Convetion-Shifting in Conditionals and Mixed Quotation

Shintaro Matsuura*

02/15/2022

Abstract

In this paper, I offer an analysis of convention-shifting phenomena based on the extension of two-dimensional semantics.[1]

1   Introduction

Consider the following question.

(1)   If 'leg' meant what 'tail' means, how many legs would a horse have?

What is the answer to this question? One might answer with four since the change in the way of using the word 'leg' does not affect the biological property of a horse. Indeed some theorists have argued that the correct answer to the question is four (Predelli, 2014). However, there is another possible answer to the question—one, as some theorists suggest (Israel & Perry, 1996; Stalnaker, 1978).

> The answer, of course, is four, since calling a tail a leg does not make it one, but one can see a different way to take the question. (Stalnaker, 1978, p. 80)

*Depertment of History and Philosophy of Science, The University of Tokyo.
email: smatsuura314@gmail.com

[1]I plan to discuss citations as well, as the title includes 'mixed quotation', but they are not included in this manuscript. However, if time permits, I will briefly discuss mixed quotations in my presentation.

When one answers one to this question, one admits that the *if*-clause in (1) brings the shift of the meaning of 'legs' in the consequent. That is, the interpretation of 'leg' is changed to mean tail, and this change makes it the case that a horse would have one leg. On the contrary, people who answer four deny that the *if*-clause does shift the meaning of 'leg'.[2]

Following Stalnaker, I assume that both readings are available and especially focus on the latter—*convention-shifting*—reading of this question. The focus of this paper is such convention-shifting phenomena. In the next section, I introduce some examples in which the meaning of a particular word is shifted from the actual convention or the meaning the speaker associated with the word. I develop the two-dimensional semantics to deal with convention-shifting in Section 3. In Section 4, I give an analysis of convention-shifting.

2 Convention-Shifting and Examples

In this section, I introduce some examples containing convention-shifting. I focus on two kinds of cases: one concerns conditionals, other concerns mixed quotation.

2.1 Conditionals

Recently, it has been argued that convention-shifting is observed in some forms of conditionals (Kocurek et al., 2020; Rabern, 2020). In particular, convention-shifting in conditionals is closely connected with the phenomena called *metalinguistic disputes* or *metalinguistic negotiation* (Plunkett & Sundell, 2013).

Consider these well-discussed examples of metalingusitic negotiation.

(2) A: Secretariat is an athlete.
    B: No, Secretariat is not an athlete.   (Ludlow, 2008, p. 118)

(3) A: Pluto is a planet.
    B: Pluto is not a planet.   (Plunkett & Sundell, 2013, p. 17)

(2) was taken from a conversation held during a radio program in which participants talked about Secretariat, an excellent racehorse. According to

---

[2]Interestingly, as Takaya (2021b, p.36) notes, it is more likely to be deemed an appropriate response to answer with one to the same question in Japanese.

Plunkett & Sundell (2013), A and B in (2) 'mutually know all of the facts about Secretariat's speed, strength, etc., and what races, awards, metals he won, etc.'(Plunkett & Sundell, 2013, p. 16), but they express different views about how the expression 'athlete' is used or what it means. For example, we can assume that speaker A thinks that 'athlete' applies to non-human animals, while B thinks it never does. That is, they associate different linguistic meanings or conventions with the expression 'athlete'. Consequently, they disagree about the meaning of 'athlete' and also engage in the metalinguistic dispute about 'athlete'. In (3), the relevant facts about Pluto (its size, orbit, and so on) are mutually known to speakers A and B. However, the participants disagree about the meaning of 'planet' and negotiate how to use this word—that is, whether or not Pluto falls under the extension of 'planet'. In the above examples, two speakers are disputing over the meaning of 'athlete' or 'planet' without explicitly mentioning these words.

In these conversational settings, some conditionals are given convention-shifting reading. Suppose, for instance, speaker B in (2) said the following conditional.[3]

(4)   If Secretariat is an athlete, one of the greatest athletes is a horse.

Since speaker B believes that the expression 'athlete' is only applicable to human beings, the occurrence of 'athlete' in the consequent of this conditional must be interpreted as A's preferred meaning of 'athlete', which includes non-human animals in its extension. That is, in order to gain the non-vacuous true reading of (4), the antecedent of (4) should work as a convention-shifter (Rabern, 2020, p. 525).

With regard to (2), similar conditionals might also emerge. Kocurek et al. (2020, p. 5) give such a dialogue between A, who disagrees with the International Astronomical Union (IAU)'s decision, and B, who advocates the IAU's definition of 'planet'.

(5)   A: Pluto is a planet.
        B: No, it's not. Pluto is not a planet because it does not clear its orbital neighborhood.
        A: I don't accept the IAU's definition! Pluto is a planet, I don't care what the IAU says.

