New Course Proposal
Linguistics 280: Language and Formal Reasoning

December 7, 2004

Contents

1 Rationale

Our modern technological society increasingly demands that we manipulate information using formal languages–computer languages and database interfaces,the language of mathematics and the technical sciences, perhaps even symbolic (mathematical) logic. While the analysis and structure of valid deductive reasoning is well codified in traditional symbolic logic, the use of information and reasoning by humans necessarily passes through the medium of natural languages like English. The relationship between the form of natural language sentences and that of formulas in logical languages seems at first blush anything but straightforward and is often thought to be ultimately erratic. But though this relationship is complicated, it is actually highly systematic and can be readily analyzed using techniques from theoretical linguistics.

There are already courses in the University in the scientific analysis of natural language syntax and semantics, viz. in the Linguistics Department (in addition to many courses that analyze natural languages by humanistic methodologies in the language departments). Likewise, there are many excellent courses that treat one kind of formal artificial language or another: the many symbolic systems of mathematics, a variety of computer languages, and, notably in the present context, mathematical logic (symbolic logic), in the Philosophy and Mathematics departments. There are also courses introducing students to the practical side of natural language processing by computer. But we believe there is no course accessible to most undergraduate students (notably those not majoring in linguistics, philosophy, computer science, etc.) that critically analyzes the structural relationships between the two kinds of languages, in meaning as well as in form.

The goal of this course is to lead students to think analytically about reasoning, meaning, and natural language (English) syntax in terms abstract enough to encompass both formal and natural languages, to see underlying structural similarities and also to understand more clearly the nature of the differences between the them. While various kinds of artificial languages could in principle be used for this comparison, *classical first-order logic* (FoL) is the best choice for a number of reasons, and English is the obvious natural language of choice. Besides the importance of deductive reasoning (here, FoL) in its own right, studying the interaction of reasoning and syntactic structure in natural language is one of the best windows into the mysterious topic of meaning in natural language.

The proposed course, *Language and Formal Reasoning*, has four components:

1.1 An introduction to essential features of first-order logic

A key ingredient in the design of this course is the use of a highly interactive computational instructional system for FoL, and at least one such instructional package is currently available that meets all our requirements. Though the course does not really depend on any unique software materials, it will be useful to note some specific features of this package in order to explain the role of interactive software in the course. This is a tutorial package developed over a period of years by logician-linguist Jon Barwise and philosopher John Etchemendy at Stanford University, which is integrated with their textbook *Language, Proof and Logic* (Chicago, 2003). As the authors emphasize, their approach is distinguished from other logic textbooks and software in that FoL is introduced not only as a formal deductive system (a formal *logic*), but also as a formal *language*, a means for communicating information about something. As they observe, too often students come away from introductory logic courses with skill in the mechanics of deduction but only a vague idea of what it all has to do with reasoning and language as they themselves use these every day. In their approach, the student interacts interacts initially not only with formulas but also with graphically-presented (three-dimensional) small finite models, with respect to which truth values of formulas are assessed, and which the student can modify. The application is highly interactive: it includes "games" that students play (based on Jakko Hintikka's *game-theoretic semantics* for FoL). At any step in such an exercise or in constructing a proof, students can get immediate feedback on their work, and their exercises can be graded computationally. This step-by-step tutorial method is essential to make this course both accessible and interesting to undergraduate students who may not otherwise have a mathematics or computer science background. (We plan to develop comparable instructional software for the English syntax portion of this course, and eventually to produce our own version of logic tutorial software that can be integrated with it.)

1.2 A context-free phrase structure grammar for (essential parts of) English syntax

English syntax is treated by means of a (context-free) phrase-structure grammar: this is (in a fairly simple presentation) the way natural language syntax is analyzed in theoretical and in computational linguistics (but this part of the course does not require any prior experience with computer languages or even any sophisticated knowledge of English syntax). This illustrative English grammar includes the logical determiners (*every, a/some, no, not every*), nouns, intransitive and multi-place verbs, verbal negation, attributive and predicate adjectives, coordination (of predicates as well as of clauses), reflexive ("bound") pronouns, and restrictive relative clauses.

