Latent Event Structure

- Atomic Object Structure:
  Formal Quale (objects expressed as basic nominal types)
- Subatomic Object Structure:
  Constitutive Quale (mereotopological structure of objects)
- Object Event Structure:
  Telic and Agentive Qualia structure (origin and functions associated with an object)
- Macro Object Structure:
  habitats, object frames, embedding object structures

The Implicit Event Structure of Things

Motivation for Qualia relations comes from the idea that there is a *hidden event* in the lexical representation associated with nouns denoting objects made for a particular purpose:

(8) a. a door is for walking through
b. a window is for seeing through
c. a book is for reading
d. a beer is for drinking
e. a cake is for eating
f. a car is for driving
g. a table is for putting things on
h. a desk is for working on
i. a pen is for writing with

Nouns encode events relating to use or function

(9) a. This pen does not work well. (does not write)
b. Can I use your pen? (for writing)
c. Have you got a red pen? (ambiguous, which writes in red)

(10) a. Any chocolate? Not after that cake! (after eating)
b. I prefer cake to biscuits. (prefer eating)
c. We skipped the cake and settled for another coffee. (skipped eating)

(11) a. There's no train till 7:00 pm. (there is no departing)
b. The train was delayed for an hour. (the departure)
c. I left in time to catch the early train. (departing early)

Adjective-Noun Telic Interpretations

(12) a. the next customer (to be taken care of)
c. the next slide (to be projected)

(13) a. This is a difficult problem (to solve).
b. This is a difficult question (to answer).

(14) Telic selectors:
fast food (to eat), a slow oven (to cook), a short novel (to read), a complex question (to answer), an easy place (to get to), useful, an effective antibiotic (to cure), agreeable, avoidable costs (to pay), enjoyable, a good doctor (to heal), a bad singer (to listen to), an interesting book (to read), ready meals (to eat).

Semantically Transparent Nominals

(15) a. functional locations: *library*, *gym*, *church*, *school*;
b. professions: *doctor*, *teacher*, *lawyer*;
c. agentive nominals (individuals engaged in an activity, either habitually or occasionally): *runner*, *passenger*, *movie goer*.

Encoding Events in Qualia Structure

$$(16)\ \begin{bmatrix} \textit{cake} \\ \text{QUALIA} = \begin{bmatrix} \text{F} = \textbf{food} \\ \text{T} = \textbf{eat(human,food)} \end{bmatrix} \end{bmatrix}$$

$$(17)\ \begin{bmatrix} \textit{pen} \\ \text{QUALIA} = \begin{bmatrix} \text{F} = \textbf{tool} \\ \text{T} = \textbf{write\_with} \end{bmatrix} \end{bmatrix}$$

$$(18)\ \begin{bmatrix} \textit{singer} \\ \text{QUALIA} = \begin{bmatrix} \text{F} = \textbf{human} \\ \text{T} = \textbf{sing(human, song)} \end{bmatrix} \end{bmatrix}$$

Teleotopology

- The function of space: the actions associated with a region or an object (inherently or opportunistically), i.e., Telic role values.
- The space of function: the regions defined by the Telic actions performed by an agent, or supervenient on the Telic state of an artifact, teleotopology.

Extending Qualia to Modeling Affordances

The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. It implies the complementarity of the animal and the environment. An affordance is neither an objective property nor a subjective property; or both if you like. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. [It] points both ways, to the environment and to the observer. (J. J. Gibson, 1979/1986)

- Gibson (1979), Turvey (1992), Steedman (2002), Sahin et al (2007), Krippendorff (2010);
- Affordance: a correlation between an agent who acts on an object with a systematic or prototypical effect.

Semantics of Function and Purpose

There are two levels of accessibility that can be identified in a Telic role value, as illustrated below.

(19) a. local modality (habitat): the conditions under which the activity can be performed on the object;
b. global modality: what is done with the object, and the resulting state.

Telic Values and Affordances

(20) $\mathcal{C} \rightarrow [\pi]\mathcal{R}$

π	π^+	$\mathcal{R}?$
$\mathcal{C}?$		$\neg\mathcal{C}?$

$\langle i,j \rangle$

Pustejovsky (2012) "The Semantics of Functional Spaces"

Telic Values and Affordances

The TELIC of *sandwich*:

$$(21)\quad \lambda x \begin{bmatrix} \textbf{sandwich} \\ \text{AS} = \begin{bmatrix} \text{ARG1} = x : e \end{bmatrix} \\ \text{QS} = \begin{bmatrix} \text{F} = phys(x) \\ \text{T} = \lambda y \lambda e[\mathcal{C} \rightarrow [eat(e,y,x)]\mathcal{R}_{eat}(x)]] \\ \text{A} = \exists z[make(z,x)] \end{bmatrix} \end{bmatrix}$$

Teleotopology

- A region created by the action(s) associated with a purposeful action by an agent;
- A region required for the performance or satisfaction of an artifact.

Action Predicates

(22) *cut*-verbs: *saw*, *ax*, *slice*

(23) $move_{dir+tr}(x) =_{df} \; loc(x) := y, b := y, p := (b) \; ; (y := z, y \neq z, p := (p, z), d(b, y) < d(b, z))^{+}$

(24)

$$\left[\!\left[\; move_{dir+tr}(x), p := (b) \;\right]\!\left[\; move_{dir+tr}(x), p := (p, z) \;\right]\!\right]^{+}_{\langle i,j \rangle}$$

Compositional Constraints in Actions 1/2

- "do π while $\neg\phi$ is true, and stop doing π when ϕ becomes true", over the interval $\langle i, j \rangle$.

π	π^+	π
$\neg\phi?$		$\phi?$

$\langle i,j \rangle$

- "do π and α while $\neg\phi$ is true and ψ is true, and stop doing π and α when ϕ and $\neg\psi$ become true", over the interval $\langle i, j \rangle$.

π	π^+	π	
α	α^+	α	
$\neg\phi?$		$\phi?$	$\langle i,j \rangle$
$\psi?$		$\neg\psi?$	

Affordance Spaces

(25) a. l_o, l_p, l_a: The location (spatial extent) defined by an object, x, its action, p, and the agent, a, respectively.

b. R_e : An embedding space, for the object-action-agent location, the convex hull of the agent using the object through time, $Conv(l_o \otimes l_p \otimes l_a)$.

c. μ: The *affordance space* is the minimal embedding space for the object:
$\forall l_o \otimes l_p \otimes l_a \exists \mu [l_o \otimes l_p \otimes l_a \subseteq \mu \rightarrow \forall R_e [l_o \otimes l_p \otimes l_a \subseteq R_e \rightarrow [R_e = \mu \vee \mu \subseteq R_e]]]$

Telic Values and Affordances

Representing the action predicate saw:

(26) a. Given an instrument of appropriate constraints, x (e.g., a saw) and an arm, y:

b. While grasping x with $hand(y)$:

c. Push x away (out) with downward pressure on object z, until extension of y is reached;

d. Pull x toward (in) with downward pressure on object z, until flexion of y is reached;

e. Repeat (c) and (d) until Goal, G is satisfied (e.g., separation of z).

Compositional Constraints for the Action of saw

$push, \mu; pull, \mu'$	$(push, \mu; pull, \mu')^+$	$push, \mu; pull, \mu'$
$grasp$		
$\neg G?$		$G?$

$\langle i,j \rangle$

Habitats for Artifacts

- table: $\mathcal{C}$ = "top oriented up", "surface is accessible", etc.
- chair: $\mathcal{C}$ = "oriented up", "seat is accessible", etc.
- table and chair: $\mathcal{C}$ = "spatially consistent", etc.
- Telic(table and chair): $\mathcal{C}$ = agent must be able to function at table from position in the chair, etc.

Habitats and Simulations

- Habitat: a representation of an object situated within a partial minimal model; Enhancements of the qualia structure.
- With multi-dimensional affordances that determine how habitats are deployed and how they modify or augment the context.
- Compositional combinations of procedural (simulation) and operational (selection, specification, refinement) knowledge.

Qualia Structure

$$
(27)\quad \lambda x \begin{bmatrix} \textbf{chair} \\ \text{AS} = \begin{bmatrix} \text{ARG1} = x : e \end{bmatrix} \\ \text{QS} = \begin{bmatrix} \text{F} = phys(x) \\ \text{T} = \lambda z \lambda e[\mathcal{C} \rightarrow [sit(e, z, x)]\mathcal{R}_{sit}(x)] \\ \text{A} = \exists w \exists e'[make(e', w, x)] \end{bmatrix} \end{bmatrix}
$$

Habitats for Artifacts

- table: $\mathcal{C}$ = "top oriented up", "surface is accessible", etc.
- chair: $\mathcal{C}$ = "oriented up", "seat is accessible", etc.
- table and chair: $\mathcal{C}$ = "spatially consistent", etc.
- Telic(table and chair): $\mathcal{C}$ = agent must be able to function at table from position in the chair, etc.

Expand the Context variable $\mathcal{C}$ to build a partial model, $\mathcal{M}$.

$$\lambda x \begin{bmatrix} \textbf{chair}_{hab} \\ \text{QS} = \begin{bmatrix} \text{F} = [phys(x), on(x, y_1), in(x, y_2), orient(x, up)] \\ \text{C} = [seat(x_1), back(x_2), legs(x_3), clear(x_1)] \\ \text{T} = \lambda z \lambda e[\mathcal{C} \rightarrow [sit(e, z, x)]\mathcal{R}_{sit}(x)] \\ \text{A} = [made(e', w, x)] \end{bmatrix} \end{bmatrix}$$

Visual Object Concept Modeling Language (VoxML)

Pustejovsky and Krishnaswamy (2014, 2016)

- Modeling language for constructing 3D visualizations of concepts denoted by natural language expressions
- Used as the platform for creating *multimodal semantic simulations*
- Encodes dynamic semantics of events and objects and object properties
- Platform independent framework for encoding and visualizing linguistic knowledge.

Visual Object Concept (Voxeme)

- Object Geometry Structure:
  Formal object characteristics in R3 space
- Habitat: Embodied and embedded object:
  Orientation
  Situated context
  Scaling
- Affordance Structure:
  What can one do to it
  What can one do with it
  What does it enable
- Voxicon: library of voxemes

VoxML Elements

Entities modeled in VoxML can be:

- Objects: Physical objects (Nouns)
- Programs: Events (Verbs)
- Attributes: Properties (Adjectives)
- Functions: Quantifiers, connectives

These entities can then compose into visualizations of natural language concepts and expressions.

VoxML Concepts

- $\mathcal{E}$ — the minimal embedding space (MES)
- $\mathcal{E}_A$ — the axis A of the MES
- $loc(x)$ — location of object x
- $orient(x)$ — orientation of object x
- $vec(A)$ — vector denoted by axis A (+ by default)
- $opp(v)$ — opposite vector of v
- $reify(x, s)$ — relabel object x (a collection ($c_1,\ldots,c_n$)) as s
- $interior(x)$ — the interior surface (and volumetric enclosed space) of object x
- $exterior(x)$ — the exterior surface of object x
- $dimension(x)$ — the number of dimensions defining entity x
- $while(\phi, e)$ — operation e is executed as long as ϕ is true
- $for(x \in y)$ — following operation is executed for each x in y
- $align(A, B)$ – for vectors A, B, defines A as parallel with B

VoxML Template: Object

(28)

$$\left[\begin{array}{l} \textbf{OBJECT} \\ \text{LEX} = \left[\begin{array}{l}\text{PRED} = \textbf{...}\\ \text{TYPE} = \textbf{...}\end{array}\right] \\ \text{TYPE} = \left[\begin{array}{l}\text{HEAD} = \textbf{...}\\ \text{COMPONENTS} = \textbf{...}\\ \text{CONCAVITY} = \textbf{...}\\ \text{ROTATSYM} = \{\ldots\}\\ \text{REFLECTSYM} = \{\ldots\}\\ \text{CONSTR} = \{\ldots\}\end{array}\right] \\ \text{HABITAT} = \left[\begin{array}{l}\text{INTR} = \ldots\\ \text{EXTR} = \ldots\end{array}\right] \\ \text{AFFORD\_STR} = \left[\,\text{A}_n = H_{[\#]} \rightarrow [E(a_{1..n})]R(a_{1..n})\,\right] \\ \text{EMBODIMENT} = \left[\begin{array}{l}\text{SCALE} = \textbf{...}\\ \text{MOVABLE} = \textbf{...}\end{array}\right] \end{array}\right]$$

VoxML Object is used for modeling nouns: 1/5

Lex	Object's lexical information
Type	Object's geometrical typing
Habitat	Object's habitat for actions
Afford_Str	Object's affordance structure
Embodiment	Object's agent-relative embodiment

Objects 2/5

- The TYPE attribute contains information to define the object geometry in terms of primitives. HEAD is a primitive 3D shape that roughly describes the object's form or the form of the object's most semantically salient subpart.

HEAD	prismatoid, pyramid, wedge, parallelepiped, cupola, frustum, cylindroid, ellipsoid, hemiellipsoid, bipyramid, rectangular_prism, toroid, sheet

Objects 3/5

- COMPONENTS: subparts of the object
- CONCAVITY: `concave`, `flat`, or `convex`; refers to any concavity that deforms the HEAD shape.
- ROTATSYM (rotational symmetry) defines any of the three orthogonal axes around which the object's geometry may be rotated for an interval of less than 360 degrees and retain identical form as the unrotated geometry.
- REFLECTSYM (Reflectional symmetry): If an object may be bisected by a plane defined by two of the three orthogonal axes and then reflected across that plane to obtain the same geometric form as the original object, it is considered to have reflectional symmetry across that plane.

HABITAT defines habitats INTRINSIC to the object, regardless of what action it participates in, such as intrinsic orientations or surfaces, as well as EXTRINSIC habitats which must be satisfied for particular actions to take place.

Objects 5/5

AFFORD_STR describes the set of specific actions, along with the requisite conditions, that the object may take part in. There are low-level affordances, called GIBSONIAN, which involve manipulation or maneuver-based actions (grasping, holding, lifting, touching); there are also TELIