Authors' final version. Appears in Sinn und Bedeutung 21, 2017.

Picture Descriptions and Centered Content[1]

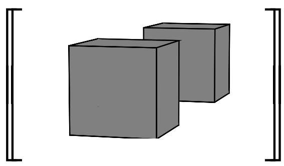
Mats ROOTH — *Cornell University*
Dorit ABUSCH — *Cornell University*

Abstract. There is an argument based on sentences that describe pictures in favor of a viewpoint-centered possible worlds semantics for pictures, over a propositional semantics (J. Ross 1997). The argument involves perspectival lexical items such as "front". We show that when a projective possible worlds semantics for pictures is employed, there is a problem with the argument coming from propositional contents being strong. The argument is reconstructed in a model modal space involving linear worlds, and it is shown that it works there, by computing the possible worlds semantics. The construction involves propositions and centered propositions that are regular sets of strings. Finally, by manipulating the marking parameter in a projective semantics for pictures, the argument is reconstructed also for 3D models.

Keywords: Semantics of pictures, linguistic descriptions of pictures, centered possible worlds.

1. Introduction

Some recent work on the semantics and pragmatics of pictures and pictorial narratives has used a framework where these artifacts, just like sentences, have propositional semantic values that are constructed as sets of possible worlds or possible situations (Greenberg 2011, 2013; Abusch 2012, 2014, 2016). We use Scott brackets for both kinds of semantic values. Just as (1a) designates the semantic value of the sentence inside the brackets, (1b) designates the semantic value of the picture inside the brackets.

(1) a. ⟦there are two cubes⟧
b. 

There are advantages in assuming information contents for pictures and linguistic phrases in the same semantic space. One comes in the analysis of multimodal messages consisting of a picture and sentence, where one wants to combine information from the two media into a whole. Another comes in the semantic analysis of sentences such as (2) that describe pictures. Ross (1997) gave a compositional semantic analysis of such sentences in terms of the propositional semantic value of the mentioned picture, and the propositional semantic value of the prejacent clause beginning with *there*.

(2) In the picture, there is a cube next to an octahedron.

[1]For comments and assistance, we thank audiences at SuB and McGill, Luis Alonso-Ovalle, Mark Steedman, Michael Wagner, and Ede Zimmermann.

Ross then pointed out a problem for this propositional theory having to do with sentences describing pictures where the prejacent clause includes perspectival constructions such as “in front of”, as in (3).

(3) In the picture, there is a cube in front of an octahedron.

Ross addressed the problem by replacing propositional semantic values for pictures and for the prejacent sentence with *viewpoint-centered* semantic values. These are analagous to the agent-centered semantic values in Lewis’s *de se* analysis of attitude semantics (Lewis 1979). The point of this paper is to construct Ross’s argument using the semantic assumptions summarized in the next section, refute it, and then reconstruct it by modifying the semantic framework.

2. Ordinary and centered pictorial contents

Classic and modern treatises on perspective, and contemporary works on computer graphics, describe mathematical recipes for mapping a three-dimensional worlds to pictures (Szeliski, 2010). We proceed here by indentifying three-dimensional worlds with data structures that specify the location, scale, and orientation of geometric objects. (4) is a world with two cubes, and nothing else.

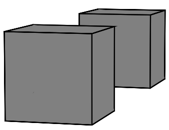
(4) Possible world w_1

type	*scale*	*translation*	*rotation*
cube	1.0	[0,0,0]	[0,0,0]
cube	1.0	[3,0,0]	[0,0,0]

Given in addition a specification of a viewpoint (a certain kind of oriented location) and a “marking rule”, a picture is mathematically determined. The viewpoint has information that determines a family of oriented projection lines and a planar region (corresponding to the picture) in the three-dimensional space. A marking rule determines how points in the picture region are to be colored. Rule R_1 combined with w_1 and a certain viewpoint results in picture (6).

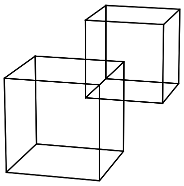
(5) R_1 : Mark a point in the picture plane black if the projection line from the viewpoint through that point intersects the edge of an object before it intersects any other part of an object, otherwise in gray if it intersects some object, and otherwise in white.