---

[3](4) is a modified version of the example in Rabern (2020, p. 525).

B: Look, I know that you think that Pluto is a planet, but there's a good reason the IAU disagrees. If Pluto were a planet, there would be dozens of planets in the solar system.

The conditional B utters in this dialogue contains convention-shift.

(6) If Pluto were a planet, there would be dozens of planets in the solar system.

A natural reading of this conditional is that If the conventional meaning of 'planet' is the previous one, which counts Pluto as a planet, many celestial bodies are also classified as a planet. Since B already agrees to use the word 'planet' in accordance with the IAU's definition, the antecedent of (6) must work as a convention-shifter in order to explain this natural reading.

While the examples like (4) and (6) do not contain quotations, there are also cases in which the particular word whose conventional meaning is shifted is explicitly quoted.

(7) If 'water' referred to gasoline, water would fuel fire.

(Kocurek et al., 2020, p. 19)

(8) If 'leg' means tail, a dog would have one leg.

(Rabern, 2020, pp. 525–526)

(9) If 'leg' meant what 'tail' means, a dog would have one leg.

In the above examples, the antecedents, which contain the explicit quotation, bring the shift in the linguistic meaning of 'water' and 'leg'.

2.2 Mixed Quotation

Convention-shifting phenomena are also observed in sentences with mixed quotations.

3 Two-Dimensional Hypersemantics

In this section, I further develop two-dimensional semantic framework as a preliminary for giving a semantics of convention-shifting. The resulting semantic framework, which I call two-dimensional hypersemantics, can offer a semantics of semantic terms.

3.1 The Basic Idea of Two-Dimensional Semantics

The basic idea behind two-dimensional semantics is that the extension of an expression is relativized with two different possible worlds—contexts of utterance and indices. For a sentence, the content or intension of the sentence is determined by the context in which the sentence is uttered. The content, which is usually a possible world proposition, is evaluated its truth value in the actual or counterfactual situation called indices.

In the following, I use $[\![\alpha]\!]^{c,i}$ to denote the semantic value of α as uttered in a context of utterance c and evaluated with an index i. As Kaplan (1989) argues, this idea enables us to comprehend the semantic behavior of indexicals. For example, the meaning of the indexical 'I', which is directly referential, is represented as follows.[4]

(10)   a. $[\![\, I \,]\!]^{c,i} = $ the speaker in $c = a_c$

For a sentence ϕ, the content the sentence expresses in a context of utterance c is given as $[\![\phi]\!]^c$:

(11) $[\![\phi]\!]^c = \{i \mid [\![\phi]\!]^{c,i} = 1\}$

Thus, if Bob utters 'I am a pianist' the content of Bob's utterance expresses the following content.

(12) $[\![\, I \, am \, a \, pianist \,]\!]^c = \{i \mid$ Bob is a pianist in $i\}$

Two-dimensional or double-indexed semantics works well for a varieties of linguistic phenomena. However, in order to treat convention-shifting phenomena, the possibility that a non-indexical expression, such as 'planet' or 'athlete', would mean something different from its actual meaning should be considered. Among the various two-dimensionalist approaches, Stalnaker's metasemantic interpretation of two-dimensional semantics explicitly admits such use of two-dimensional framework (Stalnaker, 1978, 2004). For contexts of utterance—-worlds considered as actual in Stalnaker's terminology—can include the different possibilities about what linguistic facts are obtained. Furthermore, the two-dimensional intension associated with a sentence is sensitive to alternative linguistic conventions.

---

[4]Though I sometimes use *italics* for emphasis, the difference between mention use of italics and emphatic use of italics should be clear from the context.

Indeed, Stalnaker argues about the example concerning co-referential proper names 'Hesperus' and 'Phosphorus'. Since both proper names co-refer Venus in the actual world, the sentence 'Hesperus is Phosphorus' expresses the proposition that Venus is identical to Venus. On the other hand, if the sentence is uttered in the world in which 'Hesperus' refers to Venus, but 'Phosphorus' refers to Mars, the content determined by this context of utterance is the proposition that Venus is identical to Mars. Though Stalnaker does not formally elaborate on this idea, I further develop this idea in the next subsection.

3.2   Adding Convention as a Parameter

I expand the notion of contexts of utterance and indices in two-dimensional framework to explicitly include parameters other than worlds. I add an interpretation in a model or a linguistic convention itself as the parameter m for a context of utterance and index. I call m a *convention* parameter.[5] Thus, a context of utterance $c = \langle w_c, m_c \rangle$ is a pair of a possible world and a convention, and an index $i = \langle w_i, m_i \rangle$ is also a pair of a possible world and a convention.

