Semantics & Pragmatics Volume 15, Article 14: 1–51, 2022
https://doi.org/10.3765/sp.15.14

Belief or consequences *

Christopher Tancredi
Keio Institute of Cultural and Linguistic Studies

Yael Sharvit
University of California, Los Angeles

Submitted 2020-03-11 / First decision 2020-06-08 / Revision received 2020-12-16 / Second decision 2021-03-23 / Revision received 2021-08-22 / Third decision 2022-01-31 / Revision received 2022-04-10 / Accepted 2022-07-09 / Final version received 2022-07-31 / Published 2022-10-12 / Final typesetting 2024-02-18

Abstract We argue for a new inference-based analysis of belief attribution in which the embedded proposition is inferable from, but need not directly identify, an underlying belief of the subject's. The analysis accounts for attributions of belief in necessary truths and falsities, overcoming a major difficulty facing Hintikka 1962, and goes beyond Cresswell & von Stechow 1982 in accounting for intuitively valid inferable belief attributions. The analysis is based on a novel subjective I-semantics in which extensions depend dually on extension conditions assigned by a judge, and on the judge's beliefs about what satisfies those conditions. The interpretation of *believe* uses syntactic inference over logical formulas, with premises deriving from beliefs of both the attributor and the attributee, and the conclusion derived from the clause embedded under *believe*. Unlike nearly all prior analyses of belief attribution since Hintikka, our proposal makes no commitment to possible worlds while generating *de dicto*, *de re*, *de qualitate, de translato* and other interpretations, with the only formal semantic ambiguity deriving from what gets raised out of the embedded clause.

Keywords: inference, *de dicto, de re, de qualitate, de translato*, belief attribution, I-semantics

* We would like to thank Chris Barker, Irene Heim, Magda Kaufmann, Hazel Pearson, Orin Percus, two anonymous reviewers, and audiences at Keio, MIT, NYU, Tel Aviv and UConn for helpful comments.

1 Introduction

Since Hintikka 1962, standard analyses of belief attribution have all made essential use of possible worlds or situations. Indeed, one might justifiably think that possible worlds or situations are essential for analyzing belief attribution. We show that such thinking is misguided by developing an alternative without possible worlds that is empirically superior over a wide range.[1] The success of the analysis derives in part from a restricted syntactic notion of inference and in part from a novel semantics for I-languages (Chomsky 1986), an I-semantics. The analysis derives differences among historical *de dicto, de re, de qualitate* and *de translato* interpretations from differences in syntactic raising combined with differences in attributor and attributee beliefs.

The proposed analysis overcomes the longstanding problem of formalizing attributions of belief in necessary truths and necessary falsities, not only accounting for their dependence on circumstances but also explaining intuitively available inferences among such attributions. While we do not argue for the elimination of possible worlds from semantics, our analysis suggests such a possibility.

1.1 The core examples

We examine two core examples of mistaken beliefs about mathematical facts. In the first, the attributee is assumed to have correct concepts (according to the people evaluating the situation) for all terms involved, having demonstrated this understanding in class. However, she makes a regular mistake in calculation that shows up on her test answers.

(1) **Situation A**

Mary's answers on a test:

$2+2 = 1+4$

$1+4 = 7$

Mary does not have concepts for *real number* or *imaginary number*.

Situation A renders (2a,b) potentially true, but not (2c,d).

1 This covers everything except *de se* interpretation.

(2) a. Mary believes 2+2=5
b. Mary believes 2+2=7
c. #Mary believes i^i is an imaginary number[2]
d. #Mary believes i^i is a real number

In the second example, adapted from Tancredi & Sharvit 2020 (henceforth TS), the attributee is taken to have an incorrect concept for one of the terms involved but to make no mistake in calculation.

(3) **Situation B**
Hal takes the word *prime* to be true of a number if and only if that number equals $x^3 - 1$ for some natural number x. He asserts "26, which has exactly 4 factors, is prime", correctly calculating that $26 = 3^3 - 1$.

Situation B renders (4) potentially true, arguing for an inherent ambiguity in the first conjunct.[3]

(4) Hal believes 26 is prime, but he does not believe 26 is PRIME.[4,5]

Following TS, we label the interpretation that renders the first conjunct true *de translato*. In this paper we provide a uniform analysis of the word *believe* that accounts for these two cases. We then show that the analysis covers a wide range of other examples from the literature.

2 Previous analyses

To set the stage for our analysis, we review Hintikka 1962, Cresswell & von Stechow 1982 (henceforth CvS), and TS, Hintikka a lingua franca for the analysis and discussion of attitude attribution, CvS the first formal analysis capa-

2 Mathematically, i is the square root of negative 1, an imaginary number. i^i equals approximately 0.20788, a real number, not an imaginary number.

3 Belief attributions involving idiosyncratic attributee interpretation were brought to bear on the analysis of belief attribution as early as Burge 1978.

4 We use capitals to indicate intended focus, which is helpful, but neither necessary nor sufficient, for bringing out the interpretation intended. We do not analyze the pragmatic effects of focus.

5 *Believe* is a neg-raising verb, *a doesn't believe p* potentially conveying that *a believes not p*. While acknowledging this fact, we focus in (4) and throughout on interpretations with negation scoping over *believe*. This interpretation can be brought out unambiguously by replacing *he does not believe* with the unlyrical *it is not the case that he believes*, a replacement that does not affect the observed acceptabilities.

ble of distinguishing among attributions of belief in different necessities and impossibilities, and TS the first formal analysis of *de translato* interpretation.

2.1 Problems for Hintikka

Hintikka analyzes sentences of the form *a believes that p* as true iff the totality of *a*'s beliefs entail *p*. Both the totality of *a*'s beliefs and *p* are analyzed as sets of possible worlds, with the required entailment holding if the former set is a subset of the latter. The denotations of all expressions are taken to be fixed objectively in a language that is shared by all its speakers. Under that analysis, all individuals are predicted to believe every necessary truth since necessary truths – mathematical truths among them – denote the set of all worlds, making the truths of these examples independent of the particular beliefs the subject has, and incorrectly predicting that (2d) should be true despite Mary lacking the concept of real numbers.

Hintikka's analysis can account for the truth of (2a,b). Mary's answers on the exam in Situation A indicate that some of the things she believes are impossible, which for Hintikka means the set of worlds compatible with the totality of Mary's beliefs is the empty set. This renders (2a,b) true since the empty set is a subset of any set, including the empty set itself, denoted by the embedded clauses of those examples. Unfortunately, this same property renders (2c,d) equally true, and in fact will render any sentence of the form *Mary believes p* true regardless of what is substituted for *p*. It thus predicts all of the sentences in (2) to be equivalent given Mary's beliefs, contrary to observation.

Similar problems arise for (4). Since 26 is not prime, and necessarily so, (4) can only be true under Hintikka's analysis if Hal has impossible beliefs, wrongly leading once again to the prediction that Hal can truthfully be claimed to believe any proposition. This problem can be overcome by shifting to an I-language perspective and incorporating an analysis of *de translato* interpretation that interprets the word *prime* as Hal understands it, i.e. as denoting the property of being one less than a perfect cube, adapting the analysis of TS. However, such an interpretation will render *26 is prime*, *377,933,066 is prime* and an infinity of similar sentences necessary truths, incorrectly predicting it to be impossible for Hal not to believe them.

2.2 Problems for CvS

CvS analyze the object of belief as a structured proposition, a tuple consisting of an n-place relation ω and its n-arguments: $\langle \omega, b_1,\ldots,b_n \rangle$. This tuple is constructed from the clause embedded under *believe*, each element in the tuple an intension. By appealing to structured propositions, CvS can distinguish among attributions of belief in distinct necessary truths and falsities. By taking the elements in the tuple to be intensions, they furthermore account for a limited range of inference among belief attributions, one based on substitution under intensional identity. This substitution allows them to account for the truth of (2a), where Mary answers that 2+2=1+4 but is claimed to believe 2+2=5, since *2+2=1+4* is intensionally identical with *2+2=5* on structurings differing only in whether they contain *1+4* or *5* as one of the arguments, and Mary has given indirect evidence of believing the former proposition.

There are at least two problems with this explanation. First, it extends to (2c), predicting it to be equally true based on the same fact from Situation A. Two relevant structurings generate this conclusion: the null structurings in (5), with a zero-place relation and zero arguments, and the all-focus structurings in (6).

(5) a. $\langle \lambda$w. 2+2=1+4 in w$\rangle$
 b. $\langle \lambda$w. i^i is an imaginary number in w$\rangle$

(6) a. $\langle \lambda$wλp.p(w), λw. 2+2=1+4 in w$\rangle$
 b. $\langle \lambda$wλp.p(w), λw. i^i is an imaginary number in w$\rangle$

Under either of these structurings, the embedded clauses in (2a,c) denote the same proposition, and hence trivially entail each other. If (2a) is true under either of the structurings in (5a) or (6a), then, (2c) is predicted to be true as well under the corresponding structurings in (5b) and (6b), and vice versa. A parallel problem occurs with examples involving mathematical truths: the null structuring and the all-focus structuring render all attributions to an individual of beliefs in mathematical truths identical, rendering (2d) true if Mary believes any mathematical truth under a null- or all-focus structuring.

These unwanted consequences can be avoided by blocking null and all-focus structurings. Even stipulating such a restriction, however, the potential truth of (2b) in Situation A remains problematic.

(2b) Mary believes 2+2=7

The only beliefs Mary has indicated having in Situation A are that 2+2=1+4 and that 1+4=7. Neither is intensionally identical to 2+2=7 under any structurings other than null- and all-focus structurings. Since CvS provide no way of drawing inferences among belief attributions except through intensional identity, these beliefs fail to render (2b) true for CvS, in disagreement with intuition.

CvS also cannot directly account for the truth of (4). As TS show, however, shifting to an I-language perspective and adding *de translato* interpretation makes such an account possible. Hal's utterance of *26 is prime* under this analysis counts as evidence for the truth of the first conjunct of (4) with *prime* interpreted *de translato*. This modification of CvS can also account for the truth of (2b). The I-language perspective needed to interpret (4) as true treats interpretation as speaker relative. Mary's answers on the test show that she takes *2+2, 1+4* and *7* to be identical. If mathematical truths are necessary truths for Mary like they are for other speakers, *2+2, 1+4* and *7* as understood by Mary end up intensionally identical. Since she has given direct evidence for believing 1+4=7, and since Mary's *1+4* is intensionally identical to her *2+2*, it is possible to conclude she believes 2+2=7. All that is needed for the conclusion to go through is for her belief that 1+4=7 to be structured so that her *1+4* is one of the argument terms.

Unfortunately, while adding *de translato* interpretation to an I-language version of CvS makes it possible to analyze (2b) as true, the resulting analysis is unable to account for the truth of (2a). Under CvS's original analysis, what makes (2a) true is Mary's belief that 2+2=1+4 together with the objective fact that 1+4=5. However, under an I-language perspective, Mary's belief that 2+2=1+4 cannot be understood as an attitude involving shared-language expressions: the occurrences of *2+2* and *1+4* have to be understood as denoting what Mary takes them to denote, namely (her) *7*. An I-language perspective does make it necessary to assume as a default that different speakers associate identical intensions with expressions. However, given Mary's test answers, that default gets cancelled in this case: it cannot be assumed that either her *2+2* or her *1+4* are intensionally identical to the speaker's. It follows that Mary's *1+4* cannot be assumed to be intensionally identical either to Mary's *5* or to the speaker's *5*, leaving us with no interpretation of Mary's *2+2=1+4* that is intensionally identical to *2+2=5* regardless of how the propositions are structured or of whose interpretations are used in the latter proposition. This leaves CvS able to account for (2a) but not (2b) when understood as interpreting a shared language, or able to account for (2b) but not (2a) when understood as interpreting I-languages. There is no consistent understanding of CvS, however, that can account for both (2a) and (2b) at the same time.

3 Analysis

3.1 Informal overview

To account for examples like (4), we take the target of interpretation to be subjective I-languages rather than a shared objective language. We propose a new I-language-based framework for interpretation and a uniform analysis of *believe* that accounts for historical *de dicto, de re, de qualitate* and *de translato* interpretations among others, mathematical variants included. The framework relativizes translation and interpretation to an evaluator, typically a hearer but potentially any discourse participant or even an outside observer. It also employs subjective beliefs about extensions in place of objective extensions, and encodes these beliefs in the lexicon.

3.1.1 Inference

The leading idea of our analysis is that *a believes p* is true iff *p* can be inferred by the local judge *j* from *j*'s hypothetical acceptance of a belief of *a*'s.[6,7] The four main differences between our analysis and Hintikka's are (i) we employ a restricted syntactic notion of inference rather than a semantic notion of entailment, (ii) we use individual token beliefs in the inference, not the totality of a person's beliefs, (iii) for us, judges' beliefs can also serve as premises in the inference, affecting the truth conditions of belief attributions, and (iv) for us, even outside of belief attribution contexts, formal extensions are al-

6 All interpretation for us requires determining whose beliefs fix the extensions of expressions. For predicates of personal taste like *fun*, it can also involve determining, for example, who something is fun for. We believe both these roles to be related to the judge parameter of Lasersohn (2005) and Stephenson (2007), though we only officially commit to using that parameter for fixing extensions.

7 That the embedded clause in a belief attribution can fail to identify the content of a token belief was already acknowledged by Hintikka (1962) and forms the basis for many analyses of *de re, de qualitate* and *de translato* interpretations. That it never formally identifies the content of a belief was first put forward to our knowledge by Bach (1997).

ways subjectively dependent on the beliefs of a judge and an evaluator, not fixed objectively by the properties of an objective language.

We give our analysis of *a believes that p* informally below.

(7) a. *a believes that p* is true for an evaluator *ev* with respect to a judge *j* iff according to *ev*, *j* takes *a* to have internal beliefs *b* from the hypothetical acceptance of which *j* can infer *p*.

b. The inference can involve additional premises *q* based on *j*'s beliefs, taking the form:

hypothetical acceptance of *b*
acceptance of *q*
$\therefore$ *p*

Evaluation by an evaluator *ev* is always with respect to a judge *j*, though the value for *j* can vary. When *j* = the speaker, *ev*'s evaluation determines for *ev* what the speaker commits to. When *j* = *ev*, *ev*'s evaluation determines what *ev* would commit to by accepting what the speaker says. Other choices for *j* are formally possible but play limited pragmatic roles in discourse.

We illustrate the analysis with (2a,b) in Situation A. With the facts in Situation A presumed to be known to *ev*, for both *j* as speaker and *j* as hearer (=*ev*), *ev* will take *j* to take Mary to have the following two beliefs: *2+2=1+4* (=b_1), and *1+4=7* (=b_2). *ev* will also take *j* to have the belief *1+4=5* (=*q*). The inference that justifies (2a) is then:

(8) a. 2+2=1+4 (= *j*'s hypothetical acceptance of b_1)
b. 1+4=5 (= *j*'s acceptance of *q*)
c. $\therefore$ 2+2=5 (= *p*)

The inference that justifies (2b) is:

(9) a. 2+2=1+4 (= *j*'s hypothetical acceptance of b_1)
b. 1+4=7 (= *j*'s hypothetical acceptance of b_2)
c. $\therefore$ 2+2=7 (= *p*)

It is crucial that *j*'s acceptance of Mary's beliefs be hypothetical, since for both relevant values of *j*, *j* would likely reject (8a), (9a) and (9b).

