A Note on Interpretations for Federated Languages and the Use of Disquotation

Steven O. Kimbrough
University of Pennsylvania
3730 Walnut Street, Suite 500
Philadelphia, PA 19104 USA
kimbrough@wharton.upenn.edu

ABSTRACT
Quotation and subsequent disquotation of propositional content sentences has been proposed for, and has some attractive features for, representing content sentences embedded in complex sentences by subordinating clauses. The meaning of and proper semantics for such representations, however, remain obscure, amid lurking worries that any such language will be inconsistent. This paper sketches an account of a semantics for quotation and disquotation. The key technical devices employed are (1) composing languages by federating two or more distinctly defined languages and (2) treating quotation as an invertible enumeration device, employed in the conventions governing use of the federated language.

1. BACKGROUND
Linguists and grammarians have usefully distinguished three main types of sentences (e.g., [11, pages 142–4]). First, a simple sentence consists of a single clause (verb plus noun phrase(s)) that “stands alone as a sentence.” Roughly, a simple sentence corresponds to a predicate in first order logic and its arguments properly filled in. Second, a coordinate sentence is composed from simple sentences by such coordinating operators as and, but, and or. Roughly, in first order logic coordinate sentences are composed using the logical constants on predicates.

Third, a complex sentence is one in which “a clause can be incorporated into another clause” [11, page 143]. Davidson’s famous example, “Galileo said that the earth moves” [10], serves to illustrate important distinctions for dealing with complex sentences. “the earth moves” is a clause (roughly, verb+noun phrase properly combined) that is here “incorporated” into the clause “Galileo said”. The use of “that” is diagnostic of complex sentences; it typically serves to indicate the subordinate clause indicator is dropped, because no ambiguity results, as in “Galileo said the earth moves”.

Linguistic complexity of this kind—sentences with embedded clauses, utterances that refer to other utterances—is quite common, and as noted is often signaled by the use of “that”. Here are two examples, drawn by convenience from fiction but quite representative of actual dialog.

“There are too many other things that could have happened,” I said. “That she did go away with Lavery and they split up. That she went away with some other man and with the wire is a gag. That she drank herself over the edge and is holed up in some private sanatorium taking a cure. That she got into some jam we have no idea of. That she met with foul play.”

“Good God, don’t say that,” Kingsley exclaimed.
– The Lady in the Lake, chapter 2, Raymond Chandler

“The gun, you know, is kind of queer.”

“But I told you I found it lying on the stairs,” she said angrily. “I don’t know anything else about it. I don’t know anything about guns at all. I—I never shot one in my life.” She opened a large blue bag and pulled a handkerchief out of it and sniffed.

“That’s your story,” I said. “I don’t have to get stuck with it.”
– The Lady in the Lake, chapter 15, Raymond Chandler

Talk about talk is complex in our present context (sentences with embedded clauses). It is ubiquitous and, in particular, suffuses legal discourse (see [12] for many examples). In addition, there is a well-recognized catalog of subordinating expression types (see [23] for a discussion), including those normally associated with these verbs: believes (as in X believes that P), desires, intends, and other mentalistic concepts; asserts, promises, declares, commands, and other speech acts; it is necessary that, it is possible that, it is obligatory that, it is permitted that, and other broadly modal concepts; as well as such notions as seeing to it that, feeling that, seeing that, hearing that, requiring that, and needing that.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICAIL ’05, June 6-11, 2005, Bologna, Italy.
Copyright 2005 ACM 1-59593-081-7/05/0006 ...$5.00.
Complex sentences, utterances with embedded propositional content, pose considerable challenges to formalization in logic. First order logic does not permit, at least in any straightforward way, representation of embedded clauses. Moreover, first order logic is extensional and most complex sentences are intensional to some degree. Simplifying, we might express the generic complex sentence as having the form $P$ that $Q$, where $P$ is the subordinator, the embedding clause, and $Q$ the subordinated, or embedded, clause (think: $P =$ Galileo said; $Q =$ the earth moves). In many cases, it may be true that $P$ that $Q$ and that $Q \rightarrow R$, but not that $P$ that $R$. For example, let $Q$ be $2+2=4$ and $R$ be $123-121=8-6$. Not only is it the case that $Q \rightarrow R$, but it is necessarily the case, yet it may well be true that Galileo said that $Q$ and did not ever say that $R$. Similarly, it may be true that $P$ that $Q(a)$, that $a = b$, and not true that $P$ that $Q(b)$. Standard examples are created from $a =$ the morning star and $b =$ the evening star, and from $a =$ Cicero and $b =$ Tully. The morning star is, at it happens, one and the same as the evening star. Cicero and Tully are two names for the same historical individual. Also, we have Superman and Clark Kent, and Batman and Bruce Wayne, among others. Cases are unexotically generated with definite descriptions, as the evening star.

It is obligatory, It is permitted, and It is possible contain a commendably clear and accessible treatment of the subordinated, or embedded, clause $Q$. Associated with this is the meaning postulate

$$\text{Promise}(e, s, [P]) \rightarrow (\text{Kept}(e) \leftrightarrow P)$$

which may be read as If $e$ is the event of $s$ promising that $P$, then the promise named by $e$ is kept if and only if $P$.\(^1\)

There is much to be said in principle in favor of a disquotation approach to representing embedded content. The papers cited above present evidence in this regard and prototype implementations have yielded positive results (e.g., \cite{1, 2, 3, 29}). From a theoretical perspective, note that intensionality comes in degrees. The disquotation approach defaults to the strongest level of intensionality (substitution of terms of identical meaning is not guaranteed to preserve truth) and allows, via meaning postulates, arbitrary relaxation of intensionality. This affords the modeler great flexibility. From a practical perspective, note that finding and integrating the appropriate operator-based intensional logics may often be challenging—and the required logics may not exist. The disquotation approach affords an accessible, incremental approach to modeling and implementation. It is possible to identify some of the logical structure required and to implement it with the disquotation approach, leaving open the possibility of integrating its representations with operator-oriented logics. See \cite{25} and \cite{19} for exercises of this sort. The disquotation approach is properly seen as a complement to the operator approach.