1.3 English-to-FoL Translation

As the course progresses, the student is led though the discovery of an algorithm for translating the syntactically-analyzed English sentences compositionally (i.e. exactly according to their syntactic construction) into corresponding formulas of FoL. The procedure is non-deterministic *only* in the multiple scope relations it yields for multiple quantifiers and negation. These translation relationships provide the basis for seeing the way an English sentence encodes its logical structure syntactically (for example, beginning with the English manifestations of the Aristotelian square, then embedding these for sentences with multiple quantifiers).

1.4 Differences between communicating with artificial and natural languages

(Final two weeks) The confusing differences between natural and symbolic languages that obscure their structural parallels arise mostly because human languages are used by intelligent communicators who always have in mind (if only subconsciously) the current information state and communicative goals of their interlocutors and know that those interlocutors are aware of theirs. Via reasoning that exploits these mutual assumptions, human communicators convey significant additional meaning above and beyond what is literally said (*conversational implicature*), thereby achieving an economy of words. A surprising amount of what we think of as the meaning of what we say is actually implicature. The importance of this aspect of meaning was first recognized by philosopher Paul Grice (and also philosophers who analyzed direct and indirect speech acts); it has now become a core doctrine in the field of semantics/pragmatics within linguistics. As the use of artificial languages rarely if ever involves implicature, learning to recognize where this process occurs in natural language is not only relevant for understanding natural language but also enables one to circumvent it—when for example one is trying to make a computer do what one wants, as well as to see how the language of advertising and political rhetoric literally states only the truth but misleads nonetheless because of the hearer's unconsciously cooperative communication strategies.

1.5 This course versus introduction-to-logic courses: some caveats

It is important to stress that this course does not attempt to provide the full introduction to first-order logic that constitutes the standard first course in logic. In the first place, there is not enough time in one quarter to provide such an introduction and accomplish the other main goals of this course (especially at the 200-level). Secondly, the Symbolic Logic course Philosophy 250 already exists which does an excellent job of this, as do Mathematics 345 and 366 in a different context. Thirdly, we believe that the traditional introduction to first-order logic is not needed to achieve the course goal of analyzing the similarities and differences between the two kinds of language in structure and as vehicles of reasoning: this goal requires an appreciation of what a formal deductive system is (with stress on *formal*) and an appreciation of what it can show us about reasoning, but we believe it does not require skill in turning out proofs or mastering all formal details of its construction. (By the same token, this course does not provide anything like the standard introduction to natural language syntax that one would get in many linguistics courses). However, that which is introduced will be treated fairly rigorously, and students will be made aware that they are seeing the basic parts of a mathematically well-analyzed system, which can be studied for its own sake in courses such as Philosophy 250 or Mathematics 366 and 345.

> (This course will cover: the syntax of statement logic, truth tables, validity, equivalences, standard connective elimination/introduction deductive rules, and (with limited examples) reductio and conditional proof methods, though not necessarily the full set of quantifier-connective equivalences. Exercises will be limited to fairly simple deductions, and little emphasis will be placed on strategies for finding proofs. For FoL: syntax, bound/free variables, the most useful logical equivalences; the elimination/introduction deductive rules for quantifiers, though without emphasis on detail and with only simple proofs involving multiple quantifiers in exercises. On the other hand, translation to and from English will obviously be a major focus. It should be realized however that while translation involving “tricky” English sentences of unusual syntactic or semantic structure is a familiar pedagogical device in standard logic courses (and a legitimate one there), the focus of this course is on seeing the parallels between the syntax of the two systems, hence such esoteric English syntax will for the most part be avoided. The Barwise/Etchemendy textbook, *Language, Proof and Logic*, includes a significant amount of material that will not be covered in this course but which the curious student can peruse outside of class to learn more about FoL, including the standard metatheorems—Soundness, Completeness, etc.)