3.1.2 Semantic framework (informal)

We propose a framework that relativizes semantic interpretation both to an evaluator, whose model is used, and to a judge, whose beliefs determine judge-relative subjective extensions of expressions (according to the evaluator). To implement this approach, we employ an evaluator model containing multiple constant assignment functions i that take an individual and an extension condition as arguments and return extensions, and multiple variable assignment functions g, but only a single world. The use of multiple constant assignment functions parallels the use of multiple possible worlds in a possible-worlds semantics, with different such functions encoding different ways the world could be.[8]

While syntactic expressions of English are assumed to be shared among speakers, their translations into logic and their interpretations in set theory can in principle vary from speaker to speaker. We distinguish two ways in which they can differ: (i) one expression can be associated with two different extension conditions by two different speakers, as with the expression *prime* in (4), or (ii) one common extension condition can be associated with two different believed extensions by two different speakers, as with the expression *equals* in (2b). The former distinction we model in the logical *translation* of an expression of English by associating the basic terms of the translation with distinct judge-determined extension conditions, pre-superscripted in (10) below. The latter distinction we model in the *interpretation* of an expression of logic by taking the extension to be dependent on the beliefs of the judge, subscripted in (10).

To illustrate the main properties of our framework we use $[\![\ ]\!]^{j}$ as our translation function, with j as judge, and $\|\ \|^{\mathrm{i}}$ as our interpretation function, where i is one of many functions used to assign values to constant-based expressions in a manner compatible with the subscripted judge's beliefs.

(10) a. $[\![\text{prime}]\!]^{\text{hal}} = {}^{\text{cube-1}}\text{prime}_{\text{hal}}$
$\|{}^{\text{cube-1}}\text{prime}_{\text{hal}}\|^{\mathrm{i}}$ = a function f determined by i that is true of a number x only if it is compatible with *hal*'s beliefs that x is 1 less than a perfect cube.

b. $[\![\text{prime}]\!]^{ev} = {}^{\text{2fac}}\text{prime}_{\text{ev}}$
$\|{}^{\text{2fac}}\text{prime}_{\text{ev}}\|^{\mathrm{i}}$ = a function f determined by i that is true of a number x only if it is compatible with *ev*'s beliefs that x has exactly two factors.

In both (10a) and (10b), what logical expression *prime* translates into depends on the judge superscript on the translation brackets. In particular, the super-

8 It would be possible in principle to use worlds to model believed extensions, though doing so would deprive worlds of their ontological status as possible.

scripted *hal* in (10a) determines both Hal's presumed extension conditions, *cube-1* and the subscript *hal* in the translation, while the superscripted *ev* in (10b) determines the evaluator's presumed extension conditions, *2fac*, and the subscript *ev*.[9] The function denoted by the logical translations depends on the constant assignment function *i*. Pragmatically, we take denotations to act as restrictions on $\langle i,g \rangle$ pairs. If the evaluator believes that 7 is prime and that 8 is not but is uncertain about whether 1729 is, all functions admitted by $\langle i,g \rangle$ pairs satisfying the restrictions in (10b) will be true of 7 and false of 8, but some such functions will be true of 1729 and others false of 1729. The fact that 1729 is in fact not prime is of no relevance, a fact that sets this framework apart from standard approaches to semantics that treat language objectively. It should be noted that the label *prime* in the above logical translations is strictly speaking superfluous. Nothing would be affected by analyzing the translations as ${}^{\text{cube-1}}{}_{\text{hal}}$ and ${}^{\text{2fac}}{}_{\text{ev}}$ respectively except for readability.[10]

We take beliefs, like sentences, to determine logical formulas. Beliefs, however, are not sentences. Translation of a linguistic expression into logic fixes the extension conditions as those assumed to be assigned by the relevant judge. It is not possible for translation of an expression of English to generate extension conditions on an expression of logic other than those of the associated judge. Beliefs we take to differ in this regard. I can have perfectly coherent beliefs that determine what extension Hal would assign, for example, to the extension conditions that I associate with the word *prime* even when I know my extension conditions differ from his.

While sentences get turned into logical formulas through translation, beliefs get turned into logical formulas through acceptance. We implement this by analyzing beliefs as functions from (logical or object language) expressions to logical formulas, with acceptance of a belief *b* by the judge denoted by *j* formalized as *b* applying to *j*. Formally, we analyze beliefs as Λ-expressions, with Λ-conversion resulting in substitution of the argument expression for all occurrences of the Λ-abstracted variable. In the application

9 If the evaluator is not aware of Hal's specific extension conditions but is aware of his utterance, the utterance can be seen as putting a restriction on possible extension conditions for the evaluator rather than contributing a specific one to the logical translation. We take such pragmatic restriction to be important not only for cases like (4) but for acquisition of concepts in general, but we ignore such complications here and throughout.

10 One could consider trying to eliminate extension conditions from the logic instead, recovering them when needed from the judge and label. We will see in Section 3.4.1, however, cases where that strategy would fail.

immediately below, applying a Λ-expression to a logical expression *x* results in *x* filling the role of judge, though we also make use of Λ-expressions in the translation of *believe* in (14) below. See Appendix 2 for formal details.

Consider Hal's utterance of *26 is prime* in Situation B. This utterance commits Hal to ${}^{cube\text{-}1}prime_{hal}\ ({}^{26}twenty\text{-}six_{hal})$ being true. If we take this commitment to derive from Hal's acceptance of an underlying belief of his, the underlying belief can be any of the following:

(11) a. $\Lambda z[{}^{cube\text{-}1}prime_{hal}\ ({}^{26}twenty\text{-}six_{hal})]$
b. $\Lambda z[{}^{cube\text{-}1}prime_{hal}\ ({}^{26}twenty\text{-}six_{z})]$
c. $\Lambda z[{}^{cube\text{-}1}prime_{z}\ ({}^{26}twenty\text{-}six_{hal})]$
d. $\Lambda z[{}^{cube\text{-}1}prime_{z}\ ({}^{26}twenty\text{-}six_{z})]$

Acceptance by Hal of any of these beliefs will result in the same formula that Hal committed to with his utterance. Acceptance by some other individual *j*, however, results in four distinct formulas.

In Section 3.1.1, Mary was claimed to have the belief b_1 that 2+2=1+4 based on one of her answers on a test. Given the above analysis, we now have to distinguish her public commitment from her underlying belief. Her commitment we take to be to the truth of (12), where the mathematical superscripts indicate that Mary assigns standard extension conditions to all of the terms:

(12) $${}^{=}\mathrm{equal}_{\mathrm{mary}}\ ({}^{+}\mathrm{plus}_{\mathrm{mary}}({}^{2}\mathrm{two}_{\mathrm{mary}})({}^{2}\mathrm{two}_{\mathrm{mary}}))\ ({}^{+}\mathrm{plus}_{\mathrm{mary}}({}^{1}\mathrm{one}_{\mathrm{mary}})({}^{4}\mathrm{four}_{\mathrm{mary}}))$$

Her underlying belief can be any belief of the form $\Lambda \mathrm{x}[\phi]$, where ϕ is generated from (12) by replacing zero or more occurrences of *mary* with *x*. Based on this analysis of beliefs and the logical formulas determined by their acceptance, we analyze inference by judge *j* syntactically in (11), where: the extension condition $\subseteq$ is the condition that holds of two arguments P and Q for a judge *j* iff the extension of P is a subset of the extension of Q for *j*; Premise 1 derives from acceptance of an attributee belief by the attributor; and Premise 2 derives from acceptance by the attributor of an attributee or attributor belief:

(13) **Inference**

Premise 1:	P(a)
Premise 2:	${}^{\subseteq}\mathrm{subset}_{\mathrm{j}}(\mathrm{P})(\mathrm{Q})$
Conclusion:	Q(a)

If the accepted beliefs that function as premises and the translation of the clause embedded under *believe* that functions as conclusion are fully λ-

converted formulas of the form *f(a₁)…(aₙ)*, the only way of directly splitting them into a single function *P/Q* and a single argument *a* is with *P/Q* being $f(a_1)\ldots(a_{n-1})$ and *a* being a_n. This turns out to be too restrictive for the inferences we need to admit. We overcome this limitation by allowing the premises and conclusions to be any expression of the logic that λ-converts to a relevant fully λ-converted accepted belief or embedded clause translation. This allows for an infinite number of ways of fitting a formula $f(a_1)\ldots(a_n)$ into the form *P/Q(a)*, including $\lambda x[f(x)\ldots(a_n)](a_1)$ and $\lambda P[P(a_1)](\lambda x[f(x)\ldots(a_n)])$, among others. The equivalence class of ϕ, abbreviated EC(ϕ), is the set of all formulas that λ-convert to the same formula as ϕ.

3.2 Applications I: Mary and Hal

We analyze the predicate *believe* as manipulating the judge parameter on the translation of its embedded clause, setting this parameter to the translation *s* of the subject of *believe*. This means that if nothing raises out of the embedded clause, the interpretations of all expressions of the embedded clause will only involve extensions compatible with the beliefs of (the extension of) *s*. An expression raised from the embedded clause to a higher clause, in contrast, will be translated and interpreted with respect to the judge of that higher clause.

We use *infer(A,p)* to indicate that a conclusion *p* is inferable from a set of premises *A*, where *p* and the elements of *A* are logical formulas, and the inference process is that in (13) above. We analyze *believe* as in (14), extending the domain of Λ-functions to include expressions of English as well as expressions of logic.

(14) **Translation:**[11]

$[\![\textit{believe}]\!]^{j} = \Lambda S[\lambda x_e[\text{believe}_j([\![S]\!]^{x})(x)]]$

Interpretation:

$\|\text{believe}_j(p)(a)\|^{i} = 1$ iff

$\exists b\ \exists q\ (\text{beliefs}_j(b,\|a\|^{i})\ \&\ \text{beliefs}_j(q,\|j\|^{i})\ \&\ \text{SUBSET}_j(q[j])$

$\&\ (\text{infer}_j(b[j]\cup q[j],\ p)\ \&\ \neg\text{infer}_j(q[j],\ p))$[12]

(For $b = \{b_1,\ldots,b_n\}$, $b[j] = \{b_1(j),\ldots,b_n(j)\}$)

Abbreviation:

$\|\text{believe}_j(p)(a)\|^{i} = 1$ iff INFER(a,j,p)

The interpretation of *believe* requires there to be beliefs *b* of the attributee's and beliefs *q* of *j*'s such that the attributed proposition *p* can be inferred from *j*'s acceptance of the beliefs in *b* and in *q*, with *b[j]* playing an ineliminable role in the inference and *q[j]* being restricted to *SUBSET* relations, i.e. to formulas of the form $^{\subseteq}subset_j(P)(Q)$. Based on the above assumptions, the structure of the sentence and formal analysis of the inference that render (2a) true are given in 15:

(15) $[5\ [1_e\ [\textit{Mary believes } 2{+}2{=}t_1]]]$

a. $\lambda P[P(^{+}plus_j(^{1}one_j)(^{4}four_j))]$
$(\lambda x[^{=}equal_{mary}(^{+}plus_{mary}(^{2}two_{mary})(^{2}two_{mary}))(x)])$ $(= b_{1a}(j))$

b. $^{\subseteq}subset_j(\lambda P[P(^{+}plus_j(^{1}one_j)(^{4}four_j))])(\lambda P[P(^{5}five_j)])$ $(= q(j))$

c. $\therefore\ \lambda P[P(^{5}five_j)](\lambda x[^{=}equal_{mary}(^{+}plus_{mary}(^{2}two_{mary})(^{2}two_{mary}))])$ $(= p)$

(15a) represents *j*'s hypothetical acceptance of an underlying belief consistent with Mary's public commitments, that the properties true of *j*'s 1+4 include Mary's property of equaling 2+2. (15b) represents *j*'s acceptance of *j*'s belief that whatever holds of 1+4 holds of 5 as *j* understands those expressions, and (15c) represents the translation of the clausal argument of *believes*: the properties true of *j*'s 5 include Mary's property of equaling 2+2. In (15a,c), the presented formula is a member of the relevant equivalence class in which the initial expression is type raised. Note that the syntactic raising of *5* results in *five* bearing the index *j* in (15c), the index assigned to expressions in the matrix clause. For the inference to conform to the pattern in (13), the occurrence of *five* in (15b) then has to also bear index *j*. This in turn makes it necessary for *plus, one* and *four* to bear index *j*, since nothing in the context ensures that whatever holds of 1+4 holds of *j*'s 5 when *1*, *+*, and/or *4* are understood as Mary understands them. The limitation to inferences having the form in (13) then makes it necessary for the indices on *plus, one* and *four* to all be *j* in (15a) as well.

(2b) comes out true when the embedded clause is interpreted entirely *in situ*, as in (16).

11 Note the essential use of ΛS as a function taking an object language expression as argument. This is what makes it possible to translate that argument with respect to a different judge parameter from that used to translate *believe*.

12 Thanks to Irene Heim for highlighting the need for the final conjunct.

(16) [*Mary believes 2+2=7*]

a. $\lambda P[P({}^{+}\text{plus}_{\text{mary}}({}^{1}\text{one}_{\text{mary}})({}^{4}\text{four}_{\text{mary}}))]$
$(\lambda x[{}^{=}\text{equal}_{\text{mary}}({}^{+}\text{plus}_{\text{mary}}({}^{2}\text{two}_{\text{mary}})({}^{2}\text{two}_{\text{mary}}))(x)])$ (= $b_{1b}(j)$)

b. ${}^{\subseteq}\text{subset}_{j}(\lambda P[P({}^{+}\text{plus}_{\text{mary}}({}^{1}\text{one}_{\text{mary}})({}^{4}\text{four}_{\text{mary}}))])(\lambda P[P({}^{7}\text{seven}_{\text{mary}})])$ (= $b_2(j)$)

c. $\therefore \lambda P[P({}^{7}\text{seven}_{\text{mary}})](\lambda x[{}^{=}\text{equal}_{\text{mary}}({}^{+}\text{plus}_{\text{mary}}({}^{2}\text{two}_{\text{mary}})({}^{2}\text{two}_{\text{mary}}))(x)])$ (= p)

The inference in (16) uses a belief about the extensions Mary is presumed to assign to *1+4* and *7*. This belief is justified if Mary is assumed to treat *equals* as transitive. In that case, if Mary's 1+4 equals, according to Mary, Mary's 7, then whatever Mary's 1+4 equals, Mary's 7 will also equal according to Mary. If Mary were not to treat *equals* as transitive, (16b) would not be justified, and no justified valid inference to the conclusion in (16c) would then be possible. That she does so then comes out as a commitment of the evaluator who understands (2b) to be true.

Finally, (4) comes out true when the first conjunct is analyzed without any raising, as in (17), and when *prime* in the second conjunct is raised to the matrix clause of that conjunct, as in (18). In all cases Hal is assumed to have the beliefs in (11) above, identified as b_a – b_d.