The disquotation approach is, however, not without its difficulties. Principal among them is the question of providing a semantics. Expressions (1) and (2) are not wffs (well-formed formulas) in first order logic. In what logic are they wffs? Is this logic complete? Is it consistent?

The aim of this paper is to address these questions. To this end I develop and explore the concept of what may be called a federated language. A federated language is a language comprising two or more distinct languages. I shall refer to these latter languages as federating languages. Our interest here is mainly in federated languages whose federating (sub-)languages are languages of formal logic, in particular sentence logic and first order logic.\(^2\) The goal of the present paper is to present and develop the concept, and then to apply it to the problem of making semantic sense of disquotation as used to represent embedded content.

2. PRELIMINARIES

Recall that a formal language (of logic), $L$, is specified by five elements:

1. Symbols. It will be useful to distinguish three kinds of symbols:
   
   (a) Logical constants.

   E.g., $\neg, \wedge, \rightarrow$

\(^1\)Very many details are being suppressed, or elided, in the interests of focusing on the matters to hand. For example, a valid promise is arguably to the effect that one promises that one sees to it that $P$. Also, as emphasized by Searle \cite{32, 33}, promises are arguably forward-looking; what it is one promises must be something that occurs in the future, certainly not the past. I believe that the elisions committed are without prejudice to a complete and workable theory.

\(^2\)The term federated language is a neologism so far as I know, motivated by analogy with federated database. Bowen and Kowalski \cite{7} introduced the term amalgamated language for a distinct but related concept.
2. Well-formed terms (wfts).
These are normally specified by rules of formation on the set of symbols, e.g.,
(a) If \( \phi \) is variable, then \( \phi \) is wft.
(b) If \( \phi \) is name, then \( \phi \) is wft.
Denote the set of wfts of \( \mathcal{L} \) by \( \text{wft}(\mathcal{L}) \).

3. Well-formed formulas (wffs).
These are normally specified by rules of formation on the set of symbols, e.g.,
(a) If \( \phi \) is a statement letter, then \( \phi \) is a wff.
(b) If \( \phi \) is a statement letter, then \( \neg \phi \) is a wff.
Denote the set of wffs of \( \mathcal{L} \) by \( \text{wff}(\mathcal{L}) \).

4. Axioms.
These are often given as axiom schemas, e.g., for all \( \phi \) and \( \psi \in \mathcal{L} \), \( (\phi \rightarrow (\psi \rightarrow \phi)) \).

5. Rules of derivation.
E.g., modus ponens or detachment: Given \( \phi \) and \( (\phi \rightarrow \psi) \), then \( \psi \) may be directly inferred.

Our attention here will be focused on federating languages of quite standard logic: the propositional calculus and first order logic. It matters for present purposes not at all which logical systems are used. For the sake of definiteness and accessibility, however, I shall have in mind the systems described in Richard Jeffrey’s textbook [15], which use truth tree methods. These systems have no axioms at all and rules of derivation (including tree construction rules) that are especially convenient and transparent.

The axioms and rules of derivation for a logical language might be called its logical basis. It will simplify matters, without cost of generality, to stick to a single logical basis throughout our discussion. For the same reason, it will be useful to restrict our attention to federating languages with identical sets of logical constants and punctuation marks. With these assumptions to hand, the following definitions are in order.

**Definition 1 (Language Type).** A language type is a specified logical language, which specification is unique up to its logical terms. Uniformly and unambiguously renaming (some or all of) its logical terms (e.g., renaming \( P \) as \( Q \)) does not alter the type of the language.

Note: For the sake of precluding ambiguity, new names may not duplicate old names.

**Definition 2 (Language Token).** A language token is an instance of a given language type and is identified by its logical terms. Uniformly and unambiguously renaming (some or all of) its logical terms (e.g., renaming \( P \) as \( Q \)) does not alter the type of the language, but does change one token of the language into another.

**Definition 3 (Atomically Disjoint).** Two sets of expressions, \( \Delta^1 \) and \( \Delta^2 \), are said to be atomically disjoint if they have no logical terms in common. Two language tokens, \( \mathcal{L}^1 \) and \( \mathcal{L}^2 \), are said to be atomically disjoint if they have no logical terms in common (since their respective sets of wffs are atomically disjoint).

Comments:

1. In the case of two language tokens of type propositional logic, the languages are atomically disjoint if and only if they have syntactically different sentence letters. In the case of first order logic, predicates, variables, names, and functions will have to be syntactically distinct if two language tokens are to be atomically disjoint.

2. If two language tokens of the same type are atomically disjoint, then any well-formed expression in either language can unambiguously be assigned to its proper language; no expression—whether term or formula—is well formed in more than one of the two language tokens.

3. **Federating Languages of Propositional Logic**

Let \( \mathcal{L}^1 \) be a language (token) of propositional logic, specified as in figure 1.

Let \( \mathcal{L}^2 \) be language token of the language type propositional logic and be atomically disjoint from \( \mathcal{L}^1 \). Let \( S^2 = \{ P_i^2 \mid i = 1, 2, \ldots \} \). (By atomic disjointness, \( S^1 \cap S^2 = \emptyset \).

The specifications of \( \mathcal{L}^1 \) and \( \mathcal{L}^2 \) are identical, except for their atomic sentence letters, \( S^1 \) and \( S^2 \). Thus, the specification of \( \mathcal{L}^2 \) is as in figure 1, with \( S^1 \) substituted for \( S^2 \).

Given \( \mathcal{L}^1 \) and \( \mathcal{L}^2 \) specified as indicated, the language token \( \mathcal{L}^3 \) is federated from (the atomically disjoint federating languages) \( \mathcal{L}^1 \) and \( \mathcal{L}^2 \) (we shall say \( \mathcal{L}^3 = \mathcal{L}^1 \circ \mathcal{L}^2 \)) as follows: \( \mathcal{L}^1 \) is specified as in figure 1, with \( S^1 = S^1 \cup S^2 \) uniformly substituted for \( S^1 \). Line 1(c) of figure 1 becomes

- Statement letters: Every element of \( S^3 = S^1 \cup S^2 \).