More precisely, a convention m is defined as a function from expressions to their content, that is, a function from indices to extensions. I, however, assume m is defined only for proper names and non-context-sensitive predicates.[6] That is, let D be a domain, a set of individuals, a convention m is a

---

[5]This idea might be traced back to Kaplan's logic of demonstratives *LD*. In Kaplan's formal system, the content of Γ, $\{\Gamma\}_{cf}^{\mathfrak{A}}$ in Kaplan's notation, is relativized not only with a context of utterance c and assignment function f but also structure $\mathfrak{A}$, which partly amounts to m here. In particular, structure $\mathfrak{A}$ is universally quantified when defining validity in *LD* (Kaplan, 1989, p. 549). Russell (2008) also proposes a similar function called reference determiner in her own three-dimensional semantics, which is developed to define analytic truth. Recently other theorists also have appealed to the same technique of adding parameters for convention (Krifka, 2013; Kocurek et al., 2020; Mena, 2022). Still, my use of m is slightly different from their use of m in that I regard m as an element in contexts of utterance. Those theorists, on the contrary, regard m only as an element of indices, not contexts of utterance. However, as I will discuss later, there is an independent reason to posit m in contexts of utterance.

[6]Remarking on the definition of m, first, I do not consider the possibility that functional words or logical expressions would have different meanings, such that *and* means disjunction. Second, m is not defined for context-sensitive expressions because it is already assumed in semantic theory that Kaplanian characters and some elements in contexts of utterance determine different contents of context-sensitive expressions in each context of

function which assigns α

   (i) an individual $m(\alpha)$ in D if α is a name;

   (ii) a function $m(\alpha)$ from index to n-tuple in D^n if α is a n-place predicate.

  Here, (i) says that a proper name α under convention m has the same extension $m(\alpha)$ in every index, reflecting the idea that proper names are directly referential and also rigid designators. Thus, the semantic value of α in each cell is given as follows:

(13) $[\![\alpha]\!]^{\langle w_c, m_c\rangle, \langle w_i, m_i\rangle} = m_c(\alpha)$, for any $i(= \langle w_i, m_i\rangle)$.

For example, if we consider the actual and counterfactual convention as follows,

(14) a. $m_1(Phosphorus) = $ Venus

       b. $m_2(Phosphorus) = $ Mars

then,

(15) a. $[\![\, Phosphorus \,]\!]^{\langle w_1, m_1\rangle, \langle w_i, m_i\rangle} = $ Venus, for any i

       b. $[\![\, Phosphorus \,]\!]^{\langle w_2, m_2\rangle, \langle w_i, m_i\rangle} = $ Mars, for any i

are held. Regarding (ii), in general, it gives the following denotation in each cell:

(16) $[\![\alpha]\!]^{\langle w_c, m_c\rangle, \langle w_i, m_i\rangle} = m_c(\alpha)(\langle w_i, m_i\rangle)$.

Let me give another example for a predicate. Suppose the two conventions m_{c_1} and m_{c_2} are associated with *chips*: m_{c_1} associate *chips* with the property of being french fries, and m_{c_2} associate *chips* with the property of being potato chips. For any i, the extension in each cell is given as follows.

(17) a. $[\![\, chips \,]\!]^{\langle w_c, m_{c_1}\rangle, \langle w_i, m_i\rangle} = m_{c_1}(chips)(\langle w_i, m_i\rangle)$

                                  $= \lambda x.\, x$ is part of french fries in $\langle w_i, m_i\rangle$

                                    $= \{x \mid x$ is part of french fries in $\langle w_i, m_i\rangle\}$

       b. $[\![\, chips \,]\!]^{\langle w_c, m_{c_2}\rangle, \langle w_i, m_i\rangle} = m_{c_2}(chips)(\langle w_i, m_i\rangle)$

                                    $= \lambda x.\, x$ is part of potato chips in $\langle w_i, m_i\rangle$

                                    $= \{x \mid x$ is part of potato chips in $\langle w_i, m_i\rangle\}$

---

utterance. Third, in principle, it is possible to extend m as a function from all expressions to the Kaplanian character $\lambda c \lambda i.\, [\![\alpha]\!]^{c,i}$. However, I leave this for future work.