(17) [*Hal believes 26 is prime*]
Premise 1: ${}^{cube\text{-}1}prime_{hal}({}^{26}twenty\text{-}six_{hal})$ (= $b_a(j)$)
Conclusion: ${}^{cube\text{-}1}prime_{hal}\ ({}^{26}twenty\text{-}six_{hal})$ (= p)

(18) [*prime* [1_{et} [*Hal does not believe 26 is* t_1]]]
Target Conclusion: ${}^{2fac}prime_{j}\ ({}^{26}twenty\text{-}six_{hal})$

The inference in (17) is trivial, since *j*'s acceptance of b_a from (11) generates the same logical formula as the translation of *26 is prime* with *hal* as judge. Formally, the inference can be shoehorned into the pattern in (13) by adding as a second premise ${}^{\subseteq}subset_{j}\ ({}^{cube\text{-}1}prime_{hal})\ ({}^{cube\text{-}1}prime_{hal})$, or alternatively we can accept trivial inferences as an additional acceptable inference pattern. In (18), raising of *prime* results in its logical translation bearing the superscript *2fac* and the subscript *j*, on the assumption that the judge has the correct concept of what it is to be prime. (18) comes out true because there is no inference from accepted beliefs of Hal's and subset-based beliefs of *j*'s to the conclusion ${}^{2fac}prime_{j}\ ({}^{26}twenty\text{-}six_{hal})$. Using any of the beliefs in (11) to generate the first premise will fail to generate a true inference to this conclusion since Situation B makes it clear that ${}^{cube\text{-}1}\textit{prime}_{j/hal}$ and ${}^{2fac}\textit{prime}_{j}$ do not stand in a subset relation: the former is true of 26, but the latter is not. Using the only other beliefs of Hal's indicated in Situation B to generate the first premise will lead to a premise of the form ${}^{4fac}\textit{has-4-factors}_{j/hal}$ $({}^{26}\textit{twenty-six}_{j/hal})$. However, on the assumption that a number with 4 factors cannot also have exactly 2 factors even for Hal, none of these potential premises gives rise to a true inference to the desired conclusion since ${}^{\subseteq}\textit{subset}_{j}$ $({}^{4fac}\textit{has-4-factors}_{j/hal})$ $({}^{2fac}\textit{prime}_{j})$ is false.

3.3 Formal detail

In this section we spell out the formal detail needed to support the informal analysis of Section 3.2. This includes incorporating variables into translations, adding a rule of predicate abstraction, formalizing the interpretation of logical expressions, and connecting utterances to discourse notions of truth, falsity, common ground and (dis)agreement. A summary of the formal analysis can be found in Appendix 2.

3.3.1 Translation

To determine what variables indexed expressions get translated into, we add a variable determination function h to our **translation function** $[\![\ ]\!]^{j,h}$, where j is the judge parameter that gets attached to basic expressions of the logic as a subscript, as before. Evaluation typically focuses on two fixed values for j: j=the evaluator, *ev,* and j=the speaker according to *ev*, sp_{ev}. We illustrate translation with the sentence *The*$_3$ *man smiles*, where a colon is used to separate at-issue content from presuppositional content.

(19) **Translation**

$[\![\textit{smiles}]\!]^{j,h} = \lambda x[{}^{\alpha}\text{smile}_j(x)]$
$[\![\textit{the}_3]\!]^{j,h} = \lambda Q[h(3) : Q(h(3))\ {}^{\beta}\text{and}_j\ {}^{\delta}\text{unique}_j(Q)]$
$[\![\textit{man}]\!]^{j,h} = \lambda x[{}^{\gamma}\text{man}_j(x)]$
$[\![\textit{the}_3\ \textit{man}]\!]^{j,h}$
$= h(3) : {}^{\gamma}\text{man}_j(h(3))\ {}^{\beta}\text{and}_j\ {}^{\delta}\text{unique}_j({}^{\gamma}\text{man}_j)$
$= x_3 : {}^{\gamma}\text{man}_j(x_3)\ {}^{\beta}\text{and}_j\ {}^{\delta}\text{unique}_j({}^{\gamma}\text{man}_j)$
abbreviated x_{3,man_j}

$[\![\mathit{The}_3\ \mathit{man\ smiles}]\!]^{\mathrm{j,h}}$
$= \lambda \mathrm{x}[{}^{\alpha}\mathrm{smile}_{\mathrm{j}}(\mathrm{x})](\mathrm{x}_3 : {}^{\gamma}\mathrm{man}_{\mathrm{j}}(\mathrm{x}_3)\ {}^{\beta}\mathrm{and}_{\mathrm{j}}\ {}^{\delta}\mathrm{unique}_{\mathrm{j}}({}^{\gamma}\mathrm{man}_{\mathrm{j}}))$
$= {}^{\alpha}\mathrm{smile}_{\mathrm{j}}(\mathrm{x}_3 : {}^{\gamma}\mathrm{man}_{\mathrm{j}}(\mathrm{x}_3)\ {}^{\beta}\mathrm{and}_{\mathrm{j}}\ {}^{\delta}\mathrm{unique}_{\mathrm{j}}({}^{\gamma}\mathrm{man}_{\mathrm{j}}))$
abbreviated $\mathit{smile}_j(x_{3,man_j})$, with extension conditions overt only when distinctive.

3.3.2 Variables and raising

We follow Heim & Kratzer (1998) in our syntactic analysis of raising, schematized in (20), where *exp* is the expression raised, *i* an index (a number), t_i the trace left behind, and σ the semantic type of t_i.[13]

(20) [*exp* [i_σ [...t_i ...]]]

We translate indexed expressions like definite determiners, pronouns and traces into logical expressions containing variables, and expressions of the form *[*i_σ *E]* into λ-functions. We manipulate the variable determination function *h* to accomplish these latter translations.

(21) Translation of indexed expressions:
If α is of type *e*, $[\![\alpha_i]\!]^{\mathrm{j,h}} = \mathrm{h}(i) : \mathrm{Q}(\mathrm{h}(i))$, where Q is the presupposition associated with α, if any.
If α is of type $\langle\langle e,t\rangle,e\rangle$, $[\![\alpha_i]\!]^{\mathrm{j,h}} = \lambda\mathrm{P}[\mathrm{h}(i) : \mathrm{P}(\mathrm{h}(i))\ {}^{\&}\mathrm{and}_{\mathrm{j}}\ \mathrm{Q}(\mathrm{h}(i))]$, where Q is the presupposition associated with α, if any, and & is the extension condition *j* associates with the word *and*.

(22) Predicate Abstraction:[14]
For any natural number *i*, type σ, individual *j*, variable determination function *h*, and expression *E*:
$[\![i_\sigma\ E]\!]^{\mathrm{j,h}} = \lambda x_\sigma\ [[\![E]\!]^{\mathrm{j,h}[x/i]}]$,
where $h[x/i]$ is just like h except at most that $h[x/i](i) = x$.

13 We allow two types of raising: scopeless raising where *exp* is of type σ, generalizing von Fintel & Heim's (1997–2021) analysis of scopeless raising of quantifiers, and scope-changing raising appropriate for QR, where the type of *exp* is $\langle\langle\sigma,\tau\rangle,\tau\rangle$, for some type τ. Scopeless raising allows quantifiers to raise into a higher clause without taking scope in that higher clause, a possibility exploited briefly in Section 3.5.

3.3.3 Interpretation

Interpretation is accomplished with the interpretation function $\| \|^{M_{ev},i,g}$, which generates interpretations relative to a model M_{ev} (= the evaluator *ev*'s model), a constant assignment function *i*, and a variable assignment function *g*. A model M is a tuple ⟨D,I,G⟩, where D is the set of domains D_σ for any type σ, I the set of constant assignment functions, and G the set of variable assignment functions. For any expression *a* of the logic, any model M, any i∈I and any g∈G, $\|a\|^{M,i,g}$ is the extension of *a* relative to M, i and g. Though we assume different evaluators employ different models, each evaluator *ev* employs only their own model M_{ev} for evaluation.

If *a* is a constant, ϕ a condition on its extension, and *j* a type *e* expression, $\|{}^{\phi}a_j\|^{M_{ev},i,g}$ is defined only if it is consistent with $\|j\|^{M_{ev},i,g}$'s beliefs that $i(\langle\|j\|^{M_{ev},i,g},\phi\rangle)$ satisfies ϕ, as evaluated by *ev*. If defined, $\|{}^{\phi}a_j\|^{M_{ev},i,g} = i(\langle\|j\|^{M_{ev},i,g},\phi\rangle)$.[15] More informally, $\|{}^{\phi}a_j\|^{M_{ev},i,g}$ presupposes that its extension is believed by *j* to possibly satisfy ϕ, and denotes that extension iff this presupposition is satisfied. Note that this definition allows different values for *i* to determine different extensions in cases where *j* is uncertain about what satisfies ϕ.

If *a* is a variable with index *i*, $\|a\|^{M_{ev},i,g} = g(i)$.

The logical symbol *ev* designates the evaluator: for any *x*, $\|ev\|^{M_x,i,g} = x$. This is the only basic expression of the logic that can be interpreted without a subscript or superscript. In all other cases, for $\|j\|^{M_{ev},i,g}$ to be defined *j* has to be an expression with a subscript ending in *ev*, such as *ev*, sp_{ev}, $mary_{sp_{ev}}$, etc. Where no confusion is likely to arise, however, we omit subscripts on subscripts.

If *a* is of the form $\alpha(\beta)$, then $\|a\|^{M_{ev},i,g} = \|\alpha\|^{M_{ev},i,g}(\|\beta\|^{M_{ev},i,g})$ or $\|\alpha\|^{M_{ev},i,g}(\beta)$, whichever is defined.

We illustrate below:[16,17]

14 Predicate Abstraction could equally be implemented in the interpretation by translating a bare index in the object language as a bare index in the logic and interpreting the result as identical to λ-abstraction over variables bearing that index.

15 In a semantics without extension conditions represented overtly in the logic, the same result can be obtained by analyzing $\|a_j\|^{M_{ev},i,g}$ as $i(\langle\|j\|^{M_{ev},i,g}, exc(a)(\|j\|^{M_{ev},i,g})\rangle)$, where $exc(a)(\|j\|^{M_{ev},i,g})$ is the extension condition that $\|j\|^{M_{ev},i,g}$ associates with *a*.

16 We suppress the subscript on *M* here and throughout when not relevant to the discussion.

17 Since extension conditions cannot be formalized as linguistic or logical expressions without giving rise to an infinite regress, we analyze them as abstract mental concepts. While this analysis leaves the nature of abstract mental concepts unanalyzed, it arguably improves on

(23) $\|[\![\textit{The}_3\ \textit{man smiles}]\!]^{j,h}\|^{M,i,g} =$
$\|{}^{\alpha}\text{smile}_j(x_3 : {}^{\gamma}\text{man}_j(x_3)\ {}^{\beta}\text{and}_j\ {}^{\delta}\text{unique}_j({}^{\gamma}\text{man}_j))\|^{M,i,g} = 1$ iff
$\|{}^{\alpha}\text{smile}_j\|^{M,i,g}\ (\|x_3 : {}^{\gamma}\text{man}_j(x_3)\ {}^{\beta}\text{and}_j\ {}^{\delta}\text{unique}_j({}^{\gamma}\text{man}_j)\|^{M,i,g}) = 1$ iff
$\|{}^{\alpha}\text{smile}_j\|^{M,i,g}\ (\|x_3\|^{M,i,g}) = 1$:
presupposition: $\|{}^{\gamma}\text{manj}(x_3)\ {}^{\beta}\text{and}_j\ {}^{\delta}\text{unique}_j({}^{\gamma}\text{man}_j)\|^{M,i,g} = 1$ iff
$i(\langle \|j\|^{M,i,g}, \alpha\rangle)\ (g(3)) = 1$:
presuppositions:
$i(\langle \|j\|^{M,i,g},\beta\rangle)\ (i(\langle \|j\|^{M,i,g},\gamma\rangle)\ (g(3)),\ i(\langle \|j\|^{M,i,g},\delta\rangle)\ (i(\langle \|j\|^{M,i,g},\gamma\rangle))) = 1,$
$\forall \phi \in \{\alpha,\beta,\gamma,\delta\}$, $i(\langle \|j\|^{M,i,g},\phi\rangle)$ satisfying ϕ is consistent with $\|j\|^{M,i,g}$'s beliefs

3.3.4 Discourse

In this section we show how to define pragmatic notions of truth and falsity, how to define common ground, and how to model disagreement in the proposed system. For those readers interested only in applications of the analysis to belief attribution sentences, this section can be skipped.

3.3.4.1 Truth and falsity

In the proposed semantic framework, the truth value of a declarative sentence depends on a valuation $\langle i,g\rangle$, an ordered pair consisting of a constant assignment function i and a variable assignment function g. The framework does not recognize any valuation as objectively correct, however, so intuitive truth cannot be defined relative to such a valuation. This sets our framework apart from standard possible-worlds frameworks where one world is singled out as actual and the intuitive truth of a sentence is equated with its objective semantic truth at that world.

We capture the intuitive notions of truth and falsity pragmatically, based on the belief valuations bv_x for an individual x, defined as the set of all valuations compatible with x's beliefs. Belief valuations for us are parallel to the notion of belief worlds in a possible-worlds semantics. Based on belief valuations, we define the following three-way distinction in pragmatic truth values:

Tarski's Convention T by grounding all extension conditions, which ultimately determine truth conditions, in belief.

(24) For an evaluator *ev*, a sentence S is:
true for j iff $\|[\![S]\!]^{j,h}\|^{M_{ev},i,g} = 1$ for every $\langle i,g\rangle$ in bv_{ev},
false for j iff $\|[\![S]\!]^{j,h}\|^{M_{ev},i,g} = 0$ for every $\langle i,g\rangle$ in bv_{ev}, and
of undetermined truth value for j otherwise.

ev's acceptance that a speaker who utters S believes what they say restricts bv_{ev} to valuations that render S true for the speaker, while *ev*'s acceptance of S restricts bv_{ev} to valuations that render S true for *ev*.

3.3.4.2 Common Ground

Stalnaker (2002) gives the following characterization of Common Ground:

> It is common ground that ϕ in a group if all members *accept* (for the purpose of the conversation) that ϕ, and all *believe* that all accept that ϕ, and all *believe* that all *believe* that all accept that ϕ, etc.