Line 3a of figure 1 becomes

- If \( \phi \in S^3 \), then \( \phi \) is a wff.

Given that a proper semantics for \( \mathcal{L}^1 \) (a standard propositional logic) is given via the valuation function \( V^1 \) and that a proper semantics for \( \mathcal{L}^2 \) is similarly given via \( V^2 \), it is straightforward to give a proper valuation function, \( V^3 \), for \( \mathcal{L}^3 \). Figure 2 provides the definition. A word on what it is to be a proper semantics. Even for an utterly standard propositional logic, \( \mathcal{L}^1 \) for example, not every valuation function is proper. In particular, valuation functions must be complete.
1. Symbols.
   (a) Logical constants: \( \land, \land, \Rightarrow, \lor, \Leftrightarrow \)
   (b) Punctuation: (, )
   (c) Logical terms.
      i. Statement letters: Every element of \( S^1 = \{ P^1_i \mid i = 1, 2, \ldots \} \).
2. Well-formed terms (wfts): none
3. Well-formed formulas (wffs):
   (a) If \( \phi \in S^3 \), then \( \phi \) is a wff.
   (b) If \( \phi \) is a wff, then \( \neg \phi \) is a wff.
   (c) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \land \psi) \) is a wff.
   (d) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \lor \psi) \) is a wff.
   (e) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \rightarrow \psi) \) is a wff.
   (f) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \leftrightarrow \psi) \) is a wff.
   (g) Nothing else is a wff.
4. Axioms: \( \emptyset \), as in [15].
5. Rules of derivation: truth tree method, as in [15].

**Figure 1: Specification of \( L^1 \), a token of type propositional logic**

and determine in the sense that they provide a unique valuation for each and every wff in the language.\(^3\) \( V^1 \) is not proper, then, if the following is among its rules:

- \( V^1(P^1_i) = \top \) if this sentence is false; otherwise \( V^1(P^1_i) = \bot \).

We can get into similar difficulties if we allow, say

- \( V^2(P^1_i) = \top \) if \( V^2(P^1_i) = \bot \); otherwise \( V^2(P^1_i) = \bot \).
- \( V^2(P^2_i) = \top \) if \( V^1(P^1_i) = \top \); otherwise \( V^2(P^2_i) = \bot \).

In consequence we prohibit \( V^1 \) from having functional dependence on \( L^2 \); similarly, \( V^2 \) cannot be functionally dependent on \( L^1 \) in any way. In short, we require that both \( V^1 \) and \( V^2 \) be proper valuation functions for their respective languages.

**Observation 1.** The language token \( L^3 \) composed (by federation) in this manner is of type propositional logic. This may be seen by syntactic transformation of \( L^3 \) into, e.g., \( L^4 \) as follows: \( P^1_i \sim P^2_{2i-1} \) and \( P^2_i \sim P^2_{2i} \). Then \( L^4 \) and \( L^1 \) are syntactic variants: uniformly substituting \( P^1_i \) for \( P^1_i \) transforms \( L^1 \) into \( L^1 \).

**Observation 2.** Federation does not generate inconsistency.

This may be stated more precisely and proved with the following lemma.

**Lemma 1.** Let \( \Gamma^1 \) be a consistent set of wffs in \( L^1 \) and \( \Gamma^2 \) a consistent set of wffs in \( L^2 \). Let \( \Gamma^3 = \Gamma^1 \cup \Gamma^2 \) be a set of wffs in \( L^3 = L^1 \circ L^2 \). Then \( \Gamma^3 \) is consistent.

\(^3\)Note, but elide for the sake of simplicity, the issue of evaluating first order formulas with free (unbound) variables.

1. If \( \phi \in S^3 \) then \( V^3(\phi) = V^1(\phi) \) if \( \phi \in S^1 \), and \( V^3(\phi) = V^2(\phi) \) if \( \phi \in S^2 \).
   Comment: Since \( S^1 \cap S^2 = \emptyset \) and since \( V^1 \) is a proper interpretation of \( L^1 \) and \( V^2 \) is a proper interpretation of \( L^2 \), it follows that \( V^3 \) is fully and uniquely defined for \( \phi \in S^3 \).
2. \( V^3(\neg \phi) = \top \) if \( V^3(\phi) = \bot \); otherwise \( V^3(\neg \phi) = \bot \).
3. \( V^3(\phi \land \psi) = \top \) if \( V^3(\phi) = \top \) and \( V^3(\psi) = \top \); otherwise \( V^3(\phi \land \psi) = \bot \).
4. \( V^3(\phi \lor \psi) = \top \) if \( V^3(\phi) = \top \) or (inclusively) \( V^3(\psi) = \top \); otherwise \( V^3(\phi \lor \psi) = \bot \).
5. \( V^3(\phi \rightarrow \psi) = \top \) if \( V^3(\phi) = \bot \) or \( V^3(\psi) = \top \); otherwise \( V^3(\phi \rightarrow \psi) = \bot \).
6. \( V^3(\phi \leftrightarrow \psi) = \top \) if \( V^3(\phi) = V^3(\psi) \); otherwise \( V^3(\phi \leftrightarrow \psi) = \bot \).

**Figure 2: A Federated Semantics for the Federated Propositional Language \( L^3 \)**

Since proofs serve the dual goals of establishing results and promoting insight, we offer two proofs of the lemma. In the semantic sense, consistency of a set of wffs means that there is some interpretation under which each member of the set is true. In the proof-theoretic sense, consistency of a set of wffs means that there is no formula for which both it and its denial may be proved.