2 Course Objectives

1. **Basics of FoL (First-Order Logic):** Students will be able to use correctly: the syntax of statement logic and of FoL; truth-tables for the connectives; construction of truth tables for complex formulas of statement logic; in FoL, definitions for variable binding and quantifier scope. Students will demonstrate an intuitive understanding of the interpretations of the quantifiers, including scope of two quantifiers or quantifier and negation, e.g. by giving paraphrases of quantified formulas.

2. **FoL as a Language:** Students will understand how statement logic and FoL represent information about something; specifically: (i) when presented with a simple finite model (graphically, as in the LPL software) and a set of sentences, determine which are true and which are false, (ii) when presented with a model, construct a set of formulas that correctly describe the model, (iii) when presented with a set of sentences, construct a model that they jointly describe.

3. **Equivalences and Deduction:** Students will be able to (i) use logical equivalences of FoL, e.g. to simplify formulas; (ii) determine the validity of a given (purported) fairly simple proof in statement logic or FoL (say, no more than four premises, none of them involving more than two quantifiers); (iii) when given a fairly simple argument, construct a valid proof of it or show by some method that it is invalid.

4. **Context-Free PS-Grammars and English syntax:** Students will (i) understand how simple context-free phrase-structure grammars generate trees representing the syntactic structure of English sentences,

   i.e. when given a simple tree determine whether a given grammar does or does not generate it, say which rules were involved in its derivation; (ii) when given a simple English sentence, determine whether a given simple phrase-structure grammar can generate an appropriate tree for it; (iii) be able to recognize simple instances of sentences with ambiguous constituent structures, understand how a grammar can produce both structures, and what intuitive meaning is associated with each tree; (iv) when given a simple English sentence not generated by any grammar seen so far, suggest a reasonable modification of a known grammar that will produce a reasonable tree structure for it.

5. **Relationship between English syntax and FoL:** Students will be able to (i) use correctly an algorithm for translating a restricted range of English sentences (as produced by a given phrase-structure grammar) into formulas of 1st-order logic compositionally;

   (ii) when given a simple argument in English (stated in a suitably restricted version of English), find tree diagrams for all the sentences, use the translation algorithm to translate them into first-order logic, then construct a proof of the conclusion using the deductive rules of 1st-order logic; (iii) understand how the algorithm can result in translations of simple types of ambiguous sentences (e.g. scope of subject quantifier and negation; scope of two quantifiers).

6. **Speech Acts and Indirect Speech Acts:** Students will understand the basic Austinian doctrine of speech acts and illocutionary force, and will be able to identify examples of indirect speech acts (viz. the syntactic sentence type vs. the speech act they convey).

7. **Conversational Implicature:** Students will (i) know the Gricean Conversational Maxims and Sub-Maxims and understand how they generate conversational implicature;

(ii) when given a sentence and a situation of utterance, be able to say what implicature is produced and the steps by which it comes about, especially for scalar quantity implicatures; (iii) understand what flouting a maxim is and be able to identify cases of flouting and say which maxim is being flouted; (iv) be able to use, in simpler cases, the diagnostics for distinguishing implicature from literal asserted meaning.

3 Relationship to Other Courses and Curricula

This course is intended primarily as a general education course rather than a course for majors in linguistics (although it will count as an elective option for a linguistics major or minor). Approval for the GEC Math and Logical Analysis requirement is being sought. Other courses which have related content are the following:

Philosophy 250 (Symbolic Logic): Both this course and the proposed new Ling 280 employ First-Order Logic (FoL) as applied to ordinary deductive reasoning. Phil 250 obviously covers FoL in greater depth, since that course is devoted entirely to it, whereas Ling 280 also deals with the formal properties of English syntax and their semantic relationship to FoL (a relationship essentially left unanalyzed in most introductory logic courses). Ling 280 also treats broader properties of natural languages (implicatures) that differentiate them from formal languages.