Within our proposed framework, ϕ above is best understood as a belief, that is as a function from judges (or more accurately judge-denoting expressions) to logical formulas. Intuitive beliefs about extensions we encode in belief valuations based on the evaluator's lexicon LEX, a recursive collection of sub-lexicons with a different sub-lexicon LEX_j for every judge j. The lexicon is composed of lexical items of the form $\langle {}^{\phi}exp_j, Z\rangle$, where ${}^{\phi}exp_j$ is a basic lexical term, and Z is the set of all elements that potentially satisfy ϕ according to j. We add to our lexical-term inventory terms of the form ${}^{\phi}bel_j$, where for any formula b and individual-denoting expression x, ${}^{\phi}bel_j(x)(b)$ is true iff b is an internal belief of the denotation of x's according to j.[18]

For lexical items of the form $\langle {}^{\phi}bel_j, Z\rangle$, we take Z to be a collection of lexical items that determine the beliefs of j. If j is uncertain whether Mary is a or b according to the evaluator, the lexicon will include $\langle {}^{\Psi}mary_j, \{a,b\}\rangle$. If k has the internal belief that j is certain that Mary is c according to the evaluator, the lexicon will also include $\langle {}^{\phi}bel_k, Z\rangle$, where Z contains $\langle {}^{\Psi}mary_j, \{c\}\rangle$. If l takes k to take j to think that a and b are spies and that c and d are not but to be unsure about other individuals, the lexicon will include $\langle {}^{\phi}bel_l, Z\rangle$,

18 The alternative of basing "all believe that …" on the analysis of *believe* presented in this paper would have the unwanted consequence of a proposition p counting as common ground when all speakers accept p even if some speaker thinks another does not accept p but does accept q from which that speaker can infer p.

with Z containing $\langle {}^{\phi}bel_k, Y\rangle$, Y containing $\langle {}^{X}spy_j, X\rangle$, and X being the set of functions all of which are true of a and b and all of which are false of c and d, but which differ from one another in which other individuals they are true of. This lexicon can then be used to place restrictions on the evaluator's belief valuations bv_{ev} by requiring every valuation $\langle i,g\rangle$ in bv_{ev} to assign to each lexical item in *ev*'s lexicon one of the values in the set it contains. Formally:

(25) An expression of the form ${}^{\phi}exp_j$ is a **lexical term**.

A **lexical item** is a pair $\langle \alpha, Z\rangle$, where α is a lexical term and Z is a set (of possible extensions).

The **lexicon** LEX is the set of all lexical items.

A lexical item $\langle \alpha, Z\rangle$ encodes beliefs about extensions as **belief valuations for the lexical term** α: $\mathrm{BV}_{\alpha} = \{\langle i,g\rangle : \|\alpha\|^{\mathrm{M,i,g}} \in Z\}$.

For complex expressions of the form $\alpha(\beta)$, their belief valuations are the intersection of the belief valuations for their parts: $\mathrm{BV}_{\alpha(\beta)} = \mathrm{BV}_{\alpha} \cap \mathrm{BV}_{\beta}$.

The **belief valuations of an individual** j, bv_j, can be derived by intersecting the belief valuations of all members of LEX_j, the set of all lexical items whose first element is a lexical term with subscript j.

We illustrate the restrictions imposed on bv_j with the sentence *Mary is a spy* evaluated with *j* as judge and *ev* as evaluator:

(26) Abbreviation:
$\mathrm{LEX}_{j,\alpha}$ = that tuple in LEX_j of the form $\langle \alpha, Z\rangle$ for some Z.
(= *j*'s lexical item based on α)

Mary:
j's beliefs: *mary* denotes a or b
$\mathrm{LEX}_{j,{}^{\phi}mary_j}$ = $\langle {}^{\phi}mary_j, \{a,b\}\rangle$
$\mathrm{BV}_{{}^{\phi}mary_j}$ = $\{\langle \mathrm{i,g}\rangle : \|{}^{\phi}mary_j\|^{\mathrm{M,i,g}} \in \{a,b\}\}$
Restriction: $bv_j \subseteq \mathrm{BV}_{{}^{\phi}mary_j}$

(is a) spy:[19]

j's beliefs: *spy* is true of *a* and *b*, false of *c* and *d*, and of unknown value for all other individuals.

Define: For any set A of elements of type σ:

T(A) $= \{f_{\langle\sigma,t\rangle}: \{x: f(x)=1\} \supseteq A\}$

(= the set of functions true of all members of A)

F(A) $= \{f_{\langle\sigma,t\rangle}: \{x: f(x)=1\} \cap A = \emptyset\}$

(= the set of functions false of all members of A)

$\mathrm{LEX}_{j,{}^{\Psi}spy_j} = \langle {}^{\Psi}spy_j, \mathrm{T}(\{a,b\}) \cap \mathrm{F}(\{c,d\})\rangle$

$\mathrm{BV}_{{}^{\Psi}spy_j} = \{\langle i,g\rangle: \|{}^{\Psi}spy_j\|^{M_{ev},i,g} \in (\mathrm{T}(\{a,b\}) \cap \mathrm{F}(\{c,d\}))\}$

Restriction: $bv_j \subseteq \mathrm{BV}_{{}^{\Psi}spy_j}$

Mary is a spy:

j's beliefs: as above

$\mathrm{LEX}_{j,\,{}^{\Psi}spy_j({}^{\phi}mary_j)}$ = undefined

$\mathrm{BV}_{j,\,{}^{\Psi}spy_j({}^{\phi}mary_j)} = \mathrm{BV}_{{}^{\phi}mary_j} \cap \mathrm{BV}_{{}^{\Psi}spy_j}$

Restriction: $bv_j \subseteq \mathrm{BV}_{j,\,{}^{\Psi}spy_j({}^{\phi}mary_j)}$

In the illustration above, the sentence *Mary is a spy* is true for *j* for *ev* because for every valuation $\langle i,g\rangle$ in bv_j, $\|{}^{\phi}mary_j\|^{M_{ev},i,g} \in \|{}^{\Psi}spy_j\|^{M_{ev},i,g}$. The same sentence could come out true, false or of undetermined truth value for a distinct judge *k* and/or evaluator ev_2.

We define mutual beliefs $\mathrm{MB}_{A,ev}$ and common ground $\mathrm{CG}_{A,ev}$ for a set of individuals A for an evaluator *ev* as follows:

(27) **Mutual Beliefs**

$$\mathrm{MB}_{A,ev} = \{b: b \text{ is a belief } \& \; \forall y{\in}A\, (\exists k\, (\forall \langle i,g\rangle{\in}\mathrm{BV}_{b(k)}\, (\|k\|^{M_{ev},i,g} = y \;\&\; \|b(k)\|^{M_{ev},i,g} = 1)))\}$$

Common Ground:

$$\mathrm{CG}_{A,ev} = \text{the largest subset X of } \mathrm{MB}_{A,ev} \text{ such that } \forall b{\in}\mathrm{MB}_{A,ev}\, (b{\in}X \rightarrow \Lambda x[{}^{\phi}bel_k(k)(b(x))]{\in}X)$$

By this definition, $\mathrm{CG}_{A,ev}$ consists of that subset of mutual beliefs that are mutually believed to be shared by the members of A, mutually believed to be mutually believed to be shared by the members of A, etc.

19 Here and throughout, we ignore the contributions to translation and interpretation of *is* and *a* in predicational NPs.

3.3.4.3 Disagreement

Two individuals *j* and *k* disagree about a sentence *S* according to an evaluator *ev* when for *ev*, *S* is true for *j* and false for *k* or vice versa. Since the judge ends up as a subscript on the basic terms of logical translations and determines the extension conditions associated with basic terms, $[\![S]\!]^{\mathrm{j,h}}$ and $[\![S]\!]^{\mathrm{k,h}}$ are formally distinct logical formulas with potentially distinct truth conditions. What makes opposite truth values disagreement rather than mere difference is the default expectation of an evaluator that discourse participants agree on – and should agree on – what *S* is saying, together with the recognition that opposite truth values could not be generated if both *j* and *k* had only correct beliefs (according to the evaluator) about the relevant extensions. If *ev* takes *j* and *k* to associate the same extension conditions with the basic terms in the translations of *S* but to have incompatible beliefs about the extensions of one or more terms, they disagree about the facts of the world given an agreed-on understanding of the language, according to *ev*. If *ev* takes them to associate different extension conditions with the basic terms in the translation, they disagree about the language according to *ev*.

We illustrate disagreement about facts and common ground update with a simple dialogue:

(28) C: Mary is a spy.
B: I see. I didn't know that.

We consider the first sentence evaluated from C's perspective with respect to each of B and C as judge, where in C's model M_c, $D_e = \{a,b,c,d,e,f\}$, C takes *B* to be a name of *b* for both speakers, and C identifies herself as *c*. If C believes the individuals who are spies include *a* and *b* but not *c* or *d*, and she believes B thinks they include *b* and *f* but not *a* or *c*, then all valuations $\langle i,g \rangle$ in bv_C will satisfy the following:[20]

(29) a. $\|spy_{ev}\|^{\mathrm{M}_c,\mathrm{i},\mathrm{g}}$ is a function true of all and only the members of $\{a,b\}$, of $\{a,b,e\}$, of $\{a,b,f\}$, or of $\{a,b,e,f\}$, that is, for one of the above sets as A, $\|spy_{ev}\|^{\mathrm{M}_c,\mathrm{i},\mathrm{g}}(x) = 1$ for all and only values of *x* taken from A.

b. $\|spy_B\|^{\mathrm{M}_c,\mathrm{i},\mathrm{g}}$, is a function true of all and only the members of $\{b,f\}$, of $\{b,d,f\}$, of $\{b,e,f\}$, or of $\{b,d,e,f\}$.

20 Here and below we suppress extension conditions taken to be shared by all relevant individuals.

Suppose that according to C, both B and C identify Mary as *a*. In this situation, C takes herself to be speaking truthfully, since for every $\langle i,g\rangle$ in bv_C, $\|spy_{ev}(mary_{ev})\|^{M_c,i,g} = 1$. That is, *Mary is a spy* is true for C according to C. This is so since for all such $\langle i,g\rangle$, $\|spy_{ev}\|^{M_c,i,g}$ is a function that is true of *a*. At the same time, C takes B to judge what she says as false, since for every $\langle i,g\rangle$ in bv_C, $\|spy_B(mary_B)\|^{M_c,i,g} = 0$: for all such $\langle i,g\rangle$, $\|spy_B\|^{M_c,i,g}$ is a function that is false of *a*. That is, *Mary is a spy* is false for B according to C. Thus, C evaluates B and C as initially disagreeing about whether Mary is a spy.

When B accepts C's utterance, it requires C to change the valuations in bv_C in such a way that for every $\langle i,g\rangle$ in bv_C, $\|spy_B(mary_B)\|^{M_c,i,g} = 1$. This can be done minimally by changing BV_{C,spy_B}, or by changing $BV_{C,mary_B}$ in any one of the following ways:

(30) Old BV_{C,spy_B} $\{\langle i,g\rangle: \|spy_B\|^{M_c,i,g}$ is true of *b* and *f* and false of *a* and *c*$\}$
New BV_{C,spy_B} $\{\langle i,g\rangle: \|spy_B\|^{M_c,i,g}$ is true of *a, b* and *f* and false of *c*$\}$

Old $BV_{C,mary_B}$ $\{\langle i,g\rangle: \|mary_B\|^{M_c,i,g} \in \{a\}\}$
New $BV_{C,mary_B}$ $\{\langle i,g\rangle: \|mary_B\|^{M_c,i,g} \in \{b\}\}$, OR
$\{\langle i,g\rangle: \|mary_B\|^{M_c,i,g} \in \{f\}\}$, OR
$\{\langle i,g\rangle: \|mary_B\|^{M_c,i,g} \in \{b,f\}\}$

The first change constitutes a change in who, according to C, B takes the spies to be, with *a* added to their ranks, while the other changes constitute a change in who, according to C, B takes Mary to be: definitely *b*, definitely *f*, or one of *b* or *f* without being sure which. Greater changes would be consistent with B's accepting C's utterance but would not be motivated by that acceptance in a typical context. Parallel analyses apply to B's evaluation of the same utterance.

3.3.5 Subsets

A note is in order regarding the ${}^{\subseteq}subset_j$ relation employed in the inference relation encoded in the logical translation of *believe* in (14) from Section 3.2. Suppose that for a given evaluator *ev*, for all valuations $\langle i,g\rangle \in bv_{ev}$, $\|{}^{\phi}spy_{mary}\|^{M_{ev},i,g}$ and $\|{}^{\phi}spy_{bill}\|^{M_{ev},i,g}$ are both functions true of all and only the members of $\{a\}$, or of $\{a,b\}$, that is, according to *ev*, both Bill and Mary believe *a* is a spy, they are both unsure about whether *b* is a spy, and they both think that no one else is a spy. Somewhat counterintuitively, in this situation ${}^{\subseteq}subset_j({}^{\phi}spy_{mary})({}^{\phi}spy_{bill})$ is false for *ev*. There are 4 relevant classes of constant assignment functions to consider that are contained in some valuation in bv_{ev}:

i:	$i_1(\langle \|mary\|^{M_{ev},i_1,g}, \phi\rangle) = \{a\}$,	$i_1(\langle \|bill\|^{M_{ev},i_1,g}, \phi\rangle) = \{a\}$
ii:	$i_2(\langle \|mary\|^{M_{ev},i_2,g}, \phi\rangle) = \{a\}$,	$i_2(\langle \|bill\|^{M_{ev},i_2,g}, \phi\rangle) = \{a,b\}$
iii:	$i_3(\langle \|mary\|^{M_{ev},i_3,g}, \phi\rangle) = \{a,b\}$,	$i_3(\langle \|bill\|^{M_{ev},i_3,g}, \phi\rangle) = \{a\}$
iv:	$i_4(\langle \|mary\|^{M_{ev},i_4,g}, \phi\rangle) = \{a,b\}$,	$i_4(\langle \|bill\|^{M_{ev},i_4,g}, \phi\rangle) = \{a,b\}$

The required subset relation fails to hold in case (iii). This case could be eliminated if Mary and Bill's beliefs are believed to be interdependent. For example, *ev* could believe that for any individual *z* such that Mary is uncertain about whether *z* is a spy, Bill has the belief that *z* is a spy. This would eliminate cases (i) and (iii). Alternatively, *ev* could believe that Mary's and Bill's thoughts about spyhood are entangled in the sense of quantum physics in such a way that their uncertainties about any individual's being a spy are always resolved in the same way. This would eliminate cases (ii) and (iii). Though unusual, these are not impossible beliefs for *ev* to have. However, if *ev* lacks such unusual beliefs, $^{\subseteq}subset_j(x_k)(x_l)$ will typically be false for *ev* when $k \neq l$.

3.4 Applications II: Issues from the literature

In this section we show how our analysis deals with three issues that are prevalent in the literature on belief attribution: acquaintance, transparency, and the relation of the embedded clause to an attributee's internal beliefs.

3.4.1 Acquaintance as specificity

Starting from Kaplan 1968, it has become commonplace to recognize *de re* interpretation as a separate kind of semantic interpretation that relies on the notion of acquaintance: for a sentence of the form *a believes that P(b)* to be true with *b* understood *de re*, (the denotation of) *a* has to be acquainted with (the denotation of) *b*. (See Appendix 1 for details.) Our analysis does not have a separate semantic category of *de re* interpretation, though as shown below it easily accounts for the distinctions that *de re* interpretation was invented to handle.

De re interpretation for us is a pragmatically identifiable subcase of non-*de translato* interpretation in which both the attributor and the attributee have a specific understanding of who or what the *res* is. An expression *ex* of the logic is specific for a judge *j* iff all valuations in bv_j assign to *ex* the same value. A judge *j*'s assigning specific interpretations to two distinct expressions ex_1 and ex_2 makes it possible in principle for ex_1 and ex_2 to be coreferent for *j*. *De re* interpretation will then obtain when ex_1 is contributed by a referring expression in the clause embedded under *believe* that raises to a higher clause, ex_2 is contributed by the beliefs of the attributor, and the attributor takes ex_1 and ex_2 to be coreferent.