**Proof: Semantic Version.** Since \( \Gamma^1 \) is consistent, there is an interpretation (or valuation) on \( S^1 \) (an assignment of truth values to each sentence letter in \( S^1 \)), call it \( V^1(S^1) \), under which every wff in \( \Gamma^1 \) is true. Similarly, let \( V^2(S^2) \) be an interpretation which makes every wff in \( \Gamma^2 \) true. Let \( V^1(P_j) \) be the value (true truth value) assigned by \( \Gamma^1 \) on \( S^1 \) to \( P_j \in S^1 \). Then, since \( S^1 \cap S^2 = \emptyset \),

\[
V^3(P_j) = \begin{cases} V^1(P_j) & \text{if } P_j \in S^1 \\ V^2(P_j) & \text{if } P_j \in S^2 \end{cases}
\]  \( (3) \)

is well-defined and constitutes an interpretation, \( V^3(S^3) \) (recall \( S^3 = S^1 \cup S^2 \) and \( S^1 \cap S^2 = \emptyset \), in which every wff in \( \Gamma^3 = \Gamma^1 \cup \Gamma^2 \) is true. □

**Proof: Proof-Theoretic Version.** By definition, \( \Gamma^3 \) is consistent if and only if it is not possible in \( L^3 \) to derive both \( \phi \) and \( \neg \phi \). Consider two cases: (i) \( \phi \in S^3 = S^1 \cup S^2 \) (so \( \phi \) is an atomic sentence), and (ii) \( \phi \notin S^3 = S^1 \cup S^2 \) (so \( \phi \) is not an atomic sentence). We need consider only case (i), since if we can derive \( (\phi \land \neg \phi) \) for some non-atomic wff \( \phi \), it will be trivial in the logic to hand to derive \( (\psi \land \neg \psi) \) for any atomic wff \( \psi \).

If no atomic wff \( \phi \) is derivable from \( \Gamma^3 \), then \( \Gamma^3 \) is consistent. Alternatively, let \( \phi \in S^3 \) be derivable from \( \Gamma^3 \). Since \( \Gamma^1 \) and \( \Gamma^2 \) are each consistent and \( S^1 \cap S^2 = \emptyset \), if \( \phi \in S^1 \) then \( \Gamma^2 \) is irrelevant to any derivation of \( \phi \) or \( \neg \phi \). Similarly, if \( \phi \in S^2 \) then \( \Gamma^1 \) is irrelevant to any derivation of \( \phi \) or \( \neg \phi \). Because \( \Gamma^1 \) and \( \Gamma^2 \) are atomically disjoint, no wff in \( \Gamma^1 \) or any wff produced by applications of the rules of inference on wffs in \( \Gamma^1 \) can directly contradict any wff in \( \Gamma^2 \), or any wff produced by applications of the rules of inference on wffs in \( \Gamma^2 \). Thus, if \( \Gamma^3 \vdash \phi \), then either \( \phi \in S^1 \) and \( \Gamma^1 \vdash \phi \) or
\( \phi \in S^2 \) and \( \Gamma^2 \vdash \phi \).

Assume without loss of generality that \( \phi \in S^1 \). By consistency of \( \Gamma^1 \), \( \Gamma^1 \not\vdash \neg \phi \). By consistency of \( \Gamma^2 \) and \( S^1 \cap S^2 = \emptyset \), \( \Gamma^2 \not\vdash \neg \phi \). Thus, \( \Gamma^3 \) is consistent, since for no \( \phi \) is it true that \( \Gamma^3 \vdash \phi \) and \( \Gamma^3 \vdash \neg \phi \). □

**Observation 3.** It is entirely possible to have intended interpretations of \( L^1 \) and \( L^2 \) that produce nominal but not formal contradictions in \( L^3 \).

An interpretation in sentence logic is an assignment of sentence letters to states of affairs in the world, e.g., \( P_1^1 = \) More noncombatants have been killed during the American occupation of Iraq than people were killed by the Boxing Day tsunami of 2004. While the intended interpretation serves to motivate one’s use of a language, it has no role in the semantics or proof theory for the language. For any interpretation, more than one intended interpretation is possible.

By way of illustrating this last observation, let us have the following intended interpretations:

- \( P_1^1 \): It is Tuesday.
- \( P_1^2 \): It is raining.
- \( P_2^1 \): It is Tuesday.
- \( P_2^2 \): It is raining.

The following sequent is valid in \( L^3 \):

\[
(P_1^1 \rightarrow P_2^1), (P_1^1 \rightarrow \neg P_2^2) \vdash (P_1^2 \land \neg P_2^2)
\]  

While this seeming failure to detect contraction may seem odd, it is not an anomaly. \( L^3 \) is simply an instance of sentence logic. The behavior on display here is quite available in any instance of sentence logic, as may be seen by applying our previous transformation:

- \( P_1^1 \): It is Tuesday.
- \( P_1^2 \): It is raining.
- \( P_2^1 \): It is Tuesday.
- \( P_2^2 \): It is raining.

which leads to

\[
(P_1^1 \rightarrow P_3^1), (P_2^2 \rightarrow \neg P_2^4) \vdash (P_3^1 \land \neg P_4^4)
\]  

If a composed (i.e., federated) language such as \( L^3 \) is simply a token of the type associated with its federating languages, what advantage could it possibly offer over another token, e.g., over one of its components? Any advantage would have to lie in syntax, rather than in proof theory or model theory. But one syntax may well afford us advantages not provided by another. Certainly it is possible to conceive of a hideously cumbersome syntax. On the positive side we might think of \( L^1 \) and \( L^2 \) as offering different perspectives on a given topic. Certain questions are more conveniently asked given a felicitous syntax. For example, given the intended interpretations

- \( P_2^1 \): It is raining.
- \( P_2^2 \): It is raining.

we can easily ask if, under present assumptions, the two perspectives agree on whether it is raining. It is simply convenient for this task that \( P_2^2 \) “means” in language \( i \) that it is raining. Syntactic arrangements of this sort are replaceable and augmentable by, for example, maintaining tables of correspondence between formulas in \( L^1 \) and \( L^2 \). Such tables are, of course, outside the languages in question, yet they may be useful in working with and within the languages. As we shall see, more exciting possibilities attend a composed first order language.

1. **Symbols.**
   - (a) Logical constants: \( \neg, \land, \rightarrow, \lor, \leftrightarrow \)
   - (b) Punctuation: ( , )
   - (c) Logical expressions.
     - i. Predicate letters (without regard to arity): Every element of \( S^1 = \{P_i^1 | i = 1, 2, \ldots \} \).
     - ii. Functions (without regard to arity): Every element of \( F^1 = \{f_i^1 | i = 1, 2, \ldots \} \).
     - iii. Name constants: \( N^1 = \{a_i^1 | i = 1, 2, \ldots \} \).
     - iv. Variables: \( X^1 = \{x_i^1 | i = 1, 2, \ldots \} \).

