

Intensional and Temporal Pictorial Conflation

Intentional constructions in comics and film are ones where comic panels or shots in a film are parsed as conveying what a character sees or imagines, either from an agent-centered or neutral perspective. Temporal conflation in pictures (equivalently, continuous narrative) is the phenomenon of information from different time points being combined in a single picture. This presentation argues that some intensional examples should be analyzed as intensionally conflated, where information from an attitudinal world and from the base world are combined in a single picture. (1a) is an example of continuous narrative, where a bear is shown five times. It is understood as describing a bear walking through a ravine up to a tree, where it is shot by a hunter and tumbles down. In (1b) a boy hallucinates a fairy inside a jar, with his mother standing behind him (Maier and Bimpikou 2019). (1c) is a frame from the series Poldark where the protagonist imagines confessing to his wife about an interaction with a former girlfriend. (1d) is a frame from the series Umbrella Academy, showing a character in the background who sees ghosts (the character in the foreground).

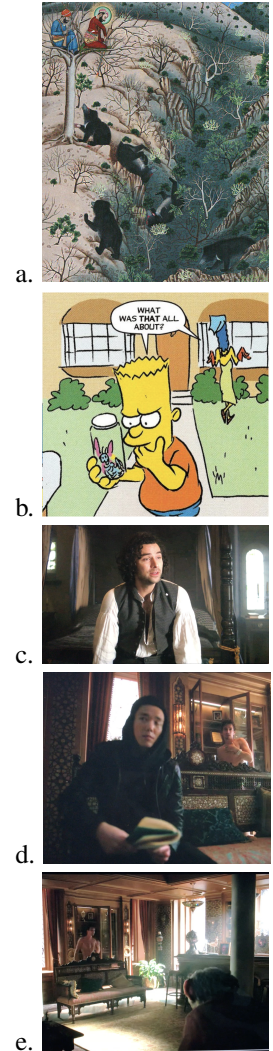
A de se semantics for perception, epistemic attitudes, and other attitudes is assumed throughout. Following Abusch and Rooth (2017) and Maier and Bimpikou (2019), intensionally construed pictures are syntactically embedded at the level which is semantically interpreted, and discourse referents can be introduced internal to the complement. A format for a basic analysis is $A_i(pd)$, where A is the attitude predicate, p is the picture, i is a discourse referent for the external agent (the subject of the attitude), and d introduces a discourse referent internal to the complement, which by convention is the de se counterpart of the external agent. The complement is interpreted as a relation between worlds, times, agents, a viewpoint, and optionally additional individuals. In free perception examples, the viewpoint coincides with the geometric viewpoint of the agent's counterpart, but in (1b-e) it does not, because the agent is depicted. A basic de se semantics requires that the relation modeling the external agent's centered attitudinal state entails the relation contributed by the complement (with some treatment of the viewpoint specified.) This is what is proposed for intensional free perception in Abusch and Rooth (2017) and for neutral-perspective examples such as (1b-e) in Maier and Bimpikou (2019).

We argue the truth conditions delivered by this basic analysis are often too strong in intensionally construed pictures with a neutral perspective. Consider the example with the jar (1b). Suppose that in the prior pictorial narrative, the mother is shown crossing her legs while the boy faces the other way, looking at the jar. Then the boy does not see the mother crossing her legs, and the boy in the base world of a described situation does not have the information that the mother has crossed legs. Therefore the epistemic state of the boy in the base world of the described situation for (1b) does not entail the content of the picture as a whole. Similarly for the other examples. A loose idea for a corrective is that a boy looking at a jar only gets information about what is in his field of view, and so the semantics of (1b) should not require that the boy in the base world of the described situation have the information that his mother is in a particular posture. The following is a starting point at weakening the information extracted from the complement.

- (2) Set of centered worlds determined by a picture p plus de-se discourse referent d .

The set of all $\langle w'', t'', x'' \rangle$ such that for $\dots \langle w', t', x', v' \rangle$ in the semantic value of pd , $\langle w'', t'' \rangle$ looks the same from the visual-geometric viewpoint of x'' in w'' at t'' as $\langle w', t' \rangle$ looks from the visual-geometric viewpoint of x' in w' at t' . (Fill in a suitable quantifier for \dots).

In general, visual-epistemic modality is a circumstantial modality that characterizes an equivalence relation on worlds centered on agents with a visual system, where $\langle w, t, x \rangle \sim \langle w', t', x' \rangle$ iff $\langle w, t \rangle$ looks the same from the visual-geometric viewpoint of x in w at t as $\langle w', t' \rangle$ looks from the visual-geometric



(1)

viewpoint of x' in w' at t' . Equivalently, a visual-epistemic state is a total answer to the question “what do things look like” as modeled in centered-world partition semantics for questions. Roughly, we would like to say that the information contributed by pd is that the agent in the base world is in the same cell of the partition as the agent depicted in the picture. This idea runs afoul of the fact that for two tuples $\langle w'_1, t'_1, x'_1, v'_1 \rangle$ and $\langle w'_2, t'_2, x'_2, v'_2 \rangle$ in the semantic value of pd , $\langle w'_1, t'_1, x'_1 \rangle$ and $\langle w'_2, t'_2, x'_2 \rangle$ need not be visual-epistemic alternatives. Suppose that in $\langle w'_1, t'_1 \rangle$ there is a smudge on the jar, on the right of the jar from x'_1 's perspective. The smudge is absent in $\langle w'_2, t'_2 \rangle$. The difference is visible from the visual perspective of the agents, though it is out of view at v'_1 and v'_2 . In this scenario, $\langle w'_1, t'_1, x'_1 \rangle$ and $\langle w'_2, t'_2, x'_2 \rangle$ are not visual-epistemic alternatives, though $\langle w'_1, t'_1, x'_1, v'_1 \rangle$ and $\langle w'_2, t'_2, x'_2, v'_2 \rangle$ are both in the semantic value of pd . This reasoning indicates that the right choice for \dots in (2) is the quantifier *some*. Equivalently, the set of agent-relevant centered worlds contributed by pd is the union of the sets of visual-epistemic alternatives to $\langle w', t', x' \rangle$ such that for some v' , $\langle w', t', x', v' \rangle$ is in the semantic value of pd .

If we merely require that the agent-centered epistemic state of the agent for the attitude entails the relation described in (2), then the mother in the background of (1b) makes no contribution at all to semantic interpretation, because she is always out of view of counterparts of the boy. The solution to this is that the depiction of the mother constrains the *base* world, rather than the worlds that are epistemic alternatives for the agent in the base world. We claim that three-dimensional worlds that are consistent with picture (1b) are split into a part that constrains epistemic alternatives (this is the part in the cone of view of the agent) and a part that constrains the base world (the rest).

This brings us to temporally conflated narratives, such as the narrative (1a) with the bear. It is uncontroversial that such pictures provide information about different world-time slices $\langle w, t \rangle$ and $\langle w, t' \rangle$. We analyze temporal conflation in the geometric framework of possible worlds semantics for pictures, and use this as a model for modal conflation. Abusch (2012) suggested syntactically splitting continuous and conflated readings of pictures and sculptures into parts at the level of the artifact (two-dimensional parts in the case of pictures, and three-dimensional parts in the case of sculpture). We instead propose an analysis where parts of worlds are sutured together as three-dimensional spaces, followed by geometric projection in the normal sense. In (1a), different spatial parts of one world history are sutured together in a way that keeps the relative spatial positions of parts of the landscape constant across parts from different times. This creates a composite 3D space, where there are bears in different locations, drawn from spatial parts (they could be vertical prisms) from different times. This composite space is projected in the standard way. Technically, the locations of boundaries between spatio-temporal parts is specified in the viewpoint parameter.

This idea is applied to modally conflated pictures in (3), which geometrically splits a described world for pd into a part of the base world and a part of a visual-epistemic alternative.

- (3) $\langle w, t, v, \mathcal{O} \rangle$ satisfies $Ai(pd)$ iff there are tuples $\langle w', t', v', x' \rangle$ and $\langle w'', t'', x'' \rangle$ s.t. $\langle w', t', v', x' \rangle$ satisfies pd , $\langle w, t, \mathcal{O}[i] \rangle \sim \langle w'', t'', x'' \rangle$, and $\langle w', t' \rangle$ is formed geometrically from $\langle w'', t'' \rangle$ and $\langle w, t \rangle$ by splicing together the oriented cone in $\langle w'', t'' \rangle$ corresponding to x'' 's field of view with the oriented spatial area in $\langle w, t \rangle$ that is the spatial complement of the cone that is x 's field of view in w at t .

This is analogous to *de re* interpretation in linguistic semantics, in that part of the complement constrains visual epistemic alternatives, and part constrains the base world. The full presentation considers other attitudes (1c,d). Any additional discourse referents are interpreted *de re* when they are indexed across the complement.

The last part motivates and formulates a constraint on intensional constructions we call the V-constraint. An intensional passage in pictorial narrative is frequently terminated by cutting to a shot that includes a principal character who is not the agent of the intensional passage. This is operative in the Poldark example (1c) where the narrative cuts to a shot including the wife, terminating the embedded passage. It is also operative in the Umbrella Academy example (1d), where cutting to a shot (1e) that includes another principal character (the chimp in the right foreground) is accompanied by the ghost becoming invisible. The V-constraint is formulated as a defeasible continuity constraint, analogous to the X-constraint in Cumming, Greenberg and Kelly (2016). The constraint can be seen as motivated by avoiding ambiguity in the indexing of the agent. The analysis of (1d,e) is complicated by the dimension of realism vs. intensionalism about perception of ghosts. We argue that the V-constraint is also operative under assumption of realism about ghosts. **References.** Abusch (2012) *Sinn und Bedeutung*. Abusch and Rooth (2017) Amsterdam Colloquium. Cumming, Greenberg, and Kelly (2016) *Philosophers' Imprint*. Maier and Bimpikou (2019) *Sinn und Bedeutung*.