3.4.1.1 Ralph and Ortcutt

We illustrate our analysis of *de re* interpretation with (32) in Situation C, modeled after Quine 1956.

(31) **Situation C**
Ralph sees a man in a dark alley involved in suspicious-looking activities. He says: "The man in the alley is a spy." Separately, he says: "Ortcutt is not a spy." Unbeknownst to Ralph, the man in the alley *is* Ortcutt.

(32) Ralph believes Ortcutt is a spy, but he does not believe ORTCUTT is a spy.

In Situation C, Ralph has an internal *the-man-in-the-alley-is-a-spy* belief but has neither an internal *Ortcutt-is-a-spy* belief nor an internal *the-man-in-the-alley-is-Ortcutt* belief. This makes it impossible to infer from Ralph's internal beliefs alone that Ortcutt is a spy. The first conjunct of (32) can nevertheless come out true, a fact accounted for by the semantics of *believe* in (14). The true interpretation is based on a representation with *Ortcutt* raised to the matrix clause.

(33) **Translation:**

$$[\![[\mathit{Ortcutt}\ [1_e\ [\mathit{Ralph\ believes}\ t_1\ \mathit{is\ a\ spy}]]]]\!]^{j,h}$$
$$= [\![[1_e\ [\mathit{Ralph\ believes}\ t_1\ \mathit{is\ a\ spy}]]]\!]^{j,h}\ ([\![\mathit{Ortcutt}]\!]^{j,h})$$
$$= [\![\mathit{Ralph\ believes}\ t_1\ \mathit{is\ a\ spy}]\!]^{j,h[[\![\mathit{Ortcutt}]\!]^{j,h}/1]}$$
$$= [\![\mathit{believes}]\!]^{j,h[ortcutt_j/1]}\ ([\![t_1\ \mathit{is\ a\ spy}]\!]^{ralph,h[ortcutt_j/1]})\ ([\![\mathit{Ralph}]\!]^{j,h[ortcutt_j/1]})$$
$$= believe_j\ (spy_{ralph}\ ([\![t_1]\!]^{ralph,h[ortcutt_j/1]}))\ (ralph_j)$$
$$= believe_j\ (spy_{ralph}(ortcutt_j))\ (ralph_j)$$

Interpretation:

$$\|believe_j\ (spy_{ralph}(ortcutt_j))\ (ralph_j)\|^{M,i,g} = 1 \text{ iff } INFER(ralph_j,j,spy_{ralph}(ortcutt_j))$$

To show that the sentence is true under this interpretation, we need to find a set of beliefs *b* of Ralph's and a set of subset-based beliefs *q* of the speaker's/*ev*'s such that $spy_{ralph}(ortcutt_j)$ can be inferred from *b*[*j*] and *q*[*j*] together but not from *q*[*j*] alone. In Situation C, Ralph has four beliefs *B* such that $B(ralph) = spy_{ralph}(x_{mia_{ralph}})$, where $x_{mia_{ralph}}$ is the abbreviated translation of *the man in the alley* with *ralph* as judge:

(34) a. $\Lambda z[\mathrm{spy}_{\mathrm{ralph}}\ (\mathrm{x}_{\mathrm{mia}_{\mathrm{ralph}}})]$
b. $\Lambda z[\mathrm{spy}_{\mathrm{ralph}}\ (\mathrm{x}_{\mathrm{mia}_{z}})]$
c. $\Lambda z[\mathrm{spy}_{z}\ (\mathrm{x}_{\mathrm{mia}_{\mathrm{ralph}}})]$
d. $\Lambda z[\mathrm{spy}_{z}\ (\mathrm{x}_{\mathrm{mia}_{z}})]$

Acceptance of these beliefs by *j* will result in the following formulas:

(35) a. $\mathrm{spy}_{\mathrm{ralph}}\ (\mathrm{x}_{\mathrm{mia}_{\mathrm{ralph}}})$
b. $\mathrm{spy}_{\mathrm{ralph}}\ (\mathrm{x}_{\mathrm{mia}_{j}})$
c. $\mathrm{spy}_{j}\ (\mathrm{x}_{\mathrm{mia}_{\mathrm{ralph}}})$
d. $\mathrm{spy}_{j}\ (\mathrm{x}_{\mathrm{mia}_{j}})$
Abbreviation of {a,b,c,d} : $\mathrm{spy}_{\mathrm{ralph/j}}\ (\mathrm{x}_{\mathrm{mia}_{\mathrm{ralph/j}}})$

We can explain the truth of the first conjunct of (32) by taking the witness belief implicit in (33) to be the singleton set containing (34b). The relevant inference, given in (36), uses a member of EC($spy_{ralph}(x_{mia_j})$) for Premise 1, and a member of EC($spy_{ralph}(ortcutt_j)$) for the conclusion. Premise 2 is the premise that all properties that are true of *j*'s man in the alley are also true of *j*'s Ortcutt according to *j*, the truth of which follows from *j*'s belief that the man in the alley is Ortcutt.

(36) Premise 1: $\lambda P[P({}^{\phi}\mathrm{x}_{\mathrm{mia}_{j}})]\ (\lambda x[\mathrm{spy}_{\mathrm{ralph}}(x)])$
Premise 2: $\mathrm{subset}_{j}\ (\lambda P[P({}^{\phi}\mathrm{x}_{\mathrm{mia}_{j}})])\ (\lambda P[P(\mathrm{ortcutt}_{j})])$
Conclusion: $\lambda P[P(\mathrm{ortcutt}_{j})]\ (\lambda x[\mathrm{spy}_{\mathrm{ralph}}(x)])$

There is no valid inference from Premise 2 alone to the same conclusion, rendering the interpretation in (33) true for an evaluator who understands Situation C, as desired.

The role that specificity plays in generating a true interpretation is in making Premise 2 in (36) possibly true. In Situation C, Ralph has beliefs *B* such that $B[ralph] = \{spy_{ralph}(x_{{}^{\phi}mia_{ralph}})\}$, where the demonstrative extension condition ϕ restricts the denotation of ${}^{\phi}mia_{ralph}$ to a specific individual. Formally this means that all valuations in bv_{ralph} assign the same unique value to ${}^{\phi}mia_{ralph}$. We take acceptance of a belief to leave the extension condition and hence specificity unchanged, so specificity of the denotation of ${}^{\phi}mia_j$ comes with *j*'s acceptance of Ralph's belief. If *j*'s Ortcutt is also specific for *j*, *j* can rationally identify *j*'s Ortcutt with *j*'s acceptance of Ralph's man-in-the-alley, rendering $subset_j(\lambda P[P(x_{\ {}^{\phi}mia_j})])(\lambda P[P(ortcutt_j)])$ in Premise 2 of (36) potentially true. Premise 2 will be actually true just in case *j* believes the man in the alley is Ortcutt, something specified in Situation C and hence believed by any individual *j* who knows the facts in Situation C.

The second conjunct of (32) is true in Situation C under an interpretation just in case the first conjunct is false under that interpretation. The first conjunct of (32) is false in Situation C when no expression raises out of the embedded clause.

(37) **Translation:**[21]
$[\![\textit{Ralph believes Ortcutt is a spy}]\!]^{j}$
$= [\![\textit{believes}]\!]^{j}\ ([\![\textit{Ortcutt is a spy}]\!]^{ralph})\ ([\![\textit{Ralph}]\!]^{j})$
$= \text{believe}_j\ (\text{spy}_{ralph}(\text{ortcutt}_{ralph}))\ (\text{ralph}_j)$

Interpretation:
$\|\text{believe}_j\ (\text{spy}_{ralph}(\text{ortcutt}_{ralph}))\ (\text{ralph}_j)\|^{M,i,g} = 1$ iff
$\text{INFER}(\text{ralph}_j, j, \text{spy}_{ralph}(\text{ortcutt}_{ralph}))$

The sole difference between this interpretation and the one from above lies in the conclusion of the inference. Here the expressions in the conclusion $spy_{ralph}(ortcutt_{ralph})$ all determine Ralph's believed extensions, while in the earlier example the expression $ortcutt_j$ determined the judge *j*'s believed extension. Given the description of Situation C, the only relevant beliefs we can justifiably take Ralph to have derive from his claims that the man in the alley is a spy and that Ortcutt is not a spy. There are again multiple beliefs B_1 and B_2 of Ralph's compatible with these respective claims, having the following properties:

(38) $B_1(\text{ralph}) = \text{spy}_{ralph}(x_{mia_{ralph}})$
$B_2(\text{ralph}) = \text{not}_{ralph}(\text{spy}_{ralph}(\text{ortcutt}_{ralph}))$

However, it is not possible to infer $spy_{ralph}(ortcutt_{ralph})$ from acceptance of any of these beliefs combined with any subset-based beliefs of the attributor's without imputing additional beliefs to Ralph that are not justified in Situation C.

21 We suppress the variable determination function when it does no work, as here.

3.4.1.2 The shortest spy

Reduction of acquaintance to specificity also allows us to predict that a *de re* interpretation for an expression ex_1 is false when the attributor assigns a specific interpretation to ex_1 but there is no relevant corresponding belief-generated expression ex_2 that is interpreted specifically. This prediction is borne out by Kaplan's (1968) shortest spy example.

We are to suppose (i) that Ralph has the internal belief that there are spies, (ii) that Ralph has the internal belief that among the spies, one, who he knows nothing else about, is shortest, and (iii) that as a matter of fact, Ortcutt is the shortest spy. It is observed that (i-iii) are not sufficient to render (39) true:

(39) Ralph believes Ortcutt is a spy.

In this case, Ralph is assumed *not* to be acquainted with a specific individual he identifies as the shortest spy. His beliefs B' that the shortest spy is a spy are such that $B'[ralph] = \{spy_{ralph}(x\ {}^{\psi}ss_{ralph})\}$, where $x\ {}^{\psi}ss_{ralph}$ is the translation of *the shortest spy* and ψ does not restrict ${}^{\psi}ss_{ralph}$ to a specific individual. This means there are distinct valuations $\langle i,g \rangle$ and $\langle i',g' \rangle$ in bv_{ralph} such that $\|x\ {}^{\psi}ss_{ralph}\|^{\mathrm{M,i,g}} \neq \|x\ {}^{\psi}ss_{ralph}\|^{\mathrm{M,i',g'}}$. Attributor j's accepting a belief among those in B' preserves (non-) specificity, resulting in j's belief valuations bv_j containing valuations that assign different values to whichever of $x\ {}^{\psi}ss_{ralph}$ or $x\ {}^{\psi}ss_j$ constitutes j's acceptance of Ralph's shortest spy. This makes it impossible for j to rationally identify j's Ortcutt with j's acceptance of Ralph's shortest spy, since doing so would result in one individual in the domain D_e being identified with multiple distinct individuals in that domain. In this case, $subset_k(\lambda P[P(x\ {}^{\psi}ss_j)])(\lambda P[P(ortcutt_j)])$ cannot be rationally believed, and plausibly no other belief of j's can underwrite the needed inference to the conclusion $spy_{ralph}(ortcutt_j)$, rendering (39) false as desired.[22,23]

22 If we modify the shortest spy scenario by stipulating *Ortcutt* to be an Evans 1979-like descriptive name for *whoever is Ralph's shortest spy* (parallel to Evans' example of the name *Julius* being coined to refer to whoever invented the zip), then (39) is predicted to be true, correctly, we believe.

23 Note that evaluation of $subset_k(\lambda P[P(x\ {}^{\psi}ss_j)])(\lambda P[P(ortcutt_j)])$ requires assigning an extension for *ss* based on j's beliefs about what satisfies ψ despite the fact that ψ does not represent the extension conditions j assigns to *ss*. This is what makes appeal to extension conditions necessary, as mentioned in footnote 10.

3.4.2 Transparency

Percus (2000) has argued that while many predicate-denoting expressions in the clausal argument of *believe* can be given a transparent interpretation, the main predicate in that argument cannot. To illustrate the claim, consider the following variant of an example of Percus's:

(40) Mary thinks the smartest man is Canadian.

Intuitively, (40) is rendered true by the smartest man being such that Mary has the belief that he is Canadian, with *the smartest man* interpreted transparently. It is not, however, rendered true by the actual Canadians being such that Mary has the belief that the smartest man is one of them if neither the evaluator nor Mary has the belief that they are Canadian, an interpretation that should be possible if *Canadian* were interpreted transparently.

Working with a possible-worlds semantics, Percus analyzes transparency as binding by a world-binding operator contributed by a higher clause. (40) comes with two world-binding λ-operators: one at the matrix level and one at the level of the embedded clause. Percus takes his observation to show that the world variable contained in the interpretation of *Canadian* cannot be bound by the higher of these two λ-operators while the world variables contained in the interpretation of *the smartest man* can be.

It should be noted that the conclusion Percus draws – that *Canadian* cannot be interpreted transparently – only follows if the higher world-binding λ-operator is implicitly assumed to apply to the actual world, since this is the only way to generate the objective actual world extension of *Canadian* under the assumptions he adopts. If that operator applies instead to the belief worlds of the speaker or evaluator, the suspect interpretation would not be generated. Rather, the sentence would be predicted to mean that Canadians as understood by the speaker/evaluator are such that Mary believes the smartest man to be among them.

Since our analysis never generates objective extensions, under our analysis, Percus's suspect interpretation cannot be generated. *In situ* interpretation of the word *Canadian* for us results in the expression having in its extension only functions true exclusively of individuals Mary considers to be possibly Canadian, not actual Canadians. Raising of *Canadian* to the matrix clause for us results in the expression having in its extension only functions true exclusively of individuals the matrix judge – typically the speaker or the evaluator – considers to be possibly Canadian, not actual Canadians. In both cases, the conclusion of the relevant inference involves (a function that determines) who someone believes to be Canadian, not who is actually, independent of anyone's beliefs, Canadian. In short, under our analysis there is no such thing as transparent interpretation in the sense assumed by Percus.

We do generate a *de qualitate* interpretation for *Canadian* in (40). This is as it should be. (40), for example, is true in a situation where Mary claims *The smartest man is Quebecois* and where being Quebecois entails being Canadian for the speaker/evaluator *j* but not for Mary. Our analysis accounts for this fact by raising *Canadian* in (40) to the matrix clause. The conclusion of the required inference then becomes $canadian_j(x_{smartest\text{-}man_{mary}})$, and the following inference underwrites the truth of (40):

(41) Premise 1: $\text{quebecois}_j(x_{smartest\text{-}man_{mary}})$
Premise 2: $\text{subset}_j\ (\text{quebecois}_j)\ (\text{canadian}_j)$
Conclusion: $\text{canadian}_j(x_{smartest\text{-}man_{mary}})$

This *de qualitate* interpretation differs crucially from Percus's suspect transparent interpretation in that the only property of being Canadian that plays a role is the matrix judge's understanding of being Canadian, not an objective understanding that is independent of the matrix judge's beliefs. We take the absence of Percus's suspect interpretation to suggest not only that interpretation *can* be made relative to a particular individual, but that it *must* be. If objective transparent interpretation were admitted without constraint, for example by allowing for basic logical expressions to be assigned interpretations depending only on their objective satisfaction of the associated extension conditions, there would be no way of blocking objective transparent interpretation of *canadian* in place of the matrix judge's interpretation assigned to $canadian_j$ for the inference in (41), and Percus's challenge could then not be met.

