WHAT IS NATURAL ABOUT NATURAL DEDUCTION

1. MAKING AND DISCHARGING HYPOTHESES I think that certain deductive systems were first called *natural* to contrast them with axiomatic deductive systems, either with or without a rule of substitution. The axiomatic systems were designed to establish results that were single formulas, or single sentences, and these “logical laws” were inferred directly from other logical laws. The systems of *Principia Mathematica* were like this, as are the systems presented in Hilbert and Ackermann’s *Principles of Mathematical Logic* and Church’s *Introduction to Mathematical Logic*. When compared to deductive reasoning carried out using expressions of natural language, these logical deductions have a somewhat “contrived” character.

Systems of natural deduction, in contrast, have arguments, or deductions, or proofs which begin with *hypotheses*, and infer consequences of these hypotheses. A system of natural deduction might establish results linking sets of premisses to conclusions; these would not be single-formula or single-sentence results. However, systems of natural deduction typically employ rules (inference principles) which *discharge*, or *cancel* hypotheses. For example, this argument from three hypotheses to the conclusion ‘C’:

$$\dfrac{B \qquad \dfrac{A \quad [A \supset [B \supset C]]}{[B \supset C]}\ \textit{Modus Ponens}}{C}\ \textit{Modus Ponens}$$

might be continued, using a rule $\supset$ *Introduction*, to obtain an argument from ‘$[A \supset [B \supset C]]$’ to this conclusion: $[B \supset [A \supset C]]$. A system of natural deduction can establish single-formula or single-sentence results. A particular system might establish only single-formula or single-sentence results.

It is essential to a system of natural deduction that it contain proofs, or arguments, from hypotheses which need not themselves be logical laws. It is also essential, if a natural deduction system is to accommodate arguments that are commonly made when using expressions of natural language, that the system contain rules which discharge hypotheses. We might show that a statement A is true, for example, by deriving a contradiction from the hypothesis $\sim A$. We discharge the negative hypothesis when we accept A. When making an inference which discharges a hypothesis, we are actually using a whole (sub-)argument as a premiss. The discharged hypothesis is “confined” to the subargument premiss.

Natural-deduction systems typically employ introduction and elimination rules which involve occurrences of a single operator. There might be the rules $\supset$ *Elimination*, or *Modus Ponens*, and *v Introduction*:

$\supset$ Elimination *v Introduction*

$$\frac{A \qquad [A \supset B]}{B} \qquad\qquad \frac{A}{[A \vee B]} \qquad \frac{B}{[A \vee B]}$$

From this perspective, *Modus Tollens*:

$$\frac{[A \supset B] \qquad {\sim}B}{{\sim}A}$$

would not be a suitable rule, because it "involves" two connectives rather than just one. A one-operator rule makes clear the contribution of that one operator, but with two operators, it isn't so clear what role each operator plays. However, it is not absolutely essential that a system of natural deduction employ one-operator introduction and elimination rules. A system of natural deduction might, after all, have *Modus Tollens* as one of its rules.

Arguments, or deductions, in a system of natural deduction resemble deductive arguments or proofs in fields such as science or mathematics. To establish a conditional result that if *A*, then *B,* we might begin by supposing *A*, and then, by a deduction employing one or more steps, reason to the conclusion *B*. This would normally be taken as sufficient to establish the conditional result, especially if the result to be established were stated before deducing *B* from *A*. The arguer isn't required to *finish* her deduction by saying, "Hence, if *A*, then *B*." But what she isn't required to say is nevertheless tacitly understood. The supposition of *A* is discharged once it has been shown that if *A*, then *B*.

Someone who begins an argument (proof) by supposing that $\sqrt{2}$ is a rational number, and then deduces a contradiction, might conclude by saying, "So $\sqrt{2}$ isn't a rational number," which cancels the initial supposition. If she had announced the result before giving the argument, she might not repeat that result, but could easily say something like "QED" after reaching the contradiction. With the proof by contradiction, she would probably say something to indicate that the conclusion has been established. However, this isn't the case with all arguments that discharge hypotheses. When establishing a conditional, or a universal claim, or even the consequence of a disjunction, once the argument from hypotheses which need to be discharged is completed, it is taken to be evident that the appropriate result has been established, and this is often, perhaps usually, not stated explicitly. If we begin by letting *ABC* be a triangle, and eventually conclude that the sum of the interior angles of this triangle is equal to two right angles, we can simply stop, and our audience will know that our result holds for all (Euclidean) triangles.

2. ILLOCUTIONARY ACTS We commonly do carry out deductive reasoning from hypotheses in natural languages, where the hypotheses must be discharged in order to establish results. We find this natural. People reasoned in this way before the development of formal logical theories.

Systems of natural deduction formalize a kind of reasoning that has long been familiar in deductive sciences. Although it is widely recognized that arguments from hypotheses are a natural form of reasoning, certain features of these arguments are frequently not recognized. We commonly mark a hypothesis in an argument by using either the word ‘suppose’ or the word ‘let’ (as in “Let *ABC* be a triangle”). I shall call these hypotheses *suppositions*; they are *initial suppositions* of deductive arguments.

Suppositions are most appropriately compared with assertions, or judgments, and denials. Assertions and denials are *illocutionary acts*. A statement can be accepted (as representing, things as they are) or rejected. Considering a statement and coming to accept it is an act. Once a person has performed this act, she continues to accept the statement until she either changes her mind, or forgets that she has accepted the statement. Continuing to accept a statement is a state rather than an act. An assertion is either an act of coming to accept a statement, or an act reaffirming one’s continued acceptance of the statement. Assertions are typically directed by a language user to an addressee. But I am using this word to cover acts of producing a statement, and accepting or reaffirming it, whether or not there is an audience. An assertion can involve a spoken statement, or a written statement, or even a statement which is merely thought. A denial is an act of rejecting a statement as one not suitable for being asserted. A *statement* is a *speech act*, or *language act*, which is true or false, and which represents things as being this or that.

A statement can be asserted or denied, it can also be (positively) supposed to be the case–this is to accept it temporarily, or provisionally. However, it isn’t only statements that can be supposed to be the case–we can also do that with schematic sentential expressions like ‘*ABC* is a triangle.’ A statement or schematic language act can also be temporarily or provisionally rejected. On this understanding, if we make a supposition, and infer a consequence of the supposition, our conclusion also has the status of a supposition, and will be called (by me) a supposition. Like assertions and denials, suppositions are illocutionary acts.

We can distinguish *initial* suppositions from *dependent* suppositions, which are derived from, and depend on, other suppositions. An assertion or denial can be an initial assertion or denial in a particular argument, but if, say, an assertion is derived from other illocutionary acts, it doesn’t depend on those acts. We can simply accept the asserted statement, and disregard the premisses from which it was derived.

A genuine deductive argument is a speech act, and its basic components are illocutionary acts. The argument might begin with assertions, denials, and suppositions, and proceed to a conclusion which is also an act of one these kinds. To understand, investigate, and evaluate speech-act arguments, we need to pay attention both to truth conditions of statements and to the illocutionary force with which statements are made. Standard systems of logic, or logical theories, fail to do this. Standard deductive systems do not “provide” for making, or representing, speech-act arguments. Instead, they license what I shall call *deductive derivations:* these are concerned solely to investigate truth conditions, and to trace truth-conditional “connections” among statements. It is easiest to see this in an axiomatic deductive system where both axioms and theorems are logical laws or (if they are sentences) logical truths.

Even standard systems of natural deduction, though their derivations mimic natural arguments to a certain extent, are exclusively focused on truth-conditional connections. These systems don't have a notational device for indicating illocutionary force, and so lack the resources to recognize or give an account of certain features that are essential to the correctness of a deductive argument. Imagine that we have a formal language with expressions for representing true or false statements, and, perhaps, with schematic sentential expressions. Then let us introduce the expressions for indicating illocutionary force. These are prefixed to the sentential expressions:

⊢ – the sign of assertion	⊢ –the sign of denial
⌞ –the sign of supposing true	¬ –the sign of supposing false

A sentence in a language of propositional logic that is composed exclusively from atomic sentences and conventional logical operators is a *plain sentence* of the logical language. The result of prefixing a plain sentence with an illocutionary operator is a *completed sentence*. Plain sentences represent statements, while completed sentences represent illocutionary acts.

A derivation like the following:

```
[A & [B & C]]
------------------ &E
    [B & C]            [D & E]
    --------- &E       ---------- &E
        B                   D
        ------------------------- &I
                [B & D]
```

can enable a person to determine that if the hypotheses '*[A & [B & C]]*' and '*[D & E]*' are true, then '*[B & D]*' must also be true. In order to modify this derivation so that the resulting construction fully represents a genuine speech-act argument, the expressions on each line must be prefixed with an expression which indicates illocutionary force.

One way of doing this yields the following representation:

```
⊢[A & [B & C]]
------------------- &E
    ⊢[B & C]            ⊢[D & E]
    ---------- &E       ----------- &E
        ⊢B                    ⊢D
        -------------------------- &I
                ⊢[B & D]
```

This represents an argument in which every step is an assertion. An argument having this form is, evidently, deductively correct. (It isn't appropriate to characterize such an argument as being valid or invalid, as those expressions are customarily understood.) It is important to realize that this representation, and the argument that is represented, is essentially *first-person*. The *illocutionary operators* indicate the illocutionary acts being performed by *the person who is making the argument*.

If the steps in the original representation are prefixed with different illocutionary operators, as in the following:

```
⊢[A & [B & C]]
------------------- &E
    ⊢[B & C]            ⌞[D & E]
    ----------- &E      ----------- &E
        ⊢B                    ⌞D
        -------------------------- &I
                ⊢[B & D]
```

the argument that is represented is not deductively correct. The last "move" is the one that is mistaken. A premiss that is asserted, when combined with a positive supposition, will not support an asserted conclusion. For an argument like this (an argument whose initial acts are all positive) to be deductively correct, any way of satisfying the truth conditions of the initial statements must also satisfy the truth conditions of the statement in the conclusion, and, in addition, the illocutionary force of the conclusion of an argument must not exceed the force of the premisses.

In a system of illocutionary logic, at least in the kind of system which I'm promoting here, the deductions are perspicuous representations of the kind of speech act arguments that are actually performed outside of logical studies. They make it convenient to distinguish considerations of truth conditions from those involving illocutionary force, and should prove useful for investigating features like cogency and rigor.

3. ILLOCUTIONARY SEMANTICS It is to some extent arbitrary how we "demarcate" the field to be called semantics. Semantics might be construed as the study of truth conditions of statements. It can be construed more broadly than this. If semantics is confined to truth

conditions and features that can be defined in terms of truth conditions, then what I have called deductive derivations pretty much exhaust the deductive techniques for exploring semantic concepts.

I favor a broader understanding of what is included in semantics. On my conception, both truth conditions and illocutionary force are semantic features of speech acts. Statements are those sentential acts which have truth conditions, and they are considered in abstraction from illocutionary force. Illocutionary acts are constituted by sentential speech acts performed with a certain force. The "contents" of assertions, denials, and suppositions, among others, can be regarded as statements. Not every sentential speech act is a statement or contains a statement. In a system of illocutionary logic which deals with arguments like those illustrated earlier, we need to recognize a semantic feature or features of illocutionary acts in terms of which we can characterize deductively correct arguments.

I think the appropriate feature is the one I call *rational commitment*. This feature, when recognized, can motivate a person to perform an intentional act. Making a decision to carry out a given action rationally commits a person to carry out that action. Performing some intentional acts can commit a person to perform, or not perform, other intentional acts. We can also be committed to remain in a certain state, like that of accepting a given statement. But rational commitment need not involve a moral requirement. I can decide to do something like buy gas on the way to work, and then fail to do it, either because I forget what I intended to do, or for some other reason, without being culpable in any way.

Some commitments are conditional, like the commitment to close the windows upstairs if it rains while I am at home, and others, like my commitment to buy gas on the way to work, are unconditional. Coming to accept, or continuing to accept, some statements, while rejecting other statements, will commit a person to accepting further statements, and to rejecting further statements. Positively or negatively supposing statements will commit a person to supposing others (either positively or negatively). If the person who accepts certain statements, and rejects others, is committed to, say, accept statement *A*, this commitment is conditional. She is committed to accept *A* if she has some interest in the matter, and gives it some thought. Although I accept many statements and reject many others, I am not interested in exploring the consequences of most of these beliefs and disbeliefs. However, it is irrational to accept a disjunction "*A or B*," reject *B*, but refuse to accept *A*.

Implication relations among statements are ontological, or *ontic*, features, not epistemic ones. A group of statements either implies some further statement or it doesn't. Some implications are more difficult to recognize, or grasp, than others, but the complicated cases of implication are not composed of, or constituted by, simpler cases. However, with rational commitment, which *is* an epistemic feature, there is an important difference between *immediate* commitment and *mediate* (or remote) commitment. Immediate commitment is evident to the person for whom it is immediate. If doing *A* immediately commits a person to doing *B*, and doing

B immediately commits her to doing *C*, then doing *A* may only mediately commit her to doing *C*. It is immediate commitment which, when recognized, motivates a person to act.

A given person, at a time, is characterized by her commitments to perform or not perform certain acts, including her commitments to accept or reject certain statements. Commitment is a natural feature of the human landscape. In a genuine deduction, a *natural* deduction, a person traces the immediate commitments of her illocutionary acts and of the states that her acts reflect. A truly natural deduction involves moves from illocutionary acts to illocutionary acts, and the correctness of such a deduction depends on both truth conditions and illocutionary force. But the commitment relation linking illocutionary acts and their associated states to further acts and states is constituted by a combination of illocutionary force and truth conditions.

Since commitment involves doing or not doing certain things, or continuing in a state that is intentionally entered into, the most appropriate way to present, and to explain, deductive commitments is by presenting deductive principles and developing a deductive system. The very idea of commitment in a system of illocutionary logic is best captured by characterizing deductions. But deduction and deductive systems are not semantic, or are not conceived as semantic. A logical semantic account employs functions which assign things to expressions, so that the values of complex expressions are partly or wholly determined by the values of their components. In a standard system of logic, or the ontic component of a system of illocutionary logic, the semantic account seems to capture the truth-conditional meanings of logical expressions. The deductive system, which is a system of what I have called deductive derivations, codifies certain semantically distinguished expressions of the formal language.

In a system of illocutionary logic, at the epistemic level, it is the deductive system that best captures the idea of commitment, but a semantic account which assigns values to completed sentences can be provided. I consider this account to be semantic because of the techniques employed in presenting and exploring it, not because this account captures meaning in some intuitive sense. The semantic account is initially a conjecture, which is established once the deductive system is shown to be sound and complete with respect to the semantic account.

To develop a semantic account for a system of illocutionary logic, we employ two kinds of function. The first kind of function, interpreting functions, determine the truth and falsity of plain sentences of the logical language in pretty much the standard way. In a system of propositional logic, for example, an interpreting function assigns truth and falsity to the atomic plain sentences of the language, and this determines a valuation of all the plain sentences. In a system of first-order logic, the interpreting function of the language for a domain assigns values to individual constants and predicates, and this determines a valuation of all the plain sentences.

The second kind of function assigns values to (some of) those completed sentences which are either assertions and denials. Since certain suppositions depend on others, we don't employ functions which assign values to suppositions. For assertions and denials, we consider an idealized language user, a woman whom I call the *designated subject*, at some particular time. At this time, there are statements which she has accepted and continues to accept, and statements which she has rejected and continues to reject. These *explicit beliefs* and *disbeliefs*, at that time, commit her to accept further statements, and to reject further statements. I use the plus sign '+' for the value of assertions and denials that she has already performed or that she is committed to perform at that time.

A *commitment valuation* is a function which assigns the value + to some completed sentences $\vdash A$ and $\dashv B$ of the logical language. We say that a commitment valuation V is *based on* an interpreting function f iff (i) for every completed sentence $\vdash A$ which is assigned + by V, the plain sentence A has value T for the valuation determined by f, and (ii) for every completed sentence $\dashv B$ which is assigned + by V, the plain sentence B has value F for the valuation determined by f.

A *coherent* commitment valuation is one that is based on an interpreting function. Someone whose beliefs and disbeliefs are characterized by a coherent commitment valuation is a person whose beliefs and disbeliefs don't conflict–the beliefs might all be true, and the disbeliefs false. If commitment valuation V is based on interpreting function f, then $< f, V >$ is a *coherent pair*.

If a coherent commitment valuation V_0 characterizes the designated subject's explicit beliefs and disbeliefs at a given time, this valuation determines a further commitment valuation, the *completion* of V_0, which assigns + to those assertions and denials to which the designated subject is committed by her explicit beliefs and disbeliefs. To provide a formal characterization of the completion of a coherent commitment valuation, it is helpful to consider a restricted class of interpreting functions, the *admissible* interpreting functions. Then we characterize the completion as follows:

Let V_0 be a coherent commitment valuation. The *completion* of V_0 is the commitment valuation V such that (i) $V(\vdash A) = +$ iff A has value T for every admissible interpreting function on which V_0 is based, (ii) $V(\dashv B) = +$ iff B has value F for every admissible interpreting function on which V_0 is based.