2. **Well-formed terms (wfts):**
   - (a) If \( \psi \in N^1 \), then \( \psi \) is a wft.
   - (b) If \( \psi \in X^1 \), then \( \psi \) is a wft.
   - (c) If \( \psi \in F^1 \) and has arity \( n \), and \( \chi_1, \chi_2, \ldots, \chi_n \) are each wfts, then \( \psi(\chi_1, \chi_2, \ldots, \chi_n) \) is a wft.

3. **Well-formed formulas (wdfs):**
   - (a) If \( \phi \in S^1 \) with arity \( n \), and \( \psi_1, \psi_2, \ldots, \psi_n \) are each wfts, then \( \phi(\psi_1, \psi_2, \ldots, \psi_n) \) is a wff.
   - (b) If \( \phi \) is a wff, then \( \neg \phi \) is a wff.
   - (c) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \land \psi) \) is a wff.
   - (d) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \lor \psi) \) is a wff.
   - (e) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \rightarrow \psi) \) is a wff.
   - (f) If \( \phi \) is a wff and \( \psi \) is a wff, then \( (\phi \leftrightarrow \psi) \) is a wff.
   - (g) If \( \phi \) is a wff and \( \psi \in X^1 \), then \( \exists \psi(\phi) \) is a wff.
   - (h) If \( \phi \) is a wff and \( \psi \in X^1 \), then \( \forall \psi(\phi) \) is a wff.
   - (i) If \( \chi_1 \) is a wff and \( \chi_2 \) is a wff, then \( (\chi_1 = \chi_2) \) is a wff.
   - (j) Nothing else is a wff.

4. **Axioms:** \( \emptyset \), as in [15].

5. **Rules of derivation:** truth tree method, as in [15].

**Figure 3:** Specification of \( L^1 \), a token of type first order logic

4. **FEDERATING FIRST ORDER LANGUAGES**

Here we discuss \( L^3 \) as composed (federated) from language tokens of first order logic, \( L^1 \) and \( L^2 \). Predicates of \( L^i \) \((i \in \{1, 2\}) \), regardless of arity will be denoted \( P_1^i, P_2^i, \ldots \). Similarly,
1. Symbols.
   (a) Logical constants: ¬, ∧, →, ∨, ↔
   (b) Punctuation: ( )
   (c) Logical expressions.
      i. Predicate letters (without regard to arity): Every element of $S^3 = S^1 \cup S^2$
      ii. Functions (without regard to arity): Every element of $F^3 = F^1 \cup F^2$
      iii. Name constants: $N^3 = N^1 \cup N^2$
      iv. Variables: $X^3 = X^1 \cup X^2$
2. Well-formed terms (wfts):
   (a) If $\psi \in \text{wft}(L^1) \cup \text{wft}(L^2)$, then $\psi$ is a wft.
3. Well-formed formulas (wffs):
   (a) If $\phi \in \text{wff}(L^1) \cup \text{wff}(L^2)$ then $\phi$ is a wff.
   (b) If $\phi$ is a wff, then $\neg(\phi)$ is a wff.
   (c) If $\phi$ is a wff and $\psi$ is a wff, then $(\phi \land \psi)$ is a wff.
   (d) If $\phi$ is a wff and $\psi$ is a wff then $(\phi \lor \psi)$ is a wff.
   (e) If $\phi$ is a wff and $\psi$ is a wff then $(\phi \rightarrow \psi)$ is a wff.
   (f) If $\phi$ is a wff and $\psi$ is a wff then $(\phi \leftrightarrow \psi)$ is a wff.
   (g) If $\phi$ is a wff and $\psi \in X^3$, then $\exists \psi(\phi)$ is a wff.
   (h) If $\phi$ is a wff and $\psi \in X^3$, then $\forall \psi(\phi)$ is a wff.
   (i) Nothing else is a wff.
4. Axioms: $\emptyset$, as in [15].
5. Rules of derivation: truth tree method, as in [15], but universal instantiation and existential instantiation are modified to block introduction of non-wffs.

Figure 4: Specification of a first order $L^3 = L^1 \circ L^2$

1. functions: $f_1, f_2, \ldots$
2. constants: $a_1, a_2, \ldots$
3. variables: $x_1, x_2, \ldots$

We may now form $L^3 = L^1 \circ L^2$, proceeding much as we did in the case of federating propositional logic tokens. Let $L^1$ be specified as in figure 3, the first order logic analog to figure 1. We specify $L^2$ by substituting in figure 3: $S^2$ for $S^1$, $F^2$ for $F^1$, $N^2$ for $N^1$, and $X^2$ for $X^1$. We compose the federated language $L^3 = L^1 \circ L^2$ by substituting in figure 3: $S^3 = (S^1 \cup S^2)$ for $S^1$, $F^3 = (F^1 \cup F^2)$ for $F^1$, $N^3 = (N^1 \cup N^2)$ for $N^1$, and $X^3 = (X^1 \cup X^2)$ for $X^1$, and in addition we use the wff-formation rules of propositional logic. Details are given in figures 4 and 5.