3.4.3 Role of the clausal argument of believe

A naïve view of belief attribution would identify the clause embedded under the verb *believe* as giving the content of one of the subject's internal beliefs, what could be called a *de dicto* interpretation. The existence of *de re* attributions like those examined by Quine (1956) already showed that this naïve view alone is not sufficient. However, it still remains common to accept that at least some belief attribution sentences are interpreted in this naïve way. Under our analysis, the embedded clause of a belief attribution sentence is never interpreted as identifying an internal belief of the subject's, putting our analysis on a par with Bach 1997.

Our approach contrasts with that of Blumberg & Lederman (2020) (henceforth BL), who argue based on the following example from Darde 1982 (cited from Blumberg & Holguín 2018) that a new mode of interpretation is needed in addition to the standard *de dicto, de re* and *de qualitate* interpretations:

(42) Ann is a six-year-old girl whom Pete, an expert in tennis pedagogy, has never met and whose existence he is unaware of. Pete believes that every six-year-old can learn to play tennis in ten lessons. Jane, Ann's aunt, is aware of Pete's feelings on the matter. Jane wants to encourage Ann's father, Jim, to sign Ann up for tennis lessons, so in conversation with Jim she asserts the following: *Pete believes Ann can learn to play tennis in ten lessons.*

We accept that this example cannot be explained using the standard modes of interpretation mentioned above. However, we reject the view that a new mode of interpretation is needed *in addition to* the standard modes. The mode of interpretation argued for in this paper accounts directly for (42), like that proposed by BL. However, unlike BL's analysis, our analysis also accounts for examples previously used to motivate the other standard modes of interpretation, and also extends to mathematical beliefs.

Under our analysis, the inference that renders the belief statement in (42) true is the following:

(43) Premise 1: $\lambda P[\forall x(\text{6-year-old}_j(x))(P(x))]$ $(\lambda y[\text{can-play-tennis-in-10-lessons}_{pete}(y)])$

Premise 2: $\text{subset}_j\ (\lambda P[\forall x(\text{6-year-old}_j(x))(P(x))])\ (\lambda P[P(\text{ann}_j)])$

Conclusion: $\lambda P[P(\text{ann}_j)]\ (\lambda y[\text{can-play-tennis-in-10-lessons}_{pete}(y)])$

Premise 1 constitutes one form of *j*'s acceptance of Pete's belief that every 6-year-old can learn to play tennis in 10 lessons. Premise 2 is the premise that the set of properties that hold of everyone *j* identifies as 6 years old is a subset of the set of properties that hold of *j*'s Ann. Since the conclusion in (43) follows from the two premises together but does not follow from Premise 2 alone, the final sentence of (42) is correctly predicted to be true under our analysis in the situation described.

3.5 Correlations with traditional analyses

While the analysis we have presented has only one source of formal ambiguity, a *de translato*/non-*de translato* ambiguity determined by presence or absence of cross-clausal syntactic movement, it has the tools to reconstruct *de dicto, de re* and *de qualitate* interpretations. This was shown for particular examples above. In this section we provide a more general analysis of these interpretations.

De dicto attribution under the proposed analysis corresponds to a typical case of interpretation without any raising out of the embedded clause, with expressions in the embedded clause having the extensions that the attributee takes them to have. *De dicto* interpretation will follow if in the accepted attributee belief of Premise 1, all basic expressions are attributee-subscripted. We call this a pure attributee belief. In this case, any attributor *j*'s contribution in Premise 2 will relate two attributee-evaluated expressions via $subset_j$, entailing that according to *j*, the attributee believes their extensions to stand in a subset relation. The conclusion of the inference then has to follow entirely from beliefs that the attributor takes the attributee to have. *De dicto* interpretation differs from Tancredi and Sharvit (2020)'s *de translato* interpretation in that in a typical case of *de dicto* interpretation the attributee and attributor agree on the extension conditions they associate with their words, whereas in TS's *de translato* interpretation they do not. Under our implementation of *de translato* interpretation here, however, *de dicto* is formally a subcase of *de translato*.

While *de dicto* interpretation derives from *in situ* interpretation of the embedded clause, *in situ* interpretation of the embedded clause does not always yield *de dicto* interpretation. *In situ* interpretation determines the conclusion of the relevant inferences, but only affects the identity of the beliefs underlying Premises 1 and 2 indirectly. Just as an attributor can believe, for example, that Ralph's *the man in the alley* has the same extension as the speaker's *Ortcutt*, the attributee too can have beliefs about the denotations of expressions evaluated by other people. Suppose that Ralph talks to Big Al, who identifies the man in the alley as Ortcutt, and that Ralph forms the belief that Big Al's Ortcutt is not Ralph's Ortcutt. Ralph may well come to accept that Big Al's Ortcutt is a spy without believing that Ralph's Ortcutt is a spy. Ralph's beliefs (as accepted by Ralph) would then be the following, with ϕ and ψ used to distinguish two homophonous names *Ortcutt* via their associated extension conditions.

(44) a. $\text{spy}_{\text{ralph}}(^{\phi}\text{ortcutt}_{\text{ralph}})$
b. $\text{not}_{\text{ralph}}(\text{spy}_{\text{ralph}}(^{\psi}\text{ortcutt}_{\text{ralph}}))$

Now suppose that Big Al understands this about Ralph, and that he further comes to the belief that Big Al's Ortcutt is identical to Ralph's Ektelpingras, an identity that Ralph himself does not hold. This belief combined with Ralph's belief in (44a) would then underwrite the inferences needed to render the following utterance of Big Al's true for anyone in full understanding of the situation, analyzed with no raising out of the embedded clause:

(45) Ralph believes Ektelpingras is a spy.

This kind of example does not have a standard name in the literature, but it is clearly not an intuitive instance of *de dicto* attribution. This highlights the fact that under our analysis, *de dicto* interpretation is not a semantic category but rather a pragmatic one. It is not determined compositionally based solely on the structure that is input to translation and interpretation. Whether something counts as *de dicto* attribution depends additionally on the nature of the beliefs assumed to underlie the relevant inferences, and these beliefs are not fixed by the semantics.

De re attribution under the proposed analysis derives from an attribution in which a referring expression has been raised out from the embedded clause. Through raising, the extension of the raised expression becomes the believed extension of the attributor rather than of the attributee. *De re* interpretation results when the attributor identifies this extension with that of a referring expression in their acceptance of the attributee's belief underlying Premise 1.

In analyzing *de re* attribution, Kaplan (1968) proposes that the attributee has to have a vivid name for the *res*.[24] Our analysis encodes no such restriction because it does not treat *de re* interpretation as a semantically distinct interpretation. Just as with *de dicto* interpretation, a belief attribution sentence with a fixed syntactic analysis can give rise to a range of distinct traditional and non-traditional interpretations. For example, the sentence *Ralph believes Ortcutt is a spy*, with *Ortcutt* (and nothing else) raised into the matrix clause, is predicted to be true not only in Situation C, but also when Ralph has the belief that whoever undertakes an interaction in a dark alley at night is a spy, provided that the person evaluating the sentence believes Ortcutt undertook an interaction in a dark alley at night. Though the semantics of

24 See Anand 2006 for relevant discussion about Kaplan's vividness requirement.

the sentence remains identical in the two situations, in Situation C it would traditionally be labeled *de re*, while in the just-described situation it would not. A distinction between such *de re* and non-*de re* understandings can, of course, be made, but on the proposed analysis it is a pragmatic distinction, not a semantic one.

Hazel Pearson (p.c.) notes additionally that the sentence *Ralph believes Ortcutt is a spy and he believes Ortcutt is not a spy* is correctly predicted to be true in a situation differing from Situation C in two ways: (i) Ralph does not have a name for Ortcutt, and (ii) Ralph sees Ortcutt at the beach and says *The man at the beach is not a spy* rather than *Ortcutt is not a spy*. For us the truth of this apparently contradictory belief attribution comes from Ralph's having two relevant beliefs whose acceptance the attributor takes to be about Ortcutt: that the man in the alley is a spy and that the man at the beach is not a spy. Raising of *Ortcutt* in both conjuncts accomplishes for us what *de re* analyses accomplish with an additional stipulated interpretation. From acceptance that the man in the alley is a spy together with the belief that the man in the alley is Ortcutt it can be concluded that Ortcutt is a spy, accounting for the truth of the first conjunct. From acceptance that the man at the beach is not a spy together with the belief that the man at the beach is Ortcutt, it can be concluded that Ortcutt is not a spy. For us, however, no independent *de re* interpretation is required to obtain this result.

De qualitate attribution is analyzed as involving raising of a function-denoting expression, with the attributee's beliefs again being pure attributee beliefs. Our analysis improves on that of Schwager (2009), which makes incorrect predictions for property expressions in the restrictive clause of a quantifier. On Schwager's analysis, summarized in Appendix 1, *de qualitate* interpretation only affects property-denoting expressions, and is independent of the environments those expressions are found in. If *X* and *Y* are property-denoting expressions and the extension of *X* is a subset of that of *Y*, Schwager's analysis licenses substitution of *X* by *Y* in any environment. We find, however, that the environment matters. Consider a situation in which John says "No tall man came". I can report on John's beliefs as in (46a), but not as (46b). Schwager, however, predicts that (46b) should be the acceptable belief report, not (46a), since the extension of *tall man* is a subset of the extension of *man*, but not of the extension of *attractive tall man*.

(46) a. John believes no attractive tall man came
b. John believes no man came

Under our analysis, these facts are predicted. The true translation of (46a) is derived by scopelessly raising the entire quantifier expression into the matrix clause, with the trace left behind being of quantifier type, and the required subset relation applying to quantifier-phrase meanings rather than property meanings. A parallel analysis of (46b) fails to render it true, and no other analysis available within the confines of our proposal does so either since syntactically, the restrictive clause of a quantifier phrase cannot be raised out of the quantifier phrase.

TS's ***de translato*** interpretation of an expression x comes for us from x being interpreted *in situ*, inside the scope of *believe*. *De translato* interpretation results from the pair of assumptions (i) that translation and interpretation are relative to a judge, and (ii) that *believe* changes the judge parameter for translation of its embedded clause to the translation of its subject. Raising out from the scope of *believe* generates non-*de translato* interpretations. The *de translato*/non-*de translato* distinction is for us a true ambiguity, in fact the only true ambiguity made possible by the semantics proposed for *believe*.

We distinguish two versions of *de translato* interpretation. The most common case involves attributor and attributee agreeing on all relevant extension conditions but potentially disagreeing on their denotations, such as in *de dicto* interpretation. The cases examined in TS are the exceptional cases where attributor and attributee mean different things by one or more of their terms. Though the distinction is important, it does not derive from an ambiguity. It rather derives from a comparison of one expression in two different sentences, *a believes S*, and *(I believe) S*. In the cases examined by TS, some expression contained in *S* will have both different extension conditions and different subscripted judges in the relevant translations of the two sentences, while in the other more normal cases the translations will differ only in the subscripted judge. The sentence *S* itself, however, does not count as ambiguous in this case since the context it occurs in fixes its translation and interpretation uniquely.

Cases of *de translato* attribution under translation across languages highlight the importance of the fact that translation and interpretation are evaluator relative as well as judge relative. If Taro, a monolingual Japanese speaker under a misunderstanding parallel to Hal's in (4), states *26-ga sosuu-desu* (=

the literal Japanese translation of *26 is prime*), the following English sentence can be true:

(47) Taro believes 26 is prime

Under the proposed analysis, this is accounted for by analyzing the embedded clause *de translato*. The logical translation of the English embedded clause is that determined by the evaluator, not by Taro. Taro's role as judge shows up only in the subscripts and superscripts attached to the logical translation ${}^{\phi}prime_{taro}({}^{\psi}26_{taro})$. As an expression of logic, ${}^{\phi}prime_{taro}$ can function both as the logical translation of the English word *prime* and as the logical translation of the Japanese word *sosuu* or its counterpart within an accepted belief, and so there is nothing blocking an evaluator from taking Taro to associate abnormal extension conditions with this logical expression.

3.5.1 A note on de translato

A reviewer raises an objection to our application of *de translato* interpretation based on an example from Davidson 1968 involving a speaker, here John, who means *orange* by the word 'hippopotamus'. If another speaker knows this fact about John, our analysis predicts the following sentence to be potentially true when evaluated by that speaker, with *hippopotamus* interpreted *in situ* and hence *de translato*.

(48) John believes he has a hippopotamus in the refrigerator.

The reviewer's objection is that "there simply is no such reading of (48), except perhaps in very special circumstances where some indication is given of a non-standard interpretation of 'hippopotamus' (in which case the sentence is crucially altered)." A relevant indication of a non-standard interpretation would be, for example, use of scare quotes when uttering the word *hippopotamus*.

We have found speakers to divide into two groups regarding whether overt indication of a non-standard interpretation is needed within the sentence. For the reviewer such indication is necessary, while for the first author it is not. Consider a situation where John throws me an orange while saying "Here, have a hippopotamus. I have another in the fridge." In this situation, for speakers in the former group using (48) to indicate John's belief that there is an orange in the fridge requires highlighting the word *hippopotamus*, e.g. with scare quotes, while for speakers in the latter group no such highlighting is needed as long as speaker and hearer share the relevant context and take themselves to do so.

The more general issue behind the reviewer's objection is that our semantics, and that of TS on which it is based, makes non-standard interpretations available without constraint, rendering the number of possible interpretations involving a single *de translato* interpretation unbounded. We accept this consequence, leaving disambiguation to the pragmatics.[25] In the vast majority of cases, we assume that disambiguation defaults to an evaluator's assigning the same extension conditions to an attributee's expression as the evaluator assigns. Departure from the default we take to require justification on behalf of the evaluator.

3.6 *De se*

The one common interpretation that we have not incorporated into our analysis is ***de se***. The standard Lewis 1979 analysis of *de se* is not available to us since we do not make use of possible worlds in the analysis of belief attribution and so cannot appeal to centered worlds either. A possible approach to *de se* interpretation would be to add judges to variables and to take *de se* interpretation to derive from a variable $x_{i,j}$, with index i and judge j, anaphoric on j. Such an analysis would render *de se* a subcase of *de translato* for second and third person pronouns. Whether this is a sufficient account of *de se* interpretation is a question that we leave to future work.

3.7 Inference

The inference relation we appeal to strictly limits the patterns of inference that can underwrite a belief attribution statement. In particular, we do not incorporate many inference patterns intuitively felt to be valid, such as *modus ponens*, *modus tollens*, inferences from a conjunction of propositions to either of the propositions conjoined, inferences from a proposition to any disjunction containing that proposition, or any other inferences based exclusively on propositional logic. This is as it should be.

25 Standard theories of semantics based on expressions having objective extensions have to contend with the parallel issue of dealing with an unbounded number of possible misinterpretations by language users. The only important difference is that under standard theories such misinterpretations have to be understood as pragmatic mis-graspings of a fixed semantic interpretation rather than as normal semantic interpretations.

