

**CSCI 699: Grounding Natural Language**

**Units: 4**

**Spring 2022 Tuesdays 2:00-5:20pm**

**Instructor: Jesse Thomason**

**Office:** SAL 244

**Office Hours:** by appointment

**Contact Info:** jessetho@usc.edu; timeline for replying to emails: within 48 hours.

|  |
| --- |
|  |

**Course Description**

Advances in representation learning for natural language via lexical embedding techniques, followed by large-scale language modeling via transformers to produce contextual embeddings, have dominated mainstream NLP for years. Yet, large language models and associated, deep contextual language representations remain unable to understand entailments such as “she took three lefts” implies “she went right.” Language is *grounded in experience* with the world and with other speakers, which no scale of text alone can ever convey to a model. This seminar class will explore the ways in which other sensory modalities, especially visual input and embodiment in 3-dimensional space, can influence and guide representation learning for language. We will focus on techniques to jointly learn representations for text and visual modalities (e.g., self-supervised masked-language and masked-region modelling), to translate between single-modality spaces (e.g., projection and manifold alignment), and to use language and vision signals to produce actions in simulated worlds (e.g., vision-and-language navigation and language-guided task execution). Finally, we will explore the use of dialogue for human-in-the-loop feedback on tasks and resulting language representations.

**Learning Objectives**

Students will interrogate the question of whether and when language is *enough* for tasks that involve language, and the nature of machine “understanding.” Students will achieve an understanding of modern language grounding techniques, their context and historical precedent, and open problems in the field. Students will review and present in detail milestone papers and techniques for class discussion. The hope is that at the end of the course, students will be able to identify shortcomings in single-modality techniques in language and beyond, and be able to develop practical solutions and insights for the application of grounded language

**Recommended Preparation**: CSCI 544 (Applied Natural Language Processing) or other graduate-level courses in computer vision or machine learning. General background on mathematical notation and optimization; familiarity with programming in Python (pytorch experience can’t hurt!), probability, and algorithms.

**Course Notes**

Lecture notes will be available online after each class.

**Technological Proficiency and Hardware/Software Required**

Ability to read modern research papers in natural language processing, computer vision, and/or robotics required; this skill will be further developed through the course. For class projects, access to lab compute (i.e., GPUs) is strongly recommended.

**Required Readings and Supplementary Materials**

All reading material will be posted on the course web page at the beginning of the course. All reading material will be freely and publicly available online.

**Description and Assessment of Assignments**

Each student will present one or two papers from the field of language grounding and complete a research project. There will be no midterm or final examinations. The main assessment will be engagement in class discussions about required paper readings and effort in the development and execution of the research project.

**Paper presentation(s) and discussions.** Students will read a paper and prepare slides to present its core motivations, insights, and conclusions to the class for discussion. The student will be expected to facilitate discussion on the paper, including not just answering class questions but encouraging them through an interrogation of the paper’s methods and conclusions. The grade for this assignment will be 5% for the student’s own presentation and discussion leadership, and 10% for the student’s engagement in discussion of other students’ presentations.

**Project proposal slides.** Students will brainstorm a project idea *individually* in the space of language grounding. The instructor may provide a set of “themes” in the course schedule from which students can draw (e.g., “modeling or analysis of the ALFRED benchmark”), but projects need not adhere to one of these themes. Early in the course, students will prepare 5-6 slides articulating the objective of their project, how the problem is approached currently, how the proposed approach addresses the limitations of existing work, and what success would contribute to the field, including considerations of the ethical implications of the contribution’s deployment.

**Project survey report.** A literature survey, analogous to the Related Work section of a conference paper, will be prepared by each student. This survey will cover the same core topics as the proposal slides, digging deeper into literature to understand and contextualizing existing work, what challenges remain to be solved, insights drawn from other work and how they might apply to a new method, and ethical implications of success. This stage may require refining the proposal to differentiate it further from existing work, if the student finds that their novel insight or new idea has been previously explored. The survey will be formatted as a 2 page literature review write-up. A template will be provided.

**Project mid-term presentation.** The student will deliver a 5 minute presentation with accompanying slides to overview their proposed approach, early stage results so far, analysis of any interesting findings, and a concrete set of next steps achievable before the end of the semester.

**Project mid-term report.** The student will first run one or more state-of-the-art and baseline methods for their proposed problem and analyze the performance of these methods. Error analysis of existing methods will provide quantitative insight into what the proposed method can and should address in achieving better results. Additionally, writing analysis and comparison code now will enable faster and more effective benchmarking of the new method against baseline and current approaches. This report will be about 3 pages. A template will be provided.

**Project final presentation.** The student will deliver a 10 minute presentation with accompanying slides, analogous to a conference talk, with 5 minutes for questions after. The presentation will briefly introduce the problem and the current approaches to that problem, then the insights and implementation of the proposed new approach, followed by experimental results and analysis of the proposed method’s performance.