Comments:
1. In figure 4, we have dropped the wff-formation rule:
   - If $\chi_1$ is a wft and $\chi_2$ is a wft, then $(\chi_1 = \chi_2)$ is a wff.
   Note that specification in figure 4 does not allow as well formed such sub-wff mixed expressions as $P^1_2(a^2_2)

1. If $\phi \in S^3$ then $V^3(\phi) = V^1(\phi)$ if $\phi \in S^1$, and $V^3(\phi) = V^2(\phi)$ if $\phi \in S^2$.
   Comment: Since $S^1 \cap S^2 = \emptyset$ and since $V^1$ is a proper interpretation of $L^1$ and $V^2$ is a proper interpretation of $L^2$, it follows that $V^3$ is fully and uniquely defined for $\phi \in S^3$.
2. $V^3(\neg \phi) = T$ if $V^3(\phi) = \bot$; otherwise $V^3(\neg \phi) = \bot$
3. $V^3(\phi \land \psi) = T$ if $V^3(\phi) = T$ and $V^3(\psi) = T$; otherwise $V^3(\phi \land \psi) = \bot$
4. $V^3(\phi \lor \psi) = T$ if $V^3(\phi) = \bot$ or (inclusively) $V^3(\psi) = T$; otherwise $V^3(\phi \lor \psi) = \bot$
5. $V^3(\phi \rightarrow \psi) = T$ if $V^3(\phi) = \bot$ or $V^3(\psi) = T$; otherwise $V^3(\phi \rightarrow \psi) = \bot$
6. $V^3(\phi \leftrightarrow \psi) = T$ if $V^3(\phi) = V^3(\psi)$; otherwise $V^3(\phi \leftrightarrow \psi) = \bot$.
7. $V^3(\exists \psi(\phi)) = V^3(\phi)$ if $\psi^i \in X^i$ does not occur freely in $\phi$. If $\psi^i \in X^i$ does occur freely in $\phi$, $V^3(\exists \psi(\phi)) = T$ if for every $n \in N^i$, $V^3(\phi[x^i/n]) = T$; otherwise $V^3(\exists \psi(\phi)) = \bot$.
   Comment: $\phi[x^i/n]$ is the wff obtained from $\phi$ by substituting $n$ for each freely occurring instance of $x^i$. Assumed is that $i \in \{1, 2\}$.
8. $V^3(\forall \psi(\phi)) = V^3(\phi)$ if $\psi^i \in X^i$ does not occur freely in $\phi$. If $\psi^i \in X^i$ does occur freely in $\phi$, $V^3(\forall \psi(\phi)) = T$ if for some $n \in N^i$, $V^3(\phi[x^i/n]) = T$; otherwise $V^3(\forall \psi(\phi)) = \bot$.

Figure 5: A Federated Semantics for the Federated First Order Language $L^3$

(although vacuous quantification is permitted, for example, $\exists x^2_1(P^2_2(a^2_2)$). The effect of this is to separate the domains of the two sublanguages, $L^1$ and $L^2$, as in fact they were separated in the case of sentence logic, above.

2. Note in particular that the composed language has no variables of its own, thus quantification only ranges over a single domain, the domain of one of the sublanguages. It is perhaps best to think of $L^3$ as a typed first order language, with the typing emerging from the composition. Put differently: there is no identity between individuals in different domains and predicates group only individuals in a single domain. Given this restriction, creating an instance of a first order language from $L^3$ (so composed) is straightforward. Thus, analogs of the observations and the lemma regarding propositional logic, in §3, obtain in the case of first order logic. In the interests of avoiding unnecessary pedantry, I leave these as an exercise for the reader.

We turn now to an application for federated languages.

5. REPRESENTATION OF propositional CONTENT

Organizations conducting business transactions over time will design relational database schemas that facilitate the
addition and deletion of records on transactions as they occur. For this purpose, the underlying schemas need to be robust in at least two senses: expressive power (they need to be able to accommodate the transaction information) and consistency (they need to remain coherent as records are added and deleted). While the relational model is strong on consistency, it is generally acceded to be not well suited for representation of complex sentences and propositional content, at least in any straightforward way. The aim for the present section is to sketch an account both of (1) how we can represent complex sentences and propositional content in a semantically proper way, and (2) how such a representation is suitable for recording business transactions, and hence for serving as a proper basis for practical inference.

The exposition proceeds first by example and demonstration, then by generalization. Starting with a concrete example, recall expressions (1), a record of the utterance of a promise, and (2), a meaning postulate pertaining to the promise. We'll proceed in steps to make sense of these expressions. For the sake of simplicity, let L1 be the L1 of §3, a standard propositional logic language. Let L2 be the L2 of §4, a standard first order logic, but with the following minor modifications. (L2 will remain a first order logic.) Let the set of name constants in L2, N2, have its elements relabeled and partitioned into three distinct sets: E2 = {e1, e2, ...} (for naming events); O2 = {o1, o2, ...} (mnemonic: other); and A2 = {a1, a2, ...}. With N2 = E2 ∪ O2 ∪ A2, we then create L3 = L1 ∪ L2 in the usual way, by federation, as specified in figure 4.

We, the users of L3, now build a correspondence convention. Under this convention, which is outside of L3, we say x ⇒ y if we want x to correspond to y. Our first correspondence convention (there will be others) is that o1, o2, ... ⇒ P1, P2, ..., or in short o1 ⇒ P1. Formally and explicitly (and for future reference):

**Correspondence Convention 1.** o1 ⇒ P1

We require of all correspondence conventions what is evident here: that they specify a 1-1 relationship between name constants in L3 and wffs in L1. Corresponding to the first correspondence convention we have the first augmentation convention.

**Augmentation Convention 1.** Given a collection of sentences in Γ3, a user, s, of L3 may add to Γ3 an expression of the form Promise(e, s, o), provided e ∈ E2, o ∈ O2, Promise ∈ S2, and s ∈ A2, and provided that the name e does not occur in Γ3.

Further, if an expression of the form Promise(e, s, o) is added, so must an expression of the form Promise(e, s, o) → (Kept(e) ↔ P), where o ⇒ P is in Correspondence Convention 1, and Kept ∈ S2.

Because we require that e be a heretofore unused name, every added expression of the form Promise(e, s, o) will be unique. In a relational database context, e could serve as the key for the corresponding record and thereby guarantee the continued consistency of the database. Here, we need a stronger condition: That Γ3 ⊬ ¬ Promise(e, s, o). Of course, for any Γ and φ, if Γ ⊬ φ then Γ ∪ {φ} is consistent. We note, however, that Promise(e, s, o) is consistent, so Γ3 ⊬ ¬ Promise(e, s, o). Further, since e is not present in Γ3, so long as Γ3 is itself consistent and there is no occurrence of Promise(x, s, o) in Γ3 where x is governed by a universal quantifier, then Γ3 ⊬ ¬ Promise(e, s, o) and Γ3 ∪ {Promise(e, s, o)} will be consistent. It is entirely possible to manage the use of the language so that this obtains.