Mathematics 366 (Discrete Mathematical Structures) and 345 (Foundations of Higher-Order Mathematics): These courses also treat FoL, but here (i) for their own mathematical properties and (ii) in their role (along with set theory, recursive function theory, number theory, etc.) in the foundations of mathematics. Neither the role of FoL in ordinary reasoning nor itsr elationship to natural languages are a focus in these courses.

Linguistics 384 (Language and Computers): This course is like the proposed 280 in treating the syntax of natural language formally, but this is done only in the service of introducing students to language technology—such as using computers for the digital storage of speech, pattern detection in large natural language corpora, machine translation, dialog systems, spelling/grammar correction, foreign language teaching. It does not deal with FoL, natural language reasoning, or the relation between form and meaning.

The relationship between natural language syntax and meaning is studied, using FoL and other logics, in the advanced Linguistics courses 502, 602.01,02 (Syntax), 680 (Formal Foundations of Linguistics) and 683.01,02 (Semantics), but these are intended for graduate students and upper-level linguistics majors. Linguistics 280 is specifically designed as a GEC Mathmatical and Logical Analysis (MLA) course appropriate for any undergraduate at the 200-level (see attached GEC Course Justification).

4 Budget Implications

No budget adjustments will be required for this course. Several current faculty members in the Linguistics Department are qualified to teach it. Other undergraduate and graduate courses in linguistics will be offered slightly less frequently in order to accommodate this new course in the Department's regular course offerings.

5 SAMPLE COURSE SYLLABUS

Linguistics 280: Language and Formal Reasoning
Course Syllabus

Course Goals

The goal of this course is to lead you to think analytically about syntax, meaning, and reasoning in terms abstract enough to encompass both natural languages (like English) and artificial formal languages (in this case, first-order logic) — to see underlying structural similarities and to understand some fundamental differences as well. This goal is accomplished by (i) introducing you to two kinds of formal systems, first-order logic and formal phrase-structure grammars, (ii) using these systems to analyze syntax and reasoning, in symbolic form and in English, and (iii) examining differences between artificial and natural language in principles of cooperative communication.

GEC Mathematical and Logical Analysis Requirement:

The Mathematical and Logical Analysis requirement of the GEC is described as *[a] course that focuses on argument in a context that emphasizes natural language, mathematics, computer science or quantitative applications* or one which *emphasize[s] the nature of correct argumentation either in natural languages or in symbolic form*, and further that such courses *emphasize the logical processes involved in mathematics, inductive or deductive reasoning*. As stated under *Goals* above, both symbolic reasoning *and* natural language reasoning are major foci of this course, for it is a course about the relationship between the two. Besides introducing (first-order) logic, which characterizes correct deductive reasoning symbolically, this course examines natural language reasoning through the strategy of analyzing English grammar symbolically as well, then transforming analyzed English sentences systematically (algorithmically) into logical formulas, enabling you to see both the similarities and differences between reasoning in the two systems.

Textbook, Software and Other Readings:

1. *Language, Proof and Logic*, by J. Barwise, J. Etchemendy, et al. This comprises a textbook and four instructional computer applications for a 1st-order logic course:

   LPL Software (included with textbook):
   Tarski's World: evaluating formulas with respect to models.
   Fitch: constructing and verifying proofs
   Boole: constructing truth tables
   Submit: submitting homework for on-line grading

   (Assigned reading in the textbook consists of pp. 1–212, except the sections marked *optional*.)

2. ”A Phrase-Structure Grammar for some English Sentence types”, Set of instructional materials written by course instructors (David Dowty, Carl Pollard, Robert Levine) (in course packet) (about 12 pages)

3. ”Syntax and Logical Form: From PS-tree to logical formula: Formal procedures for compositional translation,” Set of instructional materials written by course instructors (in course packet) (about 20 pages)

4. “Logic and Conversation”, by Paul Grice, in *Syntax and Semantics **3**: Speech Acts* (Academic Press, 1 975), 41–58 (in course packet)

5. “Conversational Implicature”, from *Meaning in Interaction : an Introduction to Pragmatics*, by Jenny Thomas (Longman, 1995), 55–78 (in course packet)

Assignments and Grading: In most of the classes, homework exercises will be assigned and will usually be due at the next class meeting. You will use your LPL software package to do many of the exercises and to submit them on-line for grading.