Note that m_i in $\langle w_i, m_i \rangle$ is irrelevant to determine the extension, namely a set of french fries or potato chips. When the token of *chips* uttered by an American is evaluated at some pub in the UK, the fact that *chips* means french fries in that pub does not affect which entities are in the extension of *chips* (the set of potato chips).[7] As this example shows, the convention parameter m_i in $\langle w_i, m_i \rangle$ basically plays no role in evaluation of extensions. It is just that by including the convention parameter in the index, the linguistic aspect of the index is explicitly represented, and a finer-grained description of the index is provided.[8]

This extended two-dimensional framework is beyond standard semantics in that the former has much more expressive power than the latter by explicitly parameterizing linguistic convention itself in contexts of utterance. Therefore, I call it two-dimensional *hypersemantics*.[9]

3.3 A Semantics of Semantic Terms

One of the distinctive features of two-dimensional hypersemantics is that it enables us to construct a semantics of *semantic terms*. Semantic terms are expressions that are used to denote the relationship between an expression and its meaning. Below are examples of sentences in which semantic terms are used, namely *refer* and *apply*.

(18) 'Phosphorus' refers to Venus.

(19) 'Athlete' applies to Secretariat.

---

[7] This is the analogous thing Kripke (1980) discusses using the examples of *gold* and *cat*. Epstein argues the relevant point for social kinds (Epstein, 2015, pp. 77–80).

[8] However, there are some exceptions. If semantic terms, such as *mean*, *refer*, and *true*, are contained in the object language, m_i in the index determines the extension. I will consider some of these words in the next subsection.

[9] This poses a question about the relationship between w and m. According to Stalnaker's original explanation, m is determined by w in accordance with the standard externalist story. More precisely, I assume here that there is flexibility between the two parameters to some extent. That is, it is possible that a world w is paired with multiple different conventions m in my view, as indicated in the *chips* example. Some theorists, like Einheuser (2006, pp. 467–468), assume there is a unique m for each w. However, it is far from obvious that there is such a functional relation between the two parameters. For, in general, a world is not likely to determine unique m, because what determines the content of an expression would be something more specific than worlds, such as facts, state of affairs, or situations. Kocurek et al. (2020, fn. 30) also indicates this point.

In addition to these examples, there are also other semantic terms like *denote*, *count as*, *mean*, and *express*, and so on.[10]

Convention parameters m_i in the index i are important when considering a semantics of these words, though they do not play a substantive role in the previous examples of *Hesperus* and *chips*. This is because the convention parameter in an index is relevant to determining a sentence's truth value when the sentence contains a semantic term.

Consider the expression *refer to* and (18). For the semantics of pure quotations, I assume the disquotational theory, according to which an enquotation is a syntactic operation and the semantic value of a quotation is the very expression quoted itself.[11] Now, I stipulate the expressions in the language are in the domain D.

(20) For any expression α, $[\![\,'\alpha'\,]\!]^{c,i} = \alpha$

---

[10]There are several things that should be noted about the discussion below. First, these expressions have other usages, so my following discussion should be taken to analyze one aspect of the usage of these words. Second, this project might seem circular and odd since, in standard model-theoretic (representationalist) semantics, the notion of reference (or denotation) precedes the notion of truth. Thus, the truth-condition of a sentence is composed based on what the words in the sentence refer to in the semantic model. Of course, it would be important to consider what the project implies about the relation between semantic notions, but I set aside these concerns here. Third, this project might have a potential problem, which also concerns the second point. It is well-known that if one includes the truth predicate and names of sentences of the object language in that object language, the semantic paradox—the liar paradox—arises. Hence, the two levels of languages—the object languages and the metalanguage—should be distinguished to avoid the object language being semantically closed. Though I do not explicitly admit such expressions in the object language, the sentences I analyze below are the sentences belonging to the metalanguage in the standard semantics. I am uncertain as to whether this poses some similar paradox, but I also set aside this potential worry here.

[11]Note that the disquotational theory of quotation above is not compositional semantics, as pointed out in Takaya (2021a, p. 53) and Rabern (2022, pp. 4–5). This fact is easily confirmed. Let α, β be the expressions, such that they are distinctive expressions but $[\![\alpha]\!]^{c,i} = [\![\beta]\!]^{c,i}$. Suppose the disquotational semantics of quotation is compositional. In that case, it must be the case that $[\![\,'\alpha'\,]\!]^{c,i} = [\![\,'\beta'\,]\!]^{c,i}$ because the same syntactic rule applies to the expressions whose semantic values are the same, as $[\![\alpha]\!]^{c,i} = [\![\beta]\!]^{c,i}$. However, it is not the case that $[\![\,'\alpha'\,]\!]^{c,i} = [\![\,'\beta'\,]\!]^{c,i}$, since $[\![\,'\alpha'\,]\!]^{c,i} = \alpha$, $[\![\,'\beta'\,]\!]^{c,i} = \beta$, and $\alpha \neq \beta$. But this problem is not so serious because compositionality is recovered by modifying the definition of semantic values, as Takaya demonstrates (Takaya, 2021a, p. 54). Rabern (2022) also proposes a different compositional analysis of pure quotations in response to this problem. Additionally, I set aside the cases of mixed quotations here. For detailed overviews of various approaches to quotation, see Cappelen et al. (2020) and Maier (2014).