We accept that a person who believes a conjoined proposition *p&q* will also be expected to believe the individual conjuncts *p* and *q*, or a person who believes both *p* and $p{\rightarrow}q$ will also be expected to believe *q*. When it comes to identifying the particular internal beliefs of an individual that can play a role in inference, we allow all legitimate inference relations, like these, to be appealed to. What is not possible under the proposed analysis is for the attributor contribution to feed a propositional logic-based inference.

To see why such a restriction is needed, consider the case of an attributor *j* holding the belief Q_4 below.

(49) $\mathrm{Q}_4 = \Lambda \mathrm{j}[\mathrm{equal}_{\mathrm{mary}}(\mathrm{plus}_{\mathrm{mary}}(1_{\mathrm{mary}})(4_{\mathrm{mary}}))(\mathrm{plus}_{\mathrm{mary}}(2_{\mathrm{mary}})(2_{\mathrm{mary}})) \rightarrow_{\mathrm{j}} \mathrm{real\text{-}number}_{\mathrm{j}}(\mathrm{i}^{\mathrm{i}}{}_{\mathrm{j}})]$

This is roughly the belief that if according to Mary 1+4=2+2, then according to *j* i^{i} is a real number. If *j* believes i^{i} is a real number, then this is a plausible belief for *j* to hold. If we combined this belief with Mary's belief that 1+4=2+2, then the truth of *i^i is a real number* would follow by *modus ponens*. Using *modus ponens* as a valid inference pattern for underwriting belief attribution would then lead to the incorrect prediction that *Mary believes that i^i is a real number* should come out true in Situation A. That it fails to do so shows that the analysis that leads to that prediction must be rejected.

3.8 Logic

In the framework we have proposed, natural language expressions and beliefs are translated into expressions of logic that serve as the input to interpretation. Expressions of the logic, however, are not given an objective interpretation. They are instead assigned extensions compatible with a judge's beliefs about what satisfies the extension conditions on the basic vocabulary. A consequence of all interpretation being belief based like this is that our logic has no room for a logical vocabulary with a fixed, objective interpretation, vocabulary like ¬, &, ∨, →, =, ∀, and ∃.[26] The English word *equals*, for example, does not translate into logic as "=", but rather typically as $^{=}equals_j$, with *j* the relevant judge and the superscripted occurrence of "=" encoding the identity relation, the relation of two expressions of logic having the same extension. (When *j* misunderstands the word *equals*, of course, "=" will have to be replaced by some other concept.) What pairs of expressions $^{=}equals_j$

26 Thanks to Lucas Champollion for discussion that helped us become clearer on this point.

applies to, however, is determined by the beliefs of the judge *j*. This is what makes it possible for a sentence like *1+4 equals 2+2* to come out true when evaluated with respect to Mary as judge and false when evaluated with respect to a mathematically adept individual as judge. What remains in our logic after the logical vocabulary has been purged is only functions composing with arguments.

It would be nice if we could reduce the restrictions on inference argued for in Section 3.7 to the use of an impoverished logic. Unfortunately, we do not see how such a reduction can be made. The difficulty lies in our dependence on the subset relation. This relation encodes a very specific concept, but the choice of concept is dictated solely by its utility. It is not forced by the adoption of our logic. The logic equally allows for a parallel inference relation making use of material implication rather than the subset relation. The only argument we have against adopting such an alternative is that it gets the facts wrong. The question of why inference needs to be restricted as we have argued thus remains at this point an open question.

4 Consequences

At the core of our I-semantics is the assumption that the extension conditions associated with an expression do not objectively determine an extension. The opposing idea that they do goes back at least to Frege 1892. Even people who explicitly reject Frege's objective view of language and adopt instead Chomsky's (1986) I-language perspective, such as Tancredi 2007a and Tancredi 2007b, have generally accepted the Fregean view that sense objectively determines reference.

Our abandonment of this idea follows in the footsteps of Kripke (1980) and Putnam (1975). Kripke famously argued that the reference of a name is not determined by an underlying concept or sense associated with the name, while Putnam made a similar case for the reference of natural kind terms. However, where Kripke proposed to eliminate senses for names, treating names as rigid designators, we opt for relativizing senses, or more accurately extension conditions and the entities taken to satisfy them, to beliefs. This allows us to analyze cases of incomplete knowledge about reference as involving restrictions on the reference of a name that do not narrow down the possible referents to one. Among other consequences, this makes it easy to explain why modal sentences like *John may be Bill and he may not be* can easily be true, an intractable problem for Kripke. If a speaker takes *John* to refer to either to *a* or *b* and *Bill* to refer to *b*, there are valuations compatible with that speaker's belief valuations for which John is Bill and there are others for which John is not Bill.

Relativizing interpretation to belief makes it a near impossibility for any two speakers to associate the exact same extensions with their expressions. However, we take it to be readily possible for them to associate identical extension conditions with their expressions. Indeed, we take the presumption of agreement on extension conditions as central to successful communication, making such agreement a pragmatic presupposition of all linguistic communication. Exceptions we take to only be considered when there is clear evidence that this presupposition cannot be upheld, cases like Hal's confusion.

This shift of perspective on the nature of meaning has wide-ranging consequences. On the philosophical side, it undermines Putnam's (1975) argument that "meaning ain't in the head". The real-world objects that are represented in the model are not in the head, but they are also not part of the meanings of expressions under our analysis. It also undermines Burge's (1978) related argument that beliefs are not in the head. Belief attribution involves more than what is in the head of the attributee, as Burge rightly noticed, but beliefs in the head of the attributee play an ineliminable role in belief attribution, and the only other things that play a role are beliefs in the head of the evaluator. It also opens up new paths for defusing Kripke's (1979) puzzle about belief, though these paths will have to tackle the thorny question of how translation across languages affects interpretation. On the linguistic side, it opens up the possibility of eliminating possible worlds from semantics altogether.

Our analysis makes a step toward a possible-worlds-free semantics by eliminating possible worlds from the semantics of belief attribution. An alternative that has been pursued by Tancredi (2007a,b) and Muñoz (2019) instead doubles down on possible worlds by allowing the semantics to manipulate multiple models, each with its own set of possible worlds. While Muñoz showed that such an approach can give a principled explanation of many cases of *de translato* interpretation, we note here that a multiple-model approach does not account for belief attributions involving impossible beliefs like those of Mary in (2), or for the intuitive inferences that can be made involving such attributions. Indeed, this is a problem affecting all objective, possible worlds-based analyses we are aware of. From our perspective, the problem stems not from the use of possible worlds themselves, but rather from the assumption that the extension of a term is fixed objectively. Adding possible worlds to the models we use for interpretation should thus be unproblematic as long as this extra assumption is not brought in with it.

The foundation of our analysis of belief attribution consists of two core parts: (i) judge- and evaluator-dependent translation and interpretation, and (ii) use of a restricted inference process. A reviewer raises the possibility of accepting (i) while maintaining a possible worlds analysis of belief attribution that makes use of standard entailment rather than restricted inference, since (i) alone makes it possible to assign non-necessary interpretations to sentences that on the surface appear to state necessary truths or falsities. We cannot reject the possibility of such an analysis out of hand. We note, however, that the interpretations that (i) gives for belief attribution sentences intricately and essentially involve the compositional structure of formulas representing beliefs. Re-analyzing the inference process as possible world entailment would involve re-analyzing those formulas as sets of worlds, a re-analysis that would obliterate the compositional structure. This would leave us with the same dilemma that possible worlds analyses have unsuccessfully grappled with for over half a century: restricting worlds to the possible makes it impossible to distinguish between different impossible beliefs, while bringing in impossible worlds renders the analysis of entailment as a subset relation among worlds unworkable. We can add to this that any entailment-based analysis will have to come to grips with the fact that not all entailment relations can underwrite belief attribution, *modus ponens* being a particularly clear mode of inference that has to be rejected. We do not see a clear way through these problems that uses a possible worlds analysis of entailment in place of (ii).

5 Conclusion

We have argued in this paper for a new inference-based analysis of belief attribution. Our analysis shares with Hintikka the idea that the embedded clause in a belief attribution has to follow from an underlying belief of the subject's. The specific "follow from" relation we employ, however, is a restricted relation of inference rather than entailment. While we hope that future work will shed light on why inference has to be restricted as it is, using our restricted inference relation rather than entailment makes it possible to account for attributions of belief in apparent necessary truths and falsities and to also account for implications that hold among such attributions, overcoming major difficulties for Hintikka (1962) and CvS.

Our analysis has only one source of ambiguity: a scope distinction that derives a semantic *de translato*/non-*de translato* distinction. We showed that in addition to its empirical superiority over a wide range of cases, the proposed analysis gives a uniform analysis of *believe* that accounts for what have historically been thought of as distinct *de dicto*, *de re* and *de qualitate* interpretations as well as of examples that have not historically warranted a separate label. *De dicto*, *de re*, and *de qualitate* interpretations end up as pragmatically identifiable subcases of *de translato* or non-*de translato* interpretation. Thus, in addition to being empirically superior, our analysis is also favored over historical alternatives by Occam's razor.

We have made no attempt to give an exhaustive comparison of our analysis with other alternatives. In comparing our analysis with a small sampling of analyses and claims from the literature, however, we hope to have provided the tools needed for the interested reader to compare its predictions with those of any other analyses of interest.

6 Appendix 1: Previous analyses

In this appendix we provide formal analyses of Hintikka (1962), Kaplan (1968), Cresswell & von Stechow (1982), Schwager (2009), and Tancredi & Sharvit (2020). We identify the examples among (2), (4), (32), (39), (40), and (42) that each analysis can account for, though we leave it to the interested reader to verify these claims.

6.1 Hintikka (1962)

Hintikka analyzes sentences of the form *a believes that S* as true iff the sum total of *a*'s beliefs entail the proposition denoted by *S*. The proposition denoted by *S* for Hintikka is the set of worlds at which *S* is true. Formally, where $[\![\ ]\!]^{w}$ is a function that takes a syntactic expression and gives back its denotation at world w, and an individual's belief worlds are the worlds compatible with everything that individual believes, for Hintikka,

(50) $[\![\textit{a believes that S}]\!]^{w} = 1$ iff
$[\![a]\!]^{w}$'s belief worlds in w are a subset of $\{w' : [\![S]\!]^{w'} = 1\}$.

The subset requirement in (50) encodes entailment: *S entails S′* iff the set of worlds at which *S* is true is a subset of the set of worlds at which *S′* is true. By analyzing belief attribution in terms of entailment, Hintikka's analysis predicts the following:

(51) If *a believes that S* is true and *S* entails *S′*, then *a believes that S′* is true.

This analysis accounts for the truth of (2a,b). It does not account for the falsity of (2c,d), or for (4), (32), (39), (40), or (42).

6.2 Kaplan (1968)

Kaplan analyzes *believe* in terms of a primitive relation between an individual and an expression, and specifies the following way of quantifying into an expression:

(52) *a believes that P(b)* is true if $\exists\alpha\,[\mathrm{R}(\alpha,\mathrm{b}',\mathrm{a}')\ \&\ \mathrm{a}'\ \mathbf{B}\ \ulcorner \mathrm{P}(\alpha)\urcorner\,]$, where:
For any expression x, x′ is the interpretation of x.
B is a primitive relation between an individual and an expression.
α ranges over expressions.
$\ulcorner \mathrm{P}(\alpha)\urcorner$ picks out the expression *P(α)*.
$\mathrm{R}(\alpha,\mathrm{b}',\mathrm{a}')$ holds iff α is a vivid name of b′ for a′.

(52) licenses substitution of an expression *b* for an expression *α* within the complement clause of *believe* when *α* is a vivid name of the denotation of *b* for the subject. This is not intended as a full analysis of *believe*, but rather adds the possibility of *de re* interpretation to whatever underlying analysis of belief one adopts. On its own, (52) can account for (32), (39) and (40) but fails to account for (2), (4) and (42).

6.3 Cresswell & von Stechow (1982) (CvS)

CvS analyze *believe* as relating to structured propositions, where a structured proposition is an n+1-tuple $\langle \omega, b_1,\ldots,b_n\rangle$ consisting of an n-place relation ω and its n arguments $b_1,\ldots,b_n$. Each of these n+1 constituents is an intension, which gives rise to one of the most significant properties of their analysis: that belief attributions allow substitution of terms in a structured proposition under intensional identity without altering the truth value of the attribution. Second, they build *de re* interpretation into their analysis of belief attribution. A statement of the form *a believes that p* is true for them under a structuring $\langle\omega, b_1,\ldots,b_n\rangle$ of p iff a ascribes ω to $b_1,\ldots,b_n$. Their formal analysis incorporates David Lewis's (1979) analysis of *de se* attribution, which here is only a distraction. A simplified formal analysis of ascription is given below, where BW_a denotes a's Belief Worlds:[27]

(53) a ascribes ω to $b_1,\ldots,b_n$ in w iff there are suitable (acquaintance) relations $\xi_o,\ldots,\xi_n$ such that

a. a stands in relation ξ_o to ω alone in w;

b. for $1 \leq i \leq n$, a stands in ξ_i to b_i alone in w; and

c. $BW_a \subseteq \{w'$: $[\exists\omega', x_1,\ldots,x_n]$ ((i) a stands in ξ_o uniquely to ω' in w', (ii) for $1 \leq i \leq n$ a stands in ξ_i uniquely to x_i in w', and (iii) ω' holds of $x_1,\ldots,x_n$ in w')$\}$

The relations $\xi_o,\ldots,\xi_n$ are generalizations of Kaplan's vivid names that can relate an individual not only to another individual but also to a property, a quantifier, or any other type of meaning. Kaplan's vividness requirement gets re-cast for CvS as a requirement of cognitive contact, though as with vividness, neither necessary nor sufficient conditions are given for being in cognitive contact with a meaning. (53) can account for (2a,c,d), (32) and (39). It does not account for (2b), (4), (40) or (42).

6.4 Schwager (2009)

Schwager posits an additional mode of interpretation she calls *de qualitate*.[28]

27 Borrowing from TS, we take CvS's official intended formulation to be the following:

(i) a ascribes ω to $b_1, \ldots, b_n$ in w iff there are suitable (acquaintance) relations $\xi_0, \ldots, \xi_n$ such that

a. $\forall y\, (w \in \xi_0(a,y) \Leftrightarrow y = \omega)$;

b. for $1 \leq i \leq n$, $\forall y\, (w \in \xi_i(a,y) \Leftrightarrow y = b_i)$; and

c. a self-ascribes in w that property ϕ such that for any world w' and any individual c, $w' \in \phi(c)$ iff $\exists\omega, x_1,\ldots,x_n$ $[\forall\omega', x'_1,\ldots,x'_n$ $((w' \in \xi_0(c,\omega') \Leftrightarrow \omega \in \omega')$ & $(w' \in \xi_i(c,x'_i) \Leftrightarrow x_i = x'_i$ for $1 \leq i \leq n))$ & $w' \in \omega(x_1,\ldots,x_n)]$

For the examples considered in this paper, the full-fledged *de se* analysis in (i) and the simplified non-*de se* analysis in (52) give the same results.