Finally, our first augmentation convention says that if we indeed add an expression of the form Promise(e, s, o) to Γ3, then we must also add a meaning postulate of the form

\[
\text{Promise}(e, s, o) \rightarrow (\text{Kept}(e) \leftrightarrow P)
\]

where o ⇒ P from our first correspondence convention, and Kept ∈ S2. Note that adding expression (6) to a consistent Γ3∪{Assert(e, s, o)} will produce a consistent set, if it is not the case the Kept(x) occurs in Γ3 where x is governed by a universal quantifier. Note as well that adding additional promises in this fashion will not generate an inconsistency. (If one addition doesn’t because these general conditions apply, neither will any finite number of such additions.)

With these conventions to hand, we can coherently express in L3 promises whose propositional contents are atomic sentences in L1. Extension to other speech acts and indeed to other subordinators is more or less straightforward. Thus, for s to assert that P we would add expressions of the form

\[
\text{Assert}(e, s, o)
\]

and

\[
\text{Assert}(e, s, o) \rightarrow (\text{Veridical}(e) \leftrightarrow P)
\]

where o ⇒ P from our first correspondence convention, and Veridical ∈ S2. Given that Γ1 is managed so that this addition process maintains consistency, we need not actually add the rule expressions (as in expressions (6) and (8)) to the database. Instead, we may abstract and use them as axiom schemas, e.g.,

\[
\text{Promise}(e, s, o_i) \rightarrow (\text{Kept}(e) \leftrightarrow P_i)
\]

subject to the usual qualifications (P_i ∈ S^1, etc.). Having taken this step it is useful to note that our (first) correspondence convention is a map from an arbitrary enumeration of certain names in O2 of L2 to an arbitrary enumeration of sentence letters in S1 of L1. Which enumerations we use are entirely immaterial, so we may as well use the expressions as their own enumerations. Thus, we can express the first correspondence convention as follows: [φ] ⇒ φ, for φ ∈ wffL3 and [φ] ∈ O2 (i.e., replace o1, o2, ... in O2 with [P1], [P2], ... for all wffs in L3). Since the wffs of L1 are denumerable, we can replace the first with the second correspondence convention:

**Correspondence Convention 2.** [φ] ⇒ φ, for [φ] ∈ O2 and φ ∈ wffL1.

Our augmentation convention is revised to reflect the new notation:

**Augmentation Convention 2.** Given a collection of sentences in Γ3, a user, s, of L3 may add to Γ3 an expression of the form Promise(e, s, [φ]), provided e ∈ E2, [φ] ∈ O2, Promise ∈ S2, and s ∈ A2, and provided that the name e does not occur in Γ3.

Further, if an expression of the form Promise(e, s, [φ]) is added, so must an expression of the form Promise(e, s, [φ]) → (Kept(e) ↔ [φ]), where [φ] ⇒ φ is in Correspondence Convention 2, and Kept ∈ S2.
Finally, given the apparatus above, there is no need to limit $L^1$ to sentence logic. So long as $L^1$ and $L^2$ are atomically disjoint we may create and operate a federated language, as sketched in this section. Nor is there need to limit the modeling to promises and assertions. But these are details best left for subsequent papers.

6. SUMMARY & DISCUSSION

In a federated language, $L^1$, every wff is either a wff of $L^1$, a wff of $L^2$, or a boolean combination of such wffs (possibly with some vacuous quantification mixed in). Because the federating languages are atomically disjoint, if $Γ^1$ is a consistent set of $L^1$ formulas and $Γ^2$ is a consistent set of $L^2$ formulas, then $Γ^3 = Γ^1 ∪ Γ^2$ is also consistent, as shown above. In the context of a federated language in which at least $L^2$ includes first order logic, or at least certain fragments of it, quotation and disquotation in the present sense is semantically unexceptionable and indeed unexotic. Quotation and disquotation are merely syntactic devices, to be used with governing conventions, for mapping an enumeration of names in $L^2$ to formulas in $L^1$. Quotation—in the sense developed here—is a function—think of a collision-free hash function—that returns the index value of an (implicit) enumeration on the domain of the function, here, wffs in a federating language. Disquotation—again in the present sense—is simply the inverse of quotation; it returns the domain item that has been quoted. Under the regime sketched here the quotation function is maintained as a convention governing the use of the federated language. This and sensible management of $Γ^3$ will permit logically accessible expressions of complex sentences while maintaining consistency.

It is interesting to compare the analysis offered here with the at most semi-formal account presented by the philosopher Donald Davidson in his perceptive essay, “On Saying That” [10]. Davidson’s views are concisely captured in the following passages, in which he focuses on “Galileo said that the earth moves” as a prototype for complex sentences (or as he calls it, indirect discourse).

The form ‘(∃x)(Galileo’s utterance $x$ and my utterance $y$ makes us samesayers)” is thus a way of attributing any saying I please to Galileo provided I find a way of replacing ‘$y$’ by a word or phrase that refers to an appropriate utterance of mine. And surely there is a way I can do this: I need only produce the required utterance and replace ‘$y$’ by a reference to it. Here goes:

The earth moves.

(∃x)(Galileo’s utterance $x$ and my last utterance makes us samesayers)

Definitional abbreviation is all that is needed to bring this little skit down to:

The earth moves.

Galileo said that.

Here the ‘that’ is a demonstrative singular term referring to an utterance (not a sentence).

From this point, Davidson finds it an innocuous step to reverse the order.

But in the present case nothing stands in the way of reversing the order of things, thus:

Galileo said that.

The earth moves.

And finally

It is now safe to allow a tiny orthographic change, a change without semantic significance, but suggesting to the eye the relation of introducer and introduced: we may suppress the stop after ‘that’ and the consequent capitalization:

Galileo said that the earth moves.

Davidson summarizes as follows:

The proposal then is this: sentences in indirect discourse [what we have called complex sentences], as it happens, wear their logical form on their sleeves (except for one small point). They consist of an expression referring to a speaker, the two-place predicate ‘said’, and a demonstrative referring to an utterance. Period.

