations
24. Friday, October 13: Review so far. Feedback on mid-term presentations. General discussion.
25. **Week 9:** Monday, October 16: Cryptography and Language - overview
26. Wednesday, October 18: Text classification - overview
27. Friday, October 20: Spam filtering - an example of text classification (Assignment 5 assigned. Deadline: 4th November)
28. **Week 10:** Monday, October 23: algorithms for classification.
29. Wednesday, October 25:
30. Friday, October 27: Dialogue Systems - Introduction
31. **Week 11:** Monday, October 30: Designing a dialog system
32. Wednesday, November 1: Evaluating a dialogue system
33. Friday, November 3: Overview of Automatic Speech Recognition (Assignment 6 assigned. Deadline: 18th November 2017)
34. **Week 12:** Monday, November 6: Speech Synthesis
35. Wednesday, November 8: ASR in real world
36. Friday, November 10: Dialog systems and ASR
37. **Week 13:** Monday, November 13: Some applications of language technologies: Discussion (
38. Wednesday, November 15: Continuation.
39. Friday, November 17: Introduction to Machine Translation & Transliteration (Final exam topics are put up).
40. **Week 14:** (November 20 –24) Thanksgiving break, no classes!
41. **Week 15:** Monday, November 27: Machine translation
42. Wednesday, November 29: Machine translation
43. Friday, December 1: Recent trends in language technologies. Impact of language technology
44. **Week 16:** (last week of classes) Monday, December 4: Recap, exercises in class
45. Wednesday, December 6: Recap, classroom projects
46. Friday, December 8, classroom projects

47. **Week 17:** (December 11-15) Final exams week. **Final exam submission date:**  
**13th December, midnight.**