(54) ***de qualitate*** (from Schwager, p.409)
$Believe_w(x,\langle P,Q\rangle)$ iff there is a property Q' s.t. at the w-closest worlds w' where $Q(w') \neq \varnothing$:
a. $Q'(w') \neq \varnothing$
b. $Q'(w') \subseteq Q(w')$
c. $Believe_w(x,\lambda w''.P_{w''}(Q'))$ is true.

Here, $\langle P,Q\rangle$ is a structured proposition, with P a function that combines with property Q to generate a proposition. This structured proposition is true at a world w iff P(w)(Q(w)) is true. Q and Q' are both properties of type $\langle s,\langle e,t\rangle\rangle$, i.e. functions from worlds to sets of individuals. The w-closest worlds w' where $Q(w') \neq \varnothing$ is the set of all worlds maximally similar to w in which $Q(w') \neq \varnothing$. If $Q(w) \neq \varnothing$, then this will be the singleton set {w}, since no world can be as similar to w as w. $\lambda w''.P_{w''}(Q')$ is a simple, non-structured proposition, the set of possible worlds w'' in which $P_{w''}(Q')$ is true, where $P_{w''}$ is the extension of P at world w''. Schwager does not specify an interpretation for $Believe_w(x,\lambda w''.P_{w''}(Q'))$, and without such a specification the analysis makes no predictions. However, analyzing it using Hintikka's analysis will overcome this problem. On this understanding, (54) adds an account of (40) to Hintikka's account of (2a,b) but does not account for (2c,d), (4), (32), (39), or (42).

6.5 Tancredi & Sharvit (2020) (TS)

TS's *de translato* interpretation analyzes *believe* as changing the judge parameter for the embedded clause to the value of the subject of *believe* and adding a subscripted T to those expressions within the embedded clause to be interpreted *de translato*. T then triggers a transformation in the language used for interpretation as follows:

(55) ***De translato* interpretation**
For any expression x, language L and judge j, $[\![x_T]\!]^{L,j} = [\![x]\!]^{T_j(L),j}$.

For a non-embedded belief attribution, *L* is the I-language of the speaker, *j* the believer, and $T_j(L)$ a transformation of *L* that makes *L* as close as possible to the believer's I-language (given what the speaker knows about that language)

28 Schwager's actual proposal applies to all attitude predicates, not only to *believe*. We have substituted *Believe* for her *Attitude* in the first line of the analysis and in (54c) to reflect the narrower focus of the present paper.

by substituting the believer's meanings of expressions for the speaker's when the two differ. (55), when added to a separate analysis of *believe*, accounts for (4), but fails to account independently for (2), (32), (39), (40), or (42).

6.6 Combinations

No combination of the above analyses accounts for all of the examples in (2), (4), (32), (39), (40) and (42) since none of the individual analyses accounts for (42). Additionally, in order to account for all four examples in (2), Hintikka, capturing (2a,b), would have to be combined with CvS, capturing (2a,c,d). Such a combination, however, gives inconsistent interpretations for attributions involving mathematical truths or falsities, and requires multiple paths to interpretation, one through structured propositions and one through unstructured propositions. A combination of these two analyses is thus not a live alternative.

7 Appendix 2: Formal analysis

types

e is a type (of *individuals*)

t is a type (of *truth values*)

exp is a type (of *expressions* (of the target language and of the logic))

ec is a type (of *extension conditions*)

If σ and τ are types, then $\langle\sigma,\tau\rangle$ is a type (of *functions* from σ-type things to τ-type things)

If $\sigma_1,\ldots,\sigma_n$ are types, then $\langle\sigma_1,\ldots,\sigma_n\rangle$ is a type (of *tuples* of things of types $\sigma_1,\ldots,\sigma_n$.)

Informal use: *An expression of type σ* is short for *an expression that either has an interpretation of or is itself of type σ.*

type-driven translation

If α is an *unindexed terminal node*, then $[\![\alpha]\!]^{\mathrm{j,h}} = {}^{\phi}\beta_j$ for some β, where ϕ is the extension condition that (the referent of) j associates with α.

If α is an *indexed terminal node* with index i and type e, then $[\![\alpha]\!]^{\mathrm{j,h}} = \mathrm{h(i)}$. (for pronouns and traces)

If α is a *non-branching node*, and β is its daughter node, then $[\![\alpha]\!]^{\mathrm{j,h}} = [\![\beta]\!]^{\mathrm{j,h}}$.

If α is a *branching node*, $\{\beta,\gamma\}$ is the set of α's daughters, and for some types σ,τ, $[\![\beta]\!]^{j,h}$ is of type $\langle\sigma,\tau\rangle$ and either $[\![\gamma]\!]^{j,h}$ or γ is of type σ, then $[\![\alpha]\!]^{j,h} = [\![\beta]\!]^{j,h}([\![\gamma]\!]^{j,h})$ or $[\![\beta]\!]^{j,h}(\gamma)$, whichever is defined.

translation of Λ-expressions

If v is a variable of type *exp* and α is an expression type τ, then $\Lambda v[\alpha]$ is a (metalinguistic) function of type $\langle exp,\tau\rangle$.

For any expression β of type *exp*, $\Lambda v[\alpha](\beta) = \alpha^{\beta/v}$, where $\alpha^{\beta/v}$ is just like α except at most in containing an occurrence of β wherever α contains v.

predicate abstraction

For any natural number *i*, type σ, individual *j*, variable determination function *h*, and expression *E*,

$[\![i_\sigma\ E]\!]^{j,h} = \Lambda x_\sigma\ [[\![E]\!]^{j,h[x/i]}]$,

where $h[x/i]$ is just like h except at most that $h[x/i](i) = x$.

interpretation

A model M is a tuple $\langle D,I,G\rangle$ where:

D is the set of domains D_σ for all types σ;

I is a set of constant assignment functions having as domain the set of all ordered pairs $\langle\alpha,\beta\rangle$ where $\alpha \in D_e$ and $\beta \in D_{ec}$.

For any constant assignment function *i*, if α is a constant, *j* an expression of type *e*, and ϕ an extension condition (of type *ec*), $\|{}^{\phi}\alpha_j\|^{M_{ev},i,g}$ is defined only if it is consistent with $\|j\|^{M_{ev},i,g}$'s beliefs that $i(\langle\|j\|^{M_{ev},i,g},\phi\rangle)$ satisfies ϕ, as evaluated by *ev*. If defined, $\|{}^{\phi}\alpha_j\|^{M_{ev},i,g} = i(\langle\|j\|^{M_{ev},i,g},\phi\rangle)$.

$\|ev\|^{M_x,i,g} = x$, the evaluator whose model is used for evaluation.

G is a set of variable assignment functions having as domain the set of all indices (natural numbers).

For any variable assignment function $g \in G$, if α is a variable with index *k*, $\|\alpha\|^{M_{ev},i,g} = g(k)$.

For any expression of the form $\alpha(\beta)$, $\|\alpha(\beta)\|^{M_{ev},i,g} = \|\alpha\|^{M_{ev},i,g}(\|\beta\|^{M_{ev},i,g})$ or $\|\alpha\|^{M_{ev},i,g}(\beta)$, whichever is defined.

interpretation of λ-expressions

If v is a variable of type σ and α is an expression of type τ, then $\lambda v[\alpha]$ is an expression of type $\langle\sigma,\tau\rangle$.

For any model M, constant assignment function i, and variable assignment function g, $\|\lambda v[\alpha]\|^{M,i,g}$ is that function f of type $\langle \sigma, \tau \rangle$ such that for any object d of type σ, $f(d) = \|\alpha\|^{M,i,g[d/v]}$, where $g[d/v]$ is that function that is exactly like g except at most that $g[d/v](v) = d$.

A **lexical term** is an expression of the form ${}^{\phi}exp_j$, where:
ϕ is an extension condition,
exp is a logical constant, and
j is a type e expression.

A **lexical item** is a pair $\langle \alpha, Z \rangle$, where α is a lexical term and Z is a set (of possible extensions).

The **lexicon** LEX is the set of all lexical items.

A lexical item $\langle \alpha, Z \rangle$ encodes beliefs about extensions as **belief valuations for the lexical term** α, BV_{α}, $= \{\langle i,g \rangle \colon \|\alpha\|^{M,i,g} \in Z\}$.

For complex expressions of the form $\alpha(\beta)$, their belief valuations are the intersection of the belief valuations for their parts: $BV_{\alpha(\beta)}$, $= BV_{\alpha} \cap BV_{\beta}$.

The **belief valuations of an individual** j, bv_j is the intersection of the belief valuations of all members of LEX_j, the set of all lexical items whose first element is a basic lexical term with subscript j.

A **belief** is a metalinguistic function $\Lambda x[c]$, where c is a formula of the logic composed from lexical terms and variables, and x occurs in c at most as a subscript on lexical terms.

Acceptance of a belief b by the individual denoted by k according to an evaluator ev results in the modification of LEX_{ev} to render $b(k)$ true for k according to ev and compatible with LEX_{ev}.

Mutual Beliefs for a set of individuals A for an evaluator ev:

$$MB_{A,ev} = \{b \colon b \text{ is a belief } \& \ \forall y \in A\ (\exists k\ (\forall \langle i,g \rangle \in BV_{b(k)}\ (\|k\|^{M_{ev},i,g} = y \ \&\ \|b(k)\|^{M_{ev},i,g} = 1)))\}$$

Common Ground for a set of individuals A for an evaluator *ev*:

$CG_{A,ev}$ = the largest subset X of $MB_{A,ev}$ such that

$$\forall b \in MB_{A,ev}\ (b \in X \rightarrow \Lambda x[^{\phi}bel_k(k)(b(x))] \in X)$$

$$[\![\textbf{\textit{believe}}]\!]^{\mathbf{j,h}} = \Lambda S[\lambda x_e[believe_j([\![S]\!]^{j,h})(x)$$

$\|\mathbf{believe_j(p)(a)}\|^{\mathbf{M,i,g}} = 1$ iff

$\exists b\ \exists q\ (beliefs_j(b,\|a\|^{M,i,g})\ \&\ beliefs_j(q,\|j\|^{M,i,g})\ \&\ SUBSET_j(q[j])$
$\&\ (infer_j(b[j] \cup q[j], p)\ \&\ \neg infer_j(q[j], p))$

(For $b = \{b_1,\ldots,b_n\}$, $b[j] = \{b_1(j),\ldots,b_n(j)\}$)

References

Anand, Pranav. 2006. De De Se. Cambridge, MA: Massachusetts Institute of Technology dissertation. https://doi.org/1721.1/37418.

Bach, Kent. 1997. Do belief reports report beliefs? *Pacific Philosophical Quarterly* 78. 215–241. https://doi.org/10.1111/1468-0114.00036.

Blumberg, Kyle & Ben Holguín. 2018. Ultra-liberal attitude reports. *Philosophical Studies* 175. 2043–2062. https://doi.org/10.1007/s11098-017-0949-7.

Blumberg, Kyle & Harvey Lederman. 2020. Revisionist reporting. *Philosophical Studies*. https://doi.org/10.1007/s11098-020-01457-4.

Burge, Tyler. 1978. Belief and synonymy. *The Journal of Philosophy* 75(3). 119–138. https://doi.org/10.2307/2025424.

Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.

Cresswell, Maxwell & Arnim von Stechow. 1982. *De re* belief generalized. *Linguistics and Philosophy* 5. 503–535. https://doi.org/10.1007/BF00355585.

Darde, Jean-Noël. 1982. *Le journal *l'humanité* et les événements du cambodge de 1975 a 1979: Les stratégies discursives d'une palinodie*. Université de Paris VIII dissertation.

Davidson, Donald. 1968. On saying that. *Synthese* 19. 130–146. https://doi.org/10.1007/BF00568054.

Evans, Gareth. 1979. Reference and contingency. *The Monist* 62. 161–189. https://doi.org/10.5840/monist197962220.

von Fintel, Kai & Irene Heim. 1997–2021. *Intensional Semantics*. MIT. https://github.com/fintelkai/fintel-heim-intensional-notes.

Frege, Gottlob. 1892. Über sinn und bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100. 25–50. Translated as 'On Sense and Reference' in Peter T. Geach and Max Black, *Translations from the Philosophical Writings of Gottlob Frege*, Blackwell, Oxford, 1960.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hintikka, Jaakko. 1962. *Knowledge and belief: An introduction to the two notions*. Ithaca, N.Y.: Cornell University Press.

Kaplan, David. 1968. Quantifying in. *Synthese* 19. 178–214. https://doi.org/10.1007/BF00568057.

Kripke, Saul. 1979. A puzzle about belief. In A. Margalit (ed.), *Meaning and use*. Dordrecht: Reidel. https://doi.org/10.1007/978-1-4020-4104-4_20.

Kripke, Saul. 1980. *Naming and necessity*. Princeton University Press.

Lasersohn, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28. 643–686. https://doi.org/10.1007/s10988-005-0596-x.

Lewis, David. 1979. Attitudes *de dicto* and *de se*. *The Philosophical Review* 88. 513–543. https://doi.org/10.2307/2184843.

Muñoz, Patrick. 2019. *Behavioral attitudes in hyperintensional grammar*. Ms., University of Chicago.

Percus, Orin. 2000. Constraints on some other variables in syntax. *Natural Language Semantics* 8. 173–229. https://doi.org/10.1023/A:1011298526791.

Putnam, Hilary. 1975. The meaning of 'meaning'. In Keith Gunderson (ed.), *Language, mind, and knowledge*. Minneapolis: University of Minnesota Press. https://doi.org/11299/185225.

Quine, Willard van Orman. 1956. Quantifiers and propositional attitudes. *Journal of Philosophy* 53(5). 177–187. https://doi.org/10.2307/2022451.

Schwager, Magdalena. 2009. Speaking of qualities. *Semantics and Linguistic Theory (SALT)* 19. 395–412. https://doi.org/10.3765/salt.v19i0.2534.

Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25. 701–721. https://doi.org/10.1023/A:1020867916902.

Stephenson, Tamina. 2007. *Towards a theory of subjective meaning*. Cambridge, MA: Massachusetts Institute of Technology dissertation. https://doi.org/1721.1/41695.

Tancredi, Christopher. 2007a. *A multi-model modal theory of I-semantics, Part I: Modals*. https://www.semanticsarchive.net/Archive/WFkYjU1Y/. Ms., University of Tokyo.

Tancredi, Christopher. 2007b. *A multi-model modal theory of I-semantics, Part II: Identity and attitudes*. https://semanticsarchive.net/Archive/jkxMDI2O/. Ms., University of Tokyo.

Tancredi, Christopher & Yael Sharvit. 2020. Qualities and translations. *Linguistics and Philosophy* 43(3). 303–343. https://doi.org/10.5840/teachphil202085125.

Christopher Tancredi
Keio Institute of Cultural and Linguistic Studies
2-15-45 Mita, Minato-ku, Tokyo-to, 108-8345 Japan
cdtancredi@gmail.com

Yael Sharvit
UCLA
Department of Linguistics
3125 Campbell Hall
Los Angeles, CA 90095
ysharvit@gmail.com