(6)

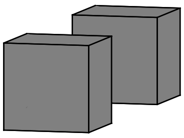


(7) describes the marking rule for a “line drawing”, resulting in pictures such as (8).

(7) R_2: Mark a point of the picture plane in black if the directed projection line intersects the edge of an object, and otherwise in white.

(8) 

Finally, projection lines can be determined in various ways with respect to the viewpoint. Using parallel projection lines instead of lines intersecting at the viewpoint results in an orthographic picture such as (9). Call the projection line parameter G.

(9) 

Summing this framework up, there is a parameterized procedure that determines a picture p from a world w in a space of geometrically constructed worlds $\mathscr{M}$, a viewpoint v, a marking rule R, and a projection-line parameter G. This is summarized in (10).

(10) $p = \Pi(\mathscr{M}, w, v, R, G)$

Starting from a picture p and a fixed $\mathscr{M}$, R and G, a semantic value as a set of worlds is now obtained by inverting projection in the way defined in (11). The semantic value of p is the set of worlds w such that using R and G, w projects to p from some viewpoint.

(11) $[\![p]\!]^{\mathscr{M},R,G} = \{w | \exists v. p = \Pi(\mathscr{M}, w, v, R, G)\}$

Alternatively, instead of existentially quantifying the viewpoint, the semantic value of a picture can be defined as a set of world-viewpoint pairs. This is a set of viewpoint-centered worlds, analogous to the believer-centered worlds discussed by Lewis (1979). The viewpoint-centered semantic value is defined in (12).

(12) $[\![p]\!]_{\triangle}^{\mathscr{M},R,G} = \{\langle w, v\rangle | p = \Pi(\mathscr{M}, w, v, R, G)\}$

3. Picture descriptions and Ross's argument

The analysis in Ross (1997) uses the formalization (13b) for sentence (13a). The operator $[x]$ is a modal necessity operator based on the propositional content of x. The formula $[x]\phi$ is true if and only if for every world w in the propositional content of x, formula ϕ is true in w.

(13) a. In one picture, there is a man on a couch.

b. $\exists x.\mathbf{picture}(x) \wedge [x]\exists y \exists z[\mathbf{man}(y) \wedge \mathbf{couch}(z) \wedge \mathbf{on}(y,z)]$

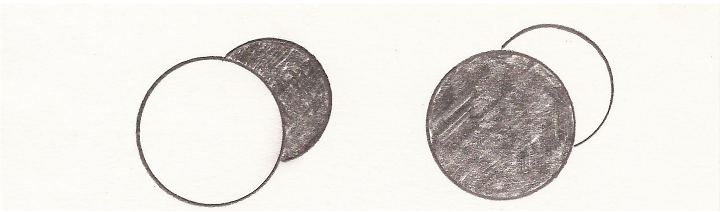
This semantics for picture descriptions is isomorphic to the subset semantics for belief descriptions. Sentence (13a) is true in a world w_0 if and only if there is an x s.t. x is a picture in w_0, and for all worlds w in $[\![x]\!]^{\mathscr{M},R,G,w_0}$, there is a y and a z such that y is a man in w, z is a couch in w, and y is on z in w. A nice aspect of the analysis that it uses the general propositional semantics for the prejacent sentence *there is a man on the couch* in (13a). There is no need to refer to a semantics for the prejacent sentence that is specific to the pictorial medium.

Ross's argument for centered contents has to do with the truth or falsity of sentence (14), construed as referring either to the picture on the left in (15), or the picture on the right. Intuitively the sentence is true with reference to the picture on the left (Picture 1), and false with reference to the picture on the right (Picture 2).

Suppose the pictures have identical propositional semantic values, along the lines of "there is a white ball and a black ball". We can't get different truth values for the sentences in (16), because the pictures enter into the subset semantics for the in-the-picture construction via their propositional semantic values.

(14) In the picture, there is a white ball in front of a black ball.

(15)



(16) a. In Picture 1, there is a white ball in front of a black ball.
b. In Picture 2, there is a white ball in front of a black ball.

