# Application

## Project Overview

Provide an explanation/background of your UROP project that includes with whom you are conducting research.

Intensional semantic theories in linguistics and philosophy have been able to model many aspects of human reasoning. However, counterfactual conditional statements— those of the form, for example, “If Nixon were to press the button, there would be a nuclear holocaust”— have proven to be difficult to adequately model with purely intensional resources. This is in part due to the non-monotonic nature of counterfactual reasoning where, for instance, Strengthening the Antecedent A ⊃ C ⊢ (A ∧ B) ⊃ C does not hold. Additionally, Simplification of Disjunctive Antecedents (A v B) ⊃ C ⊢ A ⊃ C does hold. However, these principles are interderivable in an intensional theory, motivating a hyperintensional theory as an alternative. To this point, Dr. Brast-McKie, a post-doc in the Linguistics and Philosophy department at MIT, has developed a hyperintensional semantic theory for counterfactual conditionals. In place of the possible worlds deployed in standard intensional theories, the semantics is built over a space of states which are ordered by parthood. Although the semantics captures the desired patterns of reasoning without encoding any unwanted inferences, the resulting theory is considerably more complex, making it difficult to evaluate inferences for validity. Since searching for countermodels by hand is time-consuming and extremely limited in the extent of the language that can be surveyed, I will be working with Dr. Brast-McKie to build a program that searches for countermodels up to some limit in the complexity of the models under consideration. As a result, this program will be able to rule out the existence of small counterexamples to prospective inferences, providing evidence that those inferences are valid. Additionally, when small countermodels exist, the program will specify those models. Searching the space for all inferences that do not have small countermodels will provide a much greater sense of the logic of counterfactuals that what could be established by hand. Moreover, the logic and semantics for counterfactuals is important step towards providing corresponding theories for causal inferences, where both theories have important applications in AI research. In particular, given a determination made by a neural network, it is often important to ask what caused the neural network to make the determination that it did, as well as asking what determination it would have made if things had been slightly different.

## Personal Responsibilities & Goals

Describe your planned role in the project. Be as specific as you can about your personal research duties/responsibilities, expected deliverables, and research goals you hope to accomplish by the end of term.

My responsibility in the UROP will be to make a program that checks the validity of inferences up to some limit in complexity. If there are not countermodels within that complexity limit to a given inference, there is good reason to believe that the inference in question is valid over all models, and so the inference in question may be either fed into a theorem prover, or checked for validity by hand.

Expected deliverables include the model checker program itself along with a range of printed results.

Finding countermodels in semantics to invalid rules is important, the program should ideally be intelligible and usable to those working in semantics without advanced programming experience. Accordingly, this project seeks to advance the methods employed in semantics, making computational tools accessible to a wider audience than it is currently.

## Personal Statement

Briefly state why you are interested in this UROP and explain what you hope to gain from it.

In studying linguistics and philosophy, I have found it extremely satisfying to make use of formal methods to model claims and problems that either seems vague, subjective, or otherwise intangible, rendering those claims/problems intelligible and tractable. This project aims to do exactly that for counterfactual conditional claims. Additionally, an adequate semantic theory for counterfactuals has important applications in AI research and linguistics, two fields I am interested in. During this UROP project I hope to learn more about the advanced models being developed in contemporary semantics as well as developing computational methods for managing the complexity of these theories. I also hope to continue to advance my abilities as a programmer, striving in particular to write code that is ease to interpret by others, especially those with limited programming experience. I will work with Dr. Brast-McKie in order to develop clear documentation for the program so that others may make use of these resources for related research.