The suggestion here is that if you are attracted to Davidson’s account and you want a full formalization (Davidson did not offer one), then the disquotation approach, as outlined in this essay, is one you should find congenial. The fundamental challenge to formalizing any account such as Davidson’s is the challenge of formalizing a demonstrative referring to an utterance” (and Davidson does not address the challenge). This is achieved, in the present account, by naming utterances (or rather their contents) and mapping the names to indices which afford automated recovery of the utterances. That, as noted, is but a suggestion. Unpacking it in detail is beyond the scope of this paper.

As advertised, this note is only a sketch of a concept and its development (the concept being the use of quotation as a device for encoding propositional content into a first order language). As always, and especially here, much remains to be done. I’ll conclude, then, with remarks on selected opportunities for further investigation.

Stronger federation. What has been presented and explored in this note is a minimal form of federation. Stronger forms may be possible and these may offer greater expressive power. For example, it is a requirement of our semantic evaluation functions for each federating language—the $V^i$‘s—that they be definable without reference to the other federating language. This may not be necessary. A move analogous to that made in Russell’s theory of types may work here: language $L^i$ may make reference to language $L^j$ for $i > j$. For example, we might permit $V^j(P^i_j) = V^j(P^i_j)$ as part of an interpretation of $L^2$, but not as part of the interpretation of $L^1$. Without risking a fall into inconsistency. Kaplan and Montague [22] make a similar suggestion, although they do not develop it, in their article demonstrating that certain very strong quotation languages fall into inconsistency. This theme—insolvency of quotation languages—is explored further in [28]. The point of the present paper is to demonstrate that weaker uses of quotation can be interesting without entailing inconsistency. How much these weaker
Iterated subordinators. One may safely say that all linguists and grammarians agree that subordinators may be iterated. The federation examples described here do not support iterated subordination. There is no way to express, for example, that Bob said that Carol promised that Ted requested that Alice be home by midnight. This is a limitation that needs to be addressed both for theoretical and for practical purposes. If we limit ourselves to a declared maximum depth of iteration, $D$, then it would appear that $L^{D+1} = L^3 \circ \ldots \circ L^D$, federating $D$ sub-languages, is a workable move. This would likely suffice for most computerized applications, provided the governing conventions can be worked out satisfactorily. Perhaps the depth of iteration need not be specified, provided the users of a system of languages obey a well-crafted convention, which would allow introduction of higher-order federation whenever required to express a deeper level of iteration.

Applications and reduction to practice. There are many topics of interest here. I note just two that particularly merit introduction of higher-order federation whenever required to express a deeper level of iteration.

Paraconsistency. In classical logics it is possible to prove anything from a contradiction. Letting $\vdash$ indicate a logical consequence relation, either semantic or proof-theoretic, it is said that $\vdash$ is explosive (for a given logic) if and only if for every wff $\phi$ and $\psi$,

$$\{\phi, \neg \phi\} \vdash \psi \tag{10}$$

Classical logics are thus explosive. Logics with non-explosive logical consequence relations are said to be paraconsistent. A number of such logics have been developed (see [30] for a review) and they constitute an active area of research. The trick in designing such logics, of course, is to find a principled way to contain an inconsistency. The problem is to allow certain inconsistencies without the entire system being explosive and collapsing into the pure calculus of irrelevance.

The federated languages discussed above are not paraconsistent. Recall, however, the discussion in §3 which pointed out that the following sequent is valid in $L^3$.

$$P_1^1 \circ (P_1^1 \rightarrow P_2^1), \quad (P_1^2 \rightarrow \neg P_2^2) \vdash (P_1^2 \land \neg P_2^2) \tag{14}$$

Because $P_1^2$ and $P_2^2$ may have the same intended meaning (viz., that it is raining) there is no impediment in principle to deriving a ‘contradiction’ of this sort. Might a federated language with disquotation be used to gain the benefit—or at least the effect—of paraconsistency without itself being inconsistent? Consider the following example. Let:

$$\Gamma^1 = \{B^1, \quad B^1 \rightarrow F^1\} \tag{11}$$

$$\Gamma^2 = \{B^2, (B^2 \land P^2) \rightarrow \neg F^2\} \quad \tag{12}$$

$$\Gamma^3 = \Gamma^1 \cup \Gamma^2 \quad \tag{13}$$

We set up $L^1$ and $L^2$ so that syntactically matched sentence letters have identical intended interpretations, e.g., $B^1$ and $B^2$ each are intended to mean that Tweety is a bird. Notice that $\Gamma^1 \not\vdash F^1$ and $\Gamma^3 \not\vdash F^2$. Now add to $\Gamma^3$ both $P^1$ and $P^2$. Then,

$$\Gamma^3 \cup \{P^1, P^2\} \not\vdash (F^1 \land \neg F^2) \tag{14}$$

This is the familiar Tweety example of defeasible reasoning and nonmonotonic logic, with $F^1$ for Tweety flies and $P^1$ for Tweety is a penguin.

What sense might we make of $(F^1 \land \neg F^2)$ and how might we use it? Recall the comment in §3 that we might think of $L^1$ and $L^2$ as different perspectives on a common subject. Consider by way of analogy a doctor with two different tests for diagnosing a disease. Test 1 will be positive with good probability if the patient has the disease. If the test is negative, then the patient may or may not have the disease. Test 2 is complementary. It falls positive with good probability if the patient does not have the disease, and if it is negative the patient may or may not have the disease. Neither test is entirely reliable. If they agree, are consistent in a broad sense, then very likely they give an accurate diagnosis. If test 1 is positive and test 2 negative, the patient very probably has the disease. If test 1 is negative and test 2 is positive, the patient probably does not have the disease. If both tests are negative no conclusion can be drawn by way of diagnosis. Finally, if both are positive we have the analog of expression (14). How do we choose to act on $(F^1 \land \neg F^2)$? Is a matter of policy. In the case of the Tweety example, a natural policy is to act on the assumption of $F^2$, since $L^2$ is more specific than $L^1$. In the case of a disease test, policy just might be based on evidence.

7. REFERENCES