What goes wrong? Ross pointed out that the problem comes up when the prejacent sentence contains an element such as *in front of*, the semantics of which is sensitive to a perspective. Note that the sentences in (17), where there is no perspectival lexical item, are not problematic like the sentences in (16), because they do have the same truth value.

(17) a. In Picture 1, there is a white ball next to a black ball.
b. In Picture 2, there is a white ball next to a black ball.

This can be related to an independently motivated perspectival parameter in the semantics of *in front of*. Suppose Keisha uses sentence (18a) to tell Justin where his bike is. The information conveyed is similar to what is conveyed by (18b), but stronger in that Justin gets the information that the bike rack is between the oak and the route, not simply near the oak. This motivates the hypothesis that *in front of* includes a covert perspectival parameter—it is understood as *in front from a perspective on the route*.

(18) a. On the route to school, there is bike rack in front of a big oak. The bike is locked there.
b. On the route to school, there is bike rack next to a big oak. The bike is locked there.

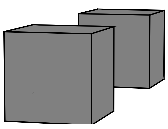
This kind analysis of *in front of* is well motivated, and there is substantial literature on it (Kemmerer, 2006; Kelleher and van Genabith, 2006). Ross used it in a solution to the puzzle of the balls. The first step is to use a viewpoint-centered semantics for pictures, as introduced in Section 2. The second step is to hypothesize that the construction in-x-ϕ binds the viewpoint parameter of the prejacent sentence ϕ. The result is that both the picture x and the prejacent ϕ contribute viewpoint-centered propositions. The semantics for in-x-ϕ does the subset check for these viewpoint-centered propositions, rather than ordinary propositions as before.

(19) Semantics for [in x, ϕ]
For all $\langle w, v\rangle$ in $[\![x]\!]^{\mathscr{M},R,G,w_0,g}$, $[\![\phi]\!]^{w,g[v0\mapsto v]} = 1$.

Notice now that the two pictures in (15) have different viewpoint-centered contents. In centered worlds $\langle w, v\rangle$ in the content of Picture 1, there is a white sphere and that is closer to the viewpoint than a black sphere. The reverse is true of centered worlds in the content of Picture 2.

4. A problem with strong pictorial contents

We want to ultimately agree with the argument summarized in Section 3, and with the conclusion. However, there is a problem. Ross assumes that pictures 1 and 2 have identical propositional contents. The same assumption is made in subsequent literature (Blumson, 2010). However, this runs afoul of the fact that pictorial contents as obtained in the projective theory are in some respects strong. For instance, the propositional content of the picture in (20a) is stronger than the propositional content of the sentence in (20b). This is shown by the fact that the possible world w_2 given in (21) is in the content of the sentence, but not in the content of the picture. In w_2, there are two cubes, so (20b) is true. But there are no edges from different cubes that are parallel. In worlds consistent with picture (20a), there are two cubes with pairs of parallel and indeed co-linear edges, assuming perspectival projection. In general, any picture of two cubes gives information about the orientation of the cubes, and there is no picture that entails sentence (20b) and has no additional entailments.

(20) a. 

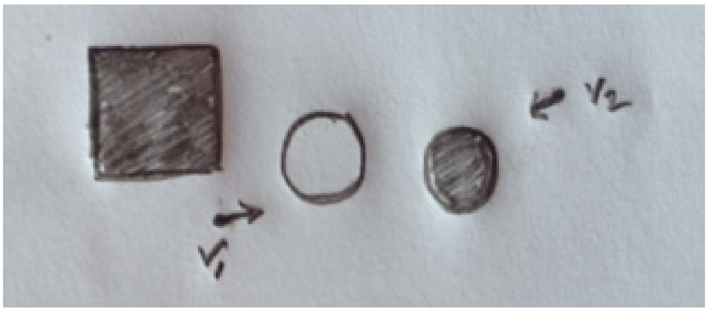
b. There are two cubes.

(21) Possible world w_2

type	*scale*	*translation*	*rotation*
cube	1.0	[0,0,0]	[0,0,0]
cube	1.0	[3,0,0]	[1,1,1]

In the same way, the content of the pictures 1 and 2 in (15) is stronger than the proposition that there is a black sphere and a white sphere. One additional entailment in this case is that there are no *other* objects visible from particular viewpoints. Consider a world w_3 with exactly three objects, which are a white sphere, a black sphere, and a black cube. The three are in line, with the white sphere in the middle. See (22).

(22) World w_3:



From the viewpoint v_1, the cube is not in view, and and $\Pi(\mathscr{M}, w_3, v_1, R, G)$ is p_1. However, from v_2, the black cube is in view in the background, and $\Pi(\mathscr{M}, w_3, v_1, R, G)$ is not p_2. In fact, there is no viewpoint v such that $\Pi(\mathscr{M}, w_3, v, R, G)$ is the picture p_2. Since the propositional denotation of p_2 is $\{w|\exists v \Pi(\mathscr{M}, w_3, v, M, G) = p_2\}$, w_3 is not an element of $[\![p_2]\!]^{\mathscr{M},R,G}$. The propositional denotation of p_1 is $\{w|\exists v \Pi(\mathscr{M}, w_3, v, R, G) = p_1\}$, and v_1 is a witness for w_3 being an element of $[\![p_1]\!]^{\mathscr{M},R,G}$. So contrary to what the argument from Section 3 has to assume, p_1 and p_2 do not have the same propositional contents, if propositional contents are defined by geometric projection as described in Section 2.

5. Pictures and projection in lineland

In this section we construct a family of models where Ross's argument does work, because pictures have weak contents that in a certain sense have no extra information. The worlds are "linelands"—worlds that have the form of a string. An additional feature is that the semantics is computable, because the propositions that are the denotations of sentences are regular sets of strings. These are the sets of strings that are representable by regular expressions and by finite state machines.

To illustrate, world s_1 as defined in (23) is a world with (from the left) a ruby, an opal, a picture of a ruby in front of an opal, two opals, and finally a ruby. A "ruby" is the character `r`, and an "opal" is the character `y`. Pictures are delimited by brackets, with a square bracket marking the front of a picture, and round bracket marking the back. Within a picture, the character `b` depicts a ruby, and the character `w` depicts an opal. Thinking of `b` and `w` as black and white, these are black and white pictures. The assumptions are listed in (24).

(23) World s_1

`_r_y_[bw)_y_y_r_`

(24)

r	ruby
y	opal
square bracket	front of picture
round bracket	back of picture
b	ruby in picture
w	opal in picture

The underlined positions seen in (23) are reserved for discourse referents, which are used for compositional semantics. The ultimate discourse referent 1 is used for the most recently mentioned object, and the discourse referent 2 for the penultimately mentioned object. The indexed world in (25) has an ultimate discourse referent for a picture, because 1 immediately precedes the start of a picture. And it has a penultimate discourse referent for a ruby, because 2 immediately precedes r.

(25) World s_2
`2r_y1[bbw)_r_y_r_y_`

Let *SitD* be the set of all indexed situations of this kind, and let *Sit* be the set of all situations without discourse referents. They are defined by terms in an extended language of regular expressions. *Sit* is a set of strings, used as the set of all worlds. Regular subsets of *Sit* are used as propositions, and regular subsets of *SitD* are used as information states in a dynamic semantics. Semantic composition is performed mainly using relation composition, using the relations listed in (26). *New* is the random choice operator of dynamic semantics. *Forget* deletes discourse referents to map to a proposition. The remaining operators are tests. *Ruby* checks that there is a ruby at the ultimate discourse referent, i.e. that r follows 1 in the string.

(26)

New	introduce random discourse referent 1, while demoting 1 to 2.
Forget	delete discourse referents
Ruby	check that 1 is an ruby
Opal	check that 1 is an opal
Pict	check that 1 is a picture
Adj	check that 1 is adjacent to 2

In these terms, the proposition denoted by (27a) is defined by (27b). The circle indicates relation composition, or restriction of a relation to a domain or co-domain. R^{CO} is the co-domain of relation R. In combination with *Forget*, it is used to map to a proposition. The dynamic semantics works by inserting 1 in a random location; checking that the object marked by 1 is an opal; inserting 1 in a random location while demoting 2 to 1; checking that the object marked by 1 is a ruby; checking that the objects marked by 1 and 2 are adjacent; and finally forgetting the discourse referents.

(27) a. There is a ruby adjacent to an opal.
b. $[Sit \circ New \circ Opal \circ New \circ Ruby \circ Adj \circ Forget]^{CO}$

These above sets and relations are defined in a language of extended regular expressions. Complete definitions are found in the replication supplement (Rooth and Abusch, 2017). The definitions can be interpreted computationally, using Xfst or Foma (Beesley and Karttunen, 2003; Hulden, 2009), to obtain a computational representation of a propositions such as (27b).

The next issue is the encoding of centered propositions. A center or viewpoint is modeled with the character “>” or “<”, located in the block where discourse referents are also placed. A centered world has exactly one center. Examples are in (28)-(29).

(28) Centered world c_1, with a center looking towards a ruby in front of an opal.

```
r_y_[bbw)_r_y>r_y_
```

(29) Centered world c_2, with a center looking towards an opal in front of a ruby.

```
r_y_[bbw)_r_y<r_y_
```

(30) describes relations that randomly insert a center and delete a center. The basic relation *Front* requires that both 1 and 2 are in the direction indicated by the center, with 1 encountered first.[2]

(30)

NewC	randomly insert a center
ForgetC	delete the center
Front	the center is >, > precedes `1`, and `1` immediately precedes `2`, or the center is <, `2` immediately precedes `1`, and `1` precedes <.

The semantics of (31a) is defined in (31b), without deleting the drefs or the center. (31c) gives a sample of the centered indexed worlds in the centered proposition.

(31)
a. There is a ruby in front of an opal.
b. $[Sit \circ NewC \circ New \circ Opal \circ New \circ Ruby \circ Front]^{co}$
c.

```
>[wb)1r2y_r_r_r_r_y_r_r_y_r_
_(bww]>r_y_r_r1r2y_
2y1r_[bb)<y_r_r_y_
>[bw)1r2y_
_r>(bwwwbbww]_y_y_r_y1r2y_y_
2y1r<y_r_[wb)_y_r_y_y_y_r_r_y_y_
>(wbwbbwb]_y_y1r2y_r_
2y1r<[bwbww)_r_r_
>[wb)1r2y_r_
```

A semantics for pictures is defined in the form of an accessibility relation. Starting from an indexed world *s* with a picture at dref 1, this is done by non-deterministically constructing a world that is consistent with the content of the picture. Transformations are made incrementally,

[2] This defines *Front* as *immediately in front*. It could be defined the other way.

using substitution and deletion relations. The steps are given in (32).[3]

(32) a. Delete everything outside the picture at dref 1, retaining the dref marker.
b. Optionally reverse the picture at dref 1.
c. If the picture is left-oriented, non-deterministically insert any element of *Sit>Sit* to the left of the dref marker 1, and non-deterministically insert any element of *Sit* to the right of the picture marked by 1.
d. If the picture is right-oriented, non-deterministically insert any element of *Sit<Sit* to the right of the picture marked by 1, and non-deterministically insert any element of *Sit* to the left of the dref marker 1.
e. Substitute "y" for "`w`", and "r" for "`b`", in the picture following 1, while deleting the brackets for that picture.
f. Delete the dref marker 1.

Each transformation is definable as a regular relation, except for the reversal. These relations are composed to define the accessibility relation $P_{\triangle}$. Steps c-d preserve oriention, with the center in the output pointing towards the former location of the front of the picture. In order to include reversal, the length of pictures that are reversed has to be bounded.[4] $P_{\triangle}$ is the composition of the six component relations. The centered semantic value of the picture at dref 1 in a world is obtained as an image, see (33).

(33) The centered semantic value of the picture at discourse referent 1 in indexed world s is $[s \circ P_{\triangle}]^{CO}$.

While this construction does not explicitly employ projection, it is natural as a version of projective semantics. Since the worlds are one-dimensional, the projection procedure should look at a zero-dimensional object, that is a point. This indicates that a "projection line" corresponds to a distance in front of the viewpoint, determining a point in the world, the properties of which are checked to determine the picture.[5]

A propositional acquaintance relation P is obtained by composing $P_{\triangle}$ with a relation that deletes the center, see (34), and a propositional semantic value is obtained as the image of P, see (35).

(34) $P \stackrel{\text{def}}{=} P_{\triangle} \circ \mathit{ForgetC}$

(35) The propositional semantic value of the picture at discourse referent 1 in indexed world s is $[s \circ P]^{CO}$.

We return now to Ross's argument. The picture at dref 1 in world p_1 given in (36a) is a picture

[3]See the definition of Pt (corresponding to $P_{\triangle}$) in the supplement.

[4]Possibly this can be finessed, depending on how accessibility is to be used. The finite state calculus includes reversal as an operation, but reversal of a string is not a regular relation. In the code, only pictures of length three or less are reversed.

[5]We have not looked into developing this in the finite state construction.

of two rubies in front of an opal. The corresponding centered semantic value is (36b). This is a certain countably infinite set, which includes the worlds listed in (36c).

(36) a. World p_1
`_ r _ y 1 [b b w) _ r _ y _ r _ y _`
b. $[p_1 \circ P_\triangle]^{CO}$
c.
```
_ (b w w w w] > r _ r _ y _ [b b) _
> r _ r _ y _ y _ [b b) _ [w b b) _
> (b w w w b b] _ r _ r _ y _ r _
_ (b b b b w] _ y _ r _ r _ [w b) <
_ (w b] _ (w b] _ [b b) _ y _ r _ r <
_ y _ r _ r < r _ (w b] _ (w w w] _
```

The picture at dref 1 in the world p_2 given in (37a) is a picture of an opal in front of two rubies. The corresponding centered semantic value is (37b). This is a certain countably infinite set, which includes the worlds listed in (37c).

(37) a. World p_2
`_ r _ y 1 [w b b) _ r _ y _ r _ y _`
b. $[p_2 \circ P_\triangle]^{CO}$
c.
```
_ y > (w b] _ y _ r _ r _ (b w b] _
_ (w b b] > y _ r _ r _ y _ (w w] _
> (w b w b w w] _ r _ y _ r _ r _
_ r _ r _ y _ [w b) _ [b b) < (w b] _
_ r _ r _ y < [w w w) _ [w w b b) _
_ r _ r _ r _ y _ [b w) < [b b w) _
```

The centered semantic values are different if and only if at least one of the set differences $[p_1 \circ P_\triangle]^{CO} - [p_2 \circ P_\triangle]^{CO}$ and $[p_2 \circ P_\triangle]^{CO} - [p_1 \circ P_\triangle]^{CO}$ is non-empty. Given the explicit computable semantics, this can be checked computationally. Both differences are non-empty. (38) lists some worlds in the content of the first picture but not the second. (39) lists some worlds in the content of the second picture but not the first. So the centered contents of the two pictures are different, as required in the argument.

(38)
```
_ y _ r _ r _ r _ [w w w w) < r _
> (w w] _ r _ r _ y _ [w b) _ [w w) _
_ r _ y _ [b w b w) > r _ r _ y _
_ y _ r _ r _ [b b w) _ [b w) _ r <
_ [w w) _ (w w] _ y _ r _ r < [b b) _
_ y _ r _ r _ [w w b b b b b b) <
```

(39)

```
_ r _ r _ y _ [b b b) _ (b b b b] <
_ [b b b) _ r _ r _ r _ y _ [w w) <
_ (b b b w] _ r _ r _ y _ (b w b] <
_ r > y _ r _ r _ (b b w] _ (b w] _
_ r _ r _ y _ [b b b b) _ (w b b] <
> y _ y _ r _ r _ (w b] _ [w w b) _
```

By the way, while the centered contents are different, they are not disjoint. (40) lists some worlds in the semantic conjunction of the two pictures.