**Project final report.** The student will present their findings, either a new method for an existing problem or a new analysis, technique, or benchmark for a novel problem. The writeup will be 8 pages in total, with up to 5 pages recycled from the survey report (Introduction and Related Work) and mid-term report (Problem/Task and Baselines). Possible additional content would be the analysis of model performance from the mid-term report can be expanded to include the new method, a new or interesting problem formulation can be evaluated against existing models, and discussions about what interesting questions the research raises. Students will return to the ethical implications of their methods, whether successful or not; considering questions such as “what would success in this domain mean for some populations or applications?” and “how could systemic failure of my model on some parts of this data affect downstream applications and users?” A template will be provided for this report.

**Grading Breakdown**

|  |  |  |
| --- | --- | --- |
| **Assessment Tool (assignments)** | **Points** | **% of Grade** |
| Paper presentations & discussions | 100 | 15% |
| Project survey report | 100 | 15% |
| Project mid-term report & pres | 100 | 20% |
| Project final presentation | 100 | 20% |
| Project final report | 100 | 30% |

**Assignment Submission Policy**

Upload/email to be decided, encouraged by 6pm (aka “quittin’ time”) on the due date; accepted until 11:59pm on the due date.

**Grading Timeline**

Assignments will be graded within one week of the due date.

**Additional Policies**

Late or rescheduled presentations and late reports will be discussed on a case-by-case basis. Submitting work late will primarily cause more stress in the generation of the final presentation and report. The due dates for the survey and mid-term reports are check-ins for student benefit.

**Course Schedule: A Weekly Breakdown**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Topics/Daily Activities** | **Readings/Preparation** | **Project Deliverables** |
| Week 1 Jan 11 | Class introduction, paper candidate discussion & presentation scheduling | **Project idea pitches; prepare a short pitch of project ideas (1-2 slides).** |  |
| Week 2 Jan 18 | Basics of deep learning, self-supervised learning, and optimization |  |  |
| Week 3 Jan 25 | Attention, transformers, and sequence prediction |  | **Proposals due Jan 27** |
| Week 4 Feb 1 | Unimodal representations in vision and language I |  |  |
| Week 5 Feb 8 | Unimodal representations in vision and language II |  | **Surveys due Feb 10** |
| Week 6 Feb 15 | Multimodal data and representations I |  |  |
| Week 7 Feb 22 | Multimodal data and representations II |  |  |
| Week 8 Feb 29 | Action and embodiment I: Action for “active” observation |  |  |
| Week 9 Mar 8 | Action and embodiment II: Action for world changes; reinforcement learning |  |  |
| Week 10 Mar 15 | *No class; Spring Recess* |  |  |
| Week 11 Mar 22 | **Project mid-term presentations** |  | **Mid-term reports due Mar 24** |
| Week 12 Mar 29 | Social grounding; dialogue as action. |  |  |
| Week 13 Apr 5 | Physical grounding; sensory modalities; robotics |  |  |
| Week 14 Apr 12 | Novel research directions in language grounding I |  |  |
| Week 15 Apr 19 | Novel research directions in language grounding II |  |  |
| Week 16 Apr 26 | **Project final presentations** |  |  |
| FINAL | **Final Report** |  | **Due on university-scheduled date of the final exam** |

**Statement on Academic Conduct and Support Systems**

**Academic Conduct:**

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” [policy.usc.edu/scampus-part-b](https://policy.usc.edu/scampus-part-b/). Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, [policy.usc.edu/scientific-misconduct](http://policy.usc.edu/scientific-misconduct).

**Support Systems:**

*Counseling and Mental Health - (213) 740-9355 – 24/7 on call*

[studenthealth.usc.edu/counseling](https://studenthealth.usc.edu/counseling/)

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

*National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call*

[suicidepreventionlifeline.org](http://www.suicidepreventionlifeline.org/)

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

*Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL), press “0” after hours – 24/7 on call*

[studenthealth.usc.edu/sexual-assault](https://studenthealth.usc.edu/sexual-assault/)

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

*Office of Equity and Diversity (OED) - (213) 740-5086 | Title IX – (213) 821-8298*

[equity.usc.edu](https://equity.usc.edu/), [titleix.usc.edu](http://titleix.usc.edu)

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

*Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298*

[usc-advocate.symplicity.com/care\_report](https://usc-advocate.symplicity.com/care_report/)

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity |Title IX for appropriate investigation, supportive measures, and response.

*The Office of Disability Services and Programs - (213) 740-0776*

[dsp.usc.edu](http://dsp.usc.edu/)

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

*USC Campus Support and Intervention - (213) 821-4710*

[campussupport.usc.edu](https://campussupport.usc.edu/)

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

*Diversity at USC - (213) 740-2101*

[diversity.usc.edu](https://diversity.usc.edu/)

Information on events, programs and training, the Provost’s Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

*USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call*

[dps.usc.edu](http://dps.usc.edu/), [emergency.usc.edu](http://emergency.usc.edu/)

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

*USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call*

[dps.usc.edu](http://dps.usc.edu/)

Non-emergency assistance or information.