In addition, I regard *to* as semantically vacuous; that is, *to* has no semantic contribution to the truth-condition of the whole sentence. Assuming m_c is our actual convention, where *Venus* refers to Venus, the semantic value of *Venus* and *'Phosphorus'* is the below.

(21) $[\![\,Venus\,]\!]^{c,i} = m_c(Venus) = \text{Venus}$

(22) $[\![\,'Phosphorus'\,]\!]^{c,i} = Phosphorus$

In the previous examples, *Phosphorus* refers to Venus under the convention m_1 and refers to Mars under m_2. Thus, the followings are expected to be held.

(23) a. $[\![\,'Phosphorus'\ refers\ to\ Venus\,]\!]^{c,\langle w_1,m_1\rangle} = 1$

b. $[\![\,'Phosphorus'\ refers\ to\ Venus\,]\!]^{c,\langle w_2,m_2\rangle} = 0$

That is, the intuitive meaning of (18) is that the referent of *Phosphorus* in $\langle w_i, m_i\rangle$ is Venus.

(24) $[\![\,'Phosphorus'\ refers\ to\ Venus\,]\!]^{c,i} = 1 \iff \text{Venus} = m_i(Phosphorus)$

From this consideration, the denotation of the 2-place predicate *refer* can be given as follows. The meaning of *refer* is a function that takes an entity as an input and returns a function that takes an expression as an input and returns 1 if the meaning of that expression in the convention m_i in index i is the entity and returns 0 otherwise.[12]

---

[12]More generally, the expression quoted in *refer* sentences might be a well-formed complex expression which is not a member of $\mathrm{dom}(m)$ because m is defined only for simple expressions. In order to accommodate these cases, the lexical entry of *refer* can be modified as follows.

(i) $[\![\,refer\,]\!]^{c,i} = \lambda x \in D.\,\lambda y \in D^{\mathcal{L}}.\,x = [\![y]\!]^{c_i,i}$

Here, $D^{\mathcal{L}} \subset D$ is a set of well-formed expressions whose semantic value is properly defined, and c_i is a context of utterance in which the value of the convention parameter is m_i. As indicated in the denotation of *refer*, this is a monstrous expression since the context of utterance in $[\![y]\!]^{c_i,i}$ is shifted. In addition, if the quoted expression is context-sensitive, the treatment would be more complicated since c_i would be the context of utterance whose value of other contextual parameters might be shifted. For instance, c_i's value of the agent parameter would be the agent a_i in the counterfactual situation i.

(ii) 'I' refers to the speaker.

(iii) $[\![\,'I'\ refers\ to\ the\ speaker\,]\!]^{c,i} = 1 \iff [\![\,the\ speaker\,]\!]^{c,i} = [\![\,I\,]\!]^{c_i,i}$
$\iff$ the person speaking at $w_i = a_c$
$\iff a_i = a_c$

(25) $[\![\ refer\ ]\!]^{c,i} = \lambda x \in D.\ \lambda y \in \mathrm{dom}(m).\ x = m_i(y)$
$\qquad\qquad = \{\langle x, y \rangle \mid x = m_i(y)\}$

(26) $[\![\ refer\ to\ Venus\ ]\!]^{c,i} = \lambda y \in \mathrm{dom}(m).\ \mathrm{Venus} = m_i(y)$
$\qquad\qquad\qquad = \{y \mid \mathrm{Venus} = m_i(y)\}$

Therefore, when α is a proper name and β is a singular term, the truth-condition of the sentence "α' refers to β' is represented as follows.

(27) $[\![\ '\alpha'\ refers\ to\ \beta\ ]\!]^{c,i} = 1 \iff [\![\beta]\!]^{c,i} = m_i(\alpha)$

In the cases in which α and β are simple predicates, the same analysis would be given.[13]

Semantic analysis of other semantic terms is obtained in an analogous way. As another example, consider a sentence in which the expression *apply to* is used. Since we assume the actual convention m_c assigns Secretariat, a superior racehorse, to the word *Secretariat*, the truth-condition of the sentence *'athlete' applies to Secretariat* is that Secretariat falls under the extension of the predicate *athlete*.

(28) $[\![(19)]\!]^{c,i} = 1 \iff \mathrm{Secretariat} \in m_i(athlete)(\langle w_i, m_i \rangle)$

In general, when α is a type $\langle e, t \rangle$ predicate and β is a singular term, the truth-condition of the sentence 'α' *applies to* β is given as follows.[14]

---

[13]Note that since c is an arbitrary constant, it is not guaranteed that c is the actual world. However, since what we want to focus on here is the particular use of *refer* that links the linguistic expression to its meaning, I leave aside for the moment the possibility that *refer* means something else. Thus, in what follows, I will consider only the intended meaning of such terms as semantic terms for other semantic terms in general. For instance, I restrict the contexts of utterance to be considered to the class of c, such that the following is held.