If the person who performs the illocutionary acts represented by completed sentences $A_1,..., A_n$ must be committed to perform the act represented by B, we say that $A_1,..., A_n$ *logically require* B. To formally characterize this relation, we use these definitions:

Let A be a completed sentence of the logical language, V_0 be a coherent commitment valuation, and V be the completion of V_0. Then V_0 *satisfies* A iff $V(A) = +$.

Let A, B be plain sentences of the logical language, and f be an interpreting function. Then (i) f *satisfies* $\llcorner A$ iff $f(A) = \mathrm{T}$, and (ii) f *satisfies* $\neg B$ iff $f(B) = \mathrm{F}$.

Let A be a completed sentence of the logical language, and $< f, V >$ be a coherent pair. Then $< f, V >$ *satisfies* A iff (i) A is an assertion or denial, and V satisfies A, or (ii) A is a supposition, and f satisfies A.

Now we can say that completed sentences $A_1, \ldots, A_n$ *logically require* (completed sentence) B iff (ii) B is an assertion or denial, and every coherent commitment valuation which satisfies the assertions and denials among $A_1, \ldots, A_n$ also satisfies B, or (ii) B is a supposition, and every coherent pair which satisfies all of $A_1, \ldots, A_n$ also satisfies B.

Just as we have an epistemic-level concept *logical requiring* which is a counterpart to the ontic-level concept *implication,* so we have a concept *incoherence* which is a counterpart to (semantic) *inconsistency*:

Given a plain sentence A, these suppositions are (*logically*) *incoherent*: $\llcorner A$, $\neg A$, and these completed sentences are (*logically*) *incoherent*: $\vdash A$, $\dashv A$. Completed sentences which logically require incoherent sentences are themselves *incoherent*.

4. DISTINCTIVE FEATURES OF ILLOCUTIONARY CONCEPTS If statements (plain sentences) $A_1, \ldots, A_n$ imply B, then the assertions $\vdash A_1, \ldots, \vdash A_n$ logically require $\vdash B$, and the suppositions $\llcorner A_1, \ldots, \llcorner A_n$ logically require $\llcorner B$. So if a single sentence A implies B, then $\vdash A$ logically requires $\vdash B$. But if a completed sentence $\vdash C$ logically requires $\vdash D$, we cannot say that C must imply D.

To see why not, consider the first person belief operator $\boldsymbol{B}$. It is evident that the premisses of the inferences below logically require the conclusions:

$$\frac{\vdash A}{\vdash \boldsymbol{B}A} \qquad\qquad \frac{\vdash \boldsymbol{B}A}{\vdash A}$$

But it isn't the case that A implies $\boldsymbol{B}A$, or that $\boldsymbol{B}A$ implies A.

Consistency also fails to line up with coherence. A sentence '$[A \; \& \sim\boldsymbol{B}A]$' which is true is consistent, no matter who the (first) person in question is. But this sentence cannot coherently be accepted/asserted by anyone. For the assertion '$\vdash [A \; \& \sim\boldsymbol{B}A]$' logically requires both $\vdash \boldsymbol{B}A$ and $\vdash \sim\boldsymbol{B}A$.

The divergence beween implication and logical requiring, and between consistency and coherence, is responsible for puzzles like Moore's Paradox and the Surprise Execution Paradox. By focusing on illocutionary acts in addition to statements, and the logical features of both statements and illocutionary acts, we can systematically explore the differences between the truth-conditional semantic features and their epistemic counterparts. However, I judge the most important feature of illocutionary logic to be that we can develop logical systems to capture, explain, and understand various features of our linguistic and deductive practice. (Some papers that present systems of illocutionary logic are listed in the *Bibliography*.)

John T. Kearns
Department of Philosophy and Center for Cognitive Science
University at Buffalo, SUNY
Buffalo, New York 14260
kearns@buffalo.edu

BIBLIOGRAPHY

John T. Kearns. 2000. "An Illocutionary Logical Explanation of the Surprise Execution," *History and Philosophy of Logic* **20**, 195-214.

___________. 2003. "The Logic of Coherent Fiction," *The Logica Yearbook 2002*, T. Childers and O. Majer eds., Filosofia, Prague, 133-146.

___________. 2006. "Conditional Assertion, Denial, and Supposition as Illocutionary Acts," *Linguistics and Philosophy 29*, 455-485.

___________. 2007. "An Illocutionary Logical Explanation of the Liar Paradox," *History and Philosophy of Logic* 28, 31-66.

___________. 2008. "Illocutionary Origins of Familiar Logical Operators," *The Logica Yearbook 2007*. Michal Peliš ed., Filosofia, Prague, 55-66.