There will be four in-class one-hour quizzes and a final exam. Grades will be determined as follows: Homework 30%, Quizzes 30%, Final exam 30%, Class participation 10%. **Final Examination Date/Time:** Tue, June 7, 2005, 9:30 am — 11:18 am.

Class meetings: two 1 1/2-hour class meetings per week, plus one 1-hour session for review and discussion of homework assignments.

Classes Times: M,W, 11:30–1:18 EN 212 **Review/Discussion Session:** F 11:30–1:18 EN 212

Instructor: David Dowty, Prof. of Linguistics
Office – Oxley 209
Email: dowty@ling.ohio-state.edu
Phone: 292-5400
Office Hours: T,W 3:30–4:30 (other times by appointment)

Week 1 Course introduction: Definition of ‘natural language’ vs. ‘formal language’. Some of the fundamental universal properties of natural languages. Some logical puzzles stated non-formally (”word puzzles” from Smullyan’s books etc.) to illustrate complexity of deductions in ordinary language. Traditional grammar and formal (generative) grammar. Readings: handouts

Week 2 Statement Logic: syntax (predicates, arguments, connectives, formulas). Truth tables for connectives and for complex formulas. Validity, contradiction, and entailment. Reading: *LPL* ch. 1,3; exercises using Tarski’s World, Boole (LPL Software).

Week 3 Statement logic: Logical Equivalences and Proofs. Reading *LPL* ch. 2,4,5, 6 (parts); exercises using Boole, Fitch.

First one-hour quiz during review/discussion period this week.

Week 4 Predicate logic: syntax, free and bound variables. Intuitive semantics of quantifiers; ambiguities of quantifier and negation scope. Quantifier equivalences and entailments; Readings: *LPL* ch. 9, exercises in Tarski’s World

Week 5 Symbolization of basic existential and universal English sentences, the Aristotelian square. (Some) deductive rules for quantifiers. Reading: *LPL* Ch. 95–9.8; 12 Exercises: using Tarski’s World, Fitch with quantificational formulas.

Second one-hour quiz during review/discussion period this week.

Week 6 Phrase-Structure Grammars and trees. Using PS-grammars to describe English syntax. Identifying sentences generated by a PS grammar and finding trees for sentences. Ambiguity. ”Templates” for quantifier translations involving *every, a(n), no*. Readings handouts 1–4 from “PSGE”

Week 7 English syntax and Logical Form: A procedure for systematically translating English into 1st-order logic. Quantificational noun phrase translation rule and multiple scope possibilities.

Quantifier Equivalence Laws of FoL. Reading *LPL* Ch. 10. (Continued use of LPL software for truth conditions and simple deductions in 1st-order logic). Readings: Handouts "Syntax and Logical Form: from PS-tree...". Exercises on translation, Fitch exercises.

Week 8 More English Syntax and translations: Negation and quantifier scope: relevant 1st-order logic equivalences. Kinds of conjunction in English and quantifier scope. Modifiers. Pronouns are bound variables. Restrictive relative clauses. Handouts 9-12 from "SaLF". Exercises Fitch, more deductions.

Third one-hour quiz during lecture/discuss session this week.

Week 9 Pragmatics: Speech acts and indirect speech acts. Paul Grice's theory of meaning, cooperativeness, and implicature. 'What is suggested' vs. 'what is at issue' vs. 'what is assumed'. The Conversational Maxims. Reading: Grice "Logic and Conversation", Horn "Presupposition and Implicature", handouts. Exercise on identifying and calculating implicatures

Week 10 The conversational maxims in detail; computing additional meaning (implicature) using the maxims. Scalar Quantity Implicatures. Normal use versus flouting of the Maxims. Tests for distinguishing literal meaning from implicature and also from presupposition. Readings: handouts, excerpts from Levinson.