(40)

```
_ [w w b) > y _ r _ r _ y _ (b w] _
> y _ y _ r _ r _ y _ (w w w w] _
_ [b b) _ y _ r _ r _ y _ (w b b] <
_ [w b w w b w) > y _ r _ r _ y _
_ [w b b) _ y _ r _ r _ y _ (w b] <
_ (b w b w] > r _ r _ y _ r _ r _
```

Second we evaluate the set differences $[p_1 \circ P]^{CO} - [p_2 \circ P]^{CO}$ and $[p_2 \circ P]^{CO} - [p_1 \circ P]^{CO}$ computationally. Both terms evaluate to the empty set, indicating that the propositional contents are the same. This is what failed in Section 4.

Finally, consider truth value of sentence (41a) in world p_1. We assume the pronoun picks up the center, marked by the dref marker 1. We are using a dynamic semantics, with the prejacent sentence formalized as in (41b). To check the truth value, we check whether any centered worlds are lost in moving from the centered proposition denoted by the picture to that proposition updated with the prejacent sentence. The original proposition is (42a), and the updated proposition is (42b). When these are compared computationally using a set difference, we find that no worlds are lost in the update; this indicates that (41a) is true in p_1. When the same experiment is done with p_2, worlds are lost, and sentence (41a) is false in p_2.

(41) a. In it, there is a ruby in front of an opal.
b. There is a ruby in front of an opal.
c. $[\mathit{NewC} \circ \mathit{New} \circ \mathit{Opal} \circ \mathit{New} \circ \mathit{Ruby} \circ \mathit{Front} \circ \mathit{Forget}]^{CO}$

(42) a. $[p_1 \circ P_{\triangle}]^{CO}$
b. $[[p_1 \circ P_{\triangle}]^{CO} \circ \mathit{NewC} \circ \mathit{New} \circ \mathit{Opal} \circ \mathit{New} \circ \mathit{Ruby} \circ \mathit{Front} \circ \mathit{Forget}]^{CO}$

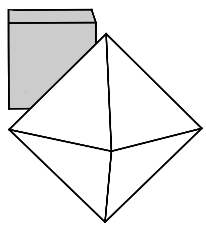
This reconstructs Ross's argument using the lineland modal space. Sentence (41a) is true with reference to the picture in world p_1, and false with reference to the picture in p_2, using a centered semantics computed as in (42). Using a non-centered semantics for the pictures can not give this result, because the propositional semantic values of the pictures are the same.

All of this shows that whether Ross's argument works or not depends on specifics of the modal space and of the projection procedure.

6. Back to 3D

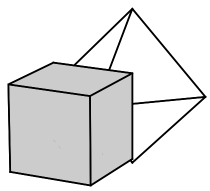
There are ways of using pictures to give information about three-dimensional worlds where the pictures are not intended to carry as much information as they do in the projective semantics from Section 2. Keisha owns some regular polytopes. She told us about them by uttering the sentence (43a), and then showing the picture (43b). She intended to give information about the shape, color and relative orientation of her polytopes, but did not intend to imply anything about what else is or is not in the vicinity. She did not, for instance, intend to rule out the possibility that her polytopes are surrounded in every direction by spheres.

(43) a. I own two regular polytopes. This is how they are oriented.

b. 

c. In the picture, my favorite polytope is in front.

This example can be turned into a Ross example by continuing as in (43c). The sequence carries the information that the favorite polytope is an octahedron. But if the picture is switched to (44), we get the information that the favorite polytope is a cube. (43b) and (44) are projected from the same world, but different viewpoints.

(44) 

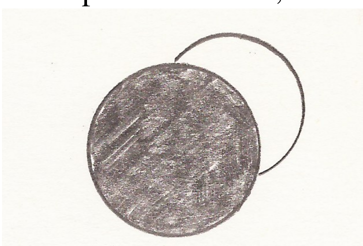
The above suggests fixing the argument by incorporating information about what objects are depicted. Keisha's intention comes down to not intending to depict anything other than her two polytopes. We suggest locating this in the marking rule. Often, pictures are used assuming marking rules with non-geometric side conditions. We can draw a map-like picture projected from above the northern reaches of our university campus, marking red for projection lines that intersect buildings belonging to the university, and gray for other buildings. Here red indicates ownership by the university, not anything geometric. In the same way, the marking rule can stipulate that only objects belonging to Keisha have an influence on marking.