   (i) $m_c(refer) = \lambda i.[\lambda x \in D.\ \lambda y \in \mathrm{dom}(m).\ x = m_i(y)\ \text{in}\ i\,]$

[14]Of course, there are other cases in which *is applied to* is used as a semantic term. If α is a complex predicate, the truth-condition is written as follows.

   (i) $[\![\ '\alpha'\ applies\ to\ \beta\ ]\!]^{c,i} = 1 \iff [\![\beta]\!]^{c,i} \in [\![\alpha]\!]^{c_i,i}$

where c_i is the same as c_i in footnote 14. Even in cases where β is a bare noun or contains a determiner, I think a similar analysis would be given based on the idea here by appropriate revision, such as adding a type-shifting rule or mereological sum in domain D.

   (ii) 'Athlete' applies to racehorses.

   (iii) 'Athlete' is applied to a cat.

(29) $\llbracket$ 'α' applies to β $\rrbracket^{c,i} = 1 \iff \llbracket\beta\rrbracket^{c,i} \in m_i(\alpha)(\langle w_i, m_i\rangle)$

4 Analysis of Conditionals with Convention-Shifting

In this section, I give an analysis of the conditionals in which a convention-shift occurred, using the two-dimensional hypersemantics I developed in the previous section.[15]

4.1 Cases without Quotation

Recall the cases introduced in Section 2.

(4) If Secretariat is an athlete, one of the greatest athletes is a horse.

(6) If Pluto were a planet, there would be dozens of planets in the solar system.

First of all, it should be confirmed that the standard possible world analysis of conditionals does not capture the intended convention-shifting reading of these conditionals. According to Stalnaker (1968)'s semantics of conditionals, $\ulcorner\phi > \psi\urcorner$ is true in a possible world w if and only if ψ is true in all possible worlds closest to w in which ϕ is true.[16] This analysis is written in the current two-dimensional setting as follows.

(30) $\llbracket\phi > \psi\rrbracket^{c,i} = 1 \iff \llbracket\psi\rrbracket^{c,i'} = 1$,
for all i' closest to i, such that $\llbracket\phi\rrbracket^{c,i'} = 1$

Furthermore, the sentences with other operators, such as (iv), would be analysed based on the idea here.

(iv) 'Athlete' might be applied to Secretariat.

[15]Though Kocurek et al. (2020) also discusses similar phenomena, there are differences between their account and my account here. First, Kocurek et al. (2020) does not appeal to the two-dimensional semantics. To use two-dimensional semantics for these conditionals is suggested in Rabern (2020), but not fully explored in the current literature. Second, Kocurek et al. (2020) does not build the semantics for the conditional involving explicit mention, which I discuss in 4.2.

[16]Stalnaker's semantics applies to both indicative and subjunctive (or counterfactual) conditionals. In addition, I assume here that the convention-shifting reading of a conditional is possible in both forms of conditionals.

Consider applying this semantics to (4). Since the speaker of (4) believes the word 'athlete' is applicable to only human beings, the antecedent and the consequent of (4) are both false. More precisely, the whole conditional would be predicted as vacuously true because there is no possible world in which the antecedent is true, as Secretariat, a racehorse, can never be an athlete. In the case of (6), a similar problem arises. (6) is uttered by the speaker who agrees with the IAU's definition and is trying to convince those who are against the IAU's decision. If (6) is interpreted with the new meaning of 'planet', this counterfactual would be false. For there would not be many planets in the closest worlds to the actual worlds except that Pluto was a planet since the fact that Pluto also meets the new strict condition for being a planet, as determined by the IAU, does not make it the case that there are also many other planets that are not planets in the actual world. Thus the intuitively true reading of (6) is not explained in the standard semantics for conditionals.

The point of these problems is that the Stalnakerian analysis does not assume a shift in the meaning of any particular expression in these examples. Rather, what is shifted in the above analysis is the index i, which evaluates the truth value of the consequent. In order to explain the on-vacuous true reading of these conditionals, the shifted meaning of 'athlete' or 'planet', which makes the antecedent true, must be used to interpret the consequent. That is, the context of utterance, which determines the content expressed by the consequent, has to be shifted.

I propose another truth-conditional analysis of the convention-shifting reading of conditionals, reflecting the above consideration. I use $\ulcorner >_+ \urcorner$ to distinguish the convention-shifting reading from the convention-fixed reading given in (30).

(31) $\quad [\![\phi >_+ \psi]\!]^{c,i} = 1 \iff [\![\psi]\!]^{c',i'} = 1$, for all $\langle c', i' \rangle$ closest to $\langle c, i \rangle$, such that $[\![\phi]\!]^{c',i'} = 1$

(34) says that the parameter c, as well as i, is shifted to make the antecedent ϕ true, and the shifted parameter c' is used as the context of utterance to interpret the consequent ψ. The similarity between the pairs consisting of c and i is naturally calculated based on the similarities between c and c', and between i and i'.

Let me confirm that this clause derives the intended reading of (4) and (6). First, consider again (4). Let $\langle w_1, m_1 \rangle$ be the original context of utter-

ance c in which (4) is uttered. w_1 would be the world in which the relevant facts about Secretariat are held, but m_1 would be the convention according to which only human beings fall under the extension of the expression 'athlete' since the speaker believes that 'athlete' has such meaning in negotiating the meaning of 'athlete'. In addition to $\langle w_1, m_1 \rangle$, we can take the alternative pair $\langle w_1, m_2 \rangle$ as making the antecedent 'Secretariat is an athlete' true. In $\langle w_1, m_2 \rangle$, the worldly facts, especially the facts concerning Secretariat, are also established, but the conventional meaning of 'athlete' provided by m_2 is different from the meaning of 'athlete' in m_1. For instance, these convention parameters give the following extension in each index $i = \langle w_i, m_i \rangle$.[17]

(32) a. $[\![athlete]\!]^{\langle w_c, m_1 \rangle, i} = m_1(athlete)(i)$
$= \lambda x.\, x$ is a physically superior human in i
$= \{x \mid x$ is a physically superior human in $i\}$

 b. $[\![athlete]\!]^{\langle w_c, m_2 \rangle, i} = m_2(athlete)(i)$
$= \lambda x.\, x$ is a physically superior animal in i
$= \{x \mid x$ is a physically superior animal in $i\}$

Given these denotations, $\langle w_1, m_2 \rangle$ would be a plausible c' which makes the antecedent 'Secretariat is an athlete' true. If the proposition that Secretariat is a physically superior animal and the consequent are true in c' and i', the whole conditional is also true. On the other hand, if the original index i is the world, in which there are a lot of greater athletes than Secretariat, the conditional would be false because Secretariat would not be counted as one of the greatest athletes.

The same goes for the counterfactual (6). The original context of utterance c is such that the convention parameter assigns to 'planet' the strict definition which excludes Pluto from the set of planets. However, if the context of utterance is shifted to make 'Pluto is a planet' true, for instance, the previous definition of 'planet', the truth of the whole conditional depends on whether or not there are many celestial bodies that also satisfy the previous definition of 'planet' in i'.

---

[17]The below is just an illustration. In fact, it is often the case that there are many m which make both speakers' utterances express true propositions.

4.2 Cases with Quotation

Aside from the convention-shifting reading of the standard forms of conditional, there are also cases involving the explicit quotation of the word whose meaning is shifted.

(7)   If 'water' referred to gasoline, water would fuel fire.

(8)   If 'leg' means tail, a dog would have one leg.

(9)   If 'leg' meant what 'tail' means, a dog would have one leg.

Intuitively, it is clear how to derive the non-vacuous true reading of these conditionals. However, how to deal with them in semantics would need to be clarified. The semantics analysis of semantic terms I give in 3.3 would be useful to calculate the truth-condition of the cases mentioning the word.

The approach is simple. Consider (7) as an example. Assuming the convention parameter m_c in the context of utterance is the convention, according to which 'gasoline' refers to gasoline, we can think that the antecedent means that the expression 'water' meant the property of being gasoline.

(33) $[\![\, \text{'water' refers to gasoline} \,]\!]^{c,j} = 1 \iff m_j(water) = m_c(gasoline).$

That is, in the counterfactual convention m_i, 'water' refers to gasoline, which is the things expressed by 'gasoline' in the actual linguistic convention. Then, the truth of the consequent is evaluated by using this counterfactual convention as the context of utterance. In other words, the clause for these cases would be the following (I use $\ulcorner >_q \urcorner$ to denote the conditionals in which an expression is mentioned to distinguish it from (30) and (31)):

(34) $[\![\, \phi >_q \psi \,]\!]^{c,i} = 1 \iff [\![\, \psi \,]\!]^{c',i} = 1,$
$\qquad\qquad\qquad$ for all c' closest to c, such that $[\![\, \phi \,]\!]^{c,c'} = 1$

The point is that the parameter of the index would not be shifted as in the standard analysis of conditionals. Rather, in the conditionals which contain mentioning a certain expression, the antecedent aims to stipulate the meaning of a particular expression explicitly. Thus the function of these examples is radically different from the usually discussed examples of conditionals, though they take the exact grammatical construction of conditionals. Following clause (34), one can take j, whose convention parameter satisfies (33) and is minimally different from m_c, as c'.

(35) $[\![\, \text{water would fuel fire} \,]\!]^{c',i} = 1$

If the various physical laws in the actual world are also obtained in i, the consequent is also true since the reference of 'water' in c' is gasoline.

The same explanation goes for the true reading of (8) and (9). the antecedent of these conditionals re-stipulate the convention of 'leg' so as to mean tail. The consequent is interpreted using this new convention.

5 Conclusion

In this paper, I give an analysis of convention-shifting by extending the two-dimensional framework. Though I try to give a two-dimensional analysis of the convention-shifting reading of the conditionals, one might pose the following worry. (31) and (34) describe the different truth-conditions, so that they are not a unified account of convention-shifting. I think this worry is right and there might be a more sophisticated semantic analysis, in which the same clause for the conditionals with quotation and without quotation is given. However, as I briefly mentioned in 4.2, the function of the two kinds of conditionals seems to be quite different. The cases like (7) are understood as a kind of meaning stipulation rather than conditionals.

I briefly mention some future issues that were not covered in detail in this thesis. The two-dimensional hypersemantics presented in Chapter 3 can still be refined in both its foundational and logical aspects. In particular, what it means to give semantic analysis of semantic terms is not clear though it is needed for practical reasons to analyze the sentences involving semantic terms.

References

Cappelen, H., Lepore, E., & McKeever, M. (2020). 'Quotation,' in Zalta, E. N. ed. *The Stanford Encyclopedia of Philosophy*: Metaphysics Research Lab, Stanford University, Summer 2020 edition.

Einheuser, I. (2006). 'Counterconventional conditionals,' *Philosophical Studies*, *127*, 3, 459–482.

Epstein, B. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*: Oxford University Press.

Israel, D. & Perry, J. (1996). 'Where monsters dwell,' in Seligman, J. & Westerståhl, D. eds. *Logic, language and computation*, *1*: CSLI Publication, Stanford, 303–316.

Kaplan, D. (1989). 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals,' in Almog, J., Perry, J., & Wettstein, H. eds. *Themes From Kaplan*: Oxford University Press, 481–563.

Kocurek, A. W., Jerzak, E., & Rudolph, R. E. (2020). 'Against Conventional Wisdom,' *Philosopher's Imprint*, *20*, 22, 1-27.

Krifka, M. (2013). 'Definitional generics,' in Del Prete, F., Beyssade, C., & Mari, A. eds. *Genericity*: Oxford University Press Oxford, 372-389.

Kripke, S. A. (1980). *Naming and Necessity*: Harvard University Press.

Ludlow, P. (2008). 'Cheap contextualism,' *Philosophical Issues*, *18*, 104–129.

Maier, E. (2014). 'Pure quotation,' *Philosophy compass*, *9*, 9, 615–630.

Mena, R. (2022). 'Metalinguistic effects,' *Mind & Language*, 1–21. DOI: 10.1111/mila.12390.

Plunkett, D. & Sundell, T. (2013). 'Disagreement and the semantics of normative and evaluative terms,' *Philosopher's Imprint*, *13*, 23, 1–37.

Predelli, S. (2014). 'Kaplan's three monsters,' *Analysis*, *74*, 3, 389–393.

Rabern, B. (2020). 'Semantic monsters,' in Biggs, S. & Geirsson, H. eds. *The Routledge Handbook of Linguistic Reference*: Routledge, 515–532.

———— (2022). 'Pure quotation in linguistic context,' *Journal of Philosophical Logic*, 1–21. DOI: 10.1007/s10992-022-09675-3.

Russell, G. (2008). *Truth in virtue of meaning: A defence of the analytic/synthetic distinction*: OUP Oxford.

Stalnaker, R. (1968). 'A Theory of Conditionals,' in Harper, W. L., Stalnaker, R., & Pearce, G. eds. *IFS: Conditionals, Belief, Decision, Chance and Time*, Dordrecht: Springer Netherlands, 41–55.

———— (1978). 'Assertion,' in Cole, P. ed. *Syntax and Semantics*, *9*: New York: New York Academic Press, 315–332. reprinted in Stalnaker (1999), 78–95.

———— (2004). 'Assertion revisited: On the interpretation of two-dimensional modal semantics,' *Philosophical Studies*, *118*, 1, 299–322.

Takaya, R. (2021a). 'Compositionality for What?: Methodological Argument from Un-synonymy and Substitution,' *Journal of the Japan Association for Philosophy of Science*, *49*, *1*, 49–63.

——— (2021b). 'Searching for Meanings: Compositionality in Semantics and Metasemantics,' Ph.D. dissertation, Keio University.