Third one-hour quiz during lecture/discussion session this week.

Students with Disabilities: Students who need an accommodation based on the impact of a disability should contact the instructor to arrange an appointment as soon as possible to discuss the course format, to anticipate needs, and to explore potential accommodations. The instructor relies on the Office of Disability Services for assistance in verifying the need for accommodations and developing accommodation strategies. Students who have not previously contacted the Office for Disability Services are encouraged to do so (614-292-3307; www.ods.ohio-state.edu).

Academic Misconduct: (I am required by the University to make this statement: Academic dishonesty of any kind is not allowed at this university. Faculty are required to turn over any evidence of possible cheating on tests or on other assignments to the University Committee on Academic Misconduct: faculty members are not allowed to decide whether academic misconduct actually has taken place or take any steps on their own to resolve a situation. The most common form of misconduct is plagiarism: remember that any time you use the ideas or the materials of another person or persons, you must acknowledge that you have done so in a citation. This includes material that you have found on the Web as well as in books from the library, textbooks, journals or magazines, or work your roommate has handed in. See http://studentaffairs.osu.edu/resource_1csc.asp for a detailed explanation of what constitutes plagiarism. The University provides guidelines for research on the Web at http://gateway.lib.ohio-state/tutor/.

GEC Mathematical and Logical Analysis (MLA) Justification for Proposed Course: Linguistics 280: Language and Formal Reasoning

The goal of the proposed new course is to lead students to think analytically about syntax, meaning, and reasoning in terms abstract enough to encompass both natural and artificial languages—to see underlying structural similarities, and also to understand some fundamental differences as well. This goal is accomplished by introducing students to two kinds of formal systems, using these systems to analyze syntax and reasoning, in symbolic form and (indirectly) in English, and by comparing artificial languages with natural language in principles of communicative use. (See accompanying new course proposal and syllabus.)

According to the Model Curriculum, a course satisfying the Mathematical and Logical Analysis requirement of the General Education Curriculum should be:

> *[a] course that focuses on argument in a context that emphasizes natural language, mathematics, computer science or quantitative applications not primarily involving data. Courses which emphasize the nature of correct argumentation either in natural languages or in symbolic form would satisfy this requirement, as would many mathematics or computer science courses. ... The courses themselves should emphasize the logical processes involved in mathematics, inductive or deductive reasoning, or computing and the theory of algorithms.*

Both symbolic reasoning *and* natural language reasoning are major foci of this course—the focus is after all the relationship between the two. We believe the course further meets the spirit of the requirement in that in introduces two additional symbolic systems, formal generative grammars for natural language syntax and formalized translation procedures for mapping analyzed English sentences into logical formulas. Specifically:

1. Formal symbolic systems studied:
   (a) Statement Logic and First-Order Logic (FoL): both as formal deductive systems and as interpreted formal languages (model theory).
   (b) Formal generative grammars, including finite-state and context-free grammars.
   (c) Algorithmic translation from trees generated by formal PS grammar (representing English sentences) into formulas of FoL.
2. Analysis of Syntax, Reasoning and Natural Language:
   (a) Purely Symbolic Reasoning: FoL
   (b) Natural language syntax (core grammar of English) analyzed as a formal language using a CF grammar (interpreted as generating phrase-structure trees).
   (c) Understanding of 'Logical Form' in English sentences via formal syntax and FoL translation; indirect analysis of natural language reasoning via this two-step system

3. (Last two weeks) Differences between meaning in formal languages (FoL) and meaning in natural languages that arise through principles of communicative use (conversational implicature); speech acts direct and indirect. (Though not a formal topic, this is highly relevant to understanding the nature of semantically interpreted formal symbolic systems, by understanding what properties they do *not* have that we take for granted in natural languages.)

Relationship to other similar MLA GEC courses:

See section 4 in main Course Proposal.