Suppose the first sentence in (43) sets up a discourse referent X_2 for two polytopes owned by Keisha. Before the picture is processed semantically, a marking rule is accommodated that marks only elements of X_2, see (45). If there are spheres surrounding the polytopes, these do not affect the picture. So with this marking rule, the picture does not carry the extra information that came up in Section 3.

(45) Mark a point in the picture plane in black if the directed line from the viewpoint through that point intersects the edge of an element of X_2 before it intersects any other element of X_2, otherwise in gray if the directed line from the viewpoint through that point intersects a black element of X_2 before it intersects any other element of X_2, otherwise in white.

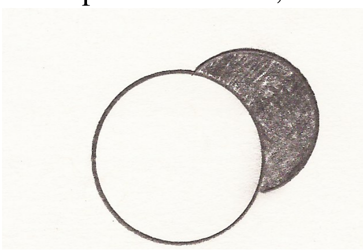
With a context and marking rule like this, we are back in business with a 3D version of Ross's argument. The pictures give the same information about the relative orientation of the objects belonging to Keisha, and the context and marking rule ensure that no other objects are depicted. The argument stated for spheres is given in (46)–(47). The centered contents are different, and the last sentence in (46) can come out true, while the last sentence in (47) comes out false. But the propositional contents are the same, indicating that no semantic rule that uses a propositional content for pictures can work.

(46) Keisha owns two spheres$_3$. This is how they are oriented.
(Accommodate a marking rule that ignores objects that are not elements of X_2, marks black spheres in black, and white spheres in white.)



In the picture, there is a black sphere in front of a white sphere. Both belong to Keisha.

(47) Keisha owns two spheres$_3$. This is how they are oriented.
(Accommodate a marking rule that ignores objects that are not elements of X_2, marks black spheres in black, and white spheres in white.)



In the picture, there is a black sphere in front of a white sphere. Both belong to Keisha.

This reconstructs Ross's argument from the spheres for centered pictorial contents. The moves we made with the marking rule are motivated, because speakers do use pictures intending that only certain specified or contextually determined objects are depicted. The contrast between this version of the argument and the original one indicates that whether or not different centered contents can collapse into a single propositional content depends on details of the marking rule.

References

Abusch, D. (2012). Applying discourse semantics and pragmatics to co-reference in picture sequences. In *Proceedings of Sinn und Bedeutung 17*. Semantics Archive.

Abusch, D. (2014). Temporal succession and aspectual type in visual narrative. In L. Crnič and U. Sauerland (Eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*. Cambridge, Mass.: MITWPL.

Abusch, D. (2016). Possible worlds semantics for pictures. Prepublication version of handbook article.

Beesley, K. R. and L. Karttunen (2003). *Finite State Morphology*. CSLI, Stanford.

Blumson, B. (2010). Pictures, perspective and possibility. *Philosophical Studies 149*(2), 135–151.

Greenberg, G. (2011). *The Semiotic Spectrum*. Ph. D. thesis, Rutgers.

Greenberg, G. (2013). Pictorial semantics. *Philosophical Review 122*.

Hulden, M. (2009). Foma: a finite-state compiler and library. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics: Demonstrations Session*, pp. 29–32. Association for Computational Linguistics.

Kelleher, J. and J. van Genabith (2006). A computational model of the referential semantics of projective prepositions. In *Syntax and Semantics of Prepositions*, pp. 211–228. Springer.

Kemmerer, D. (2006). The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia 44*(9), 1607–1621.

Lewis, D. (1979). Attitudes de dicto and de se. *Philosophical Review 88*(4), 513–543.

Rooth, M. and D. Abusch (2017). Replication data for: Picture descriptions and centered content. Harvard Dataverse Network.

Ross, J. (1997). *The Semantics of Media*. Number 00 in Studies in Linguistics and Philosophy. Dordrecht: Kluwer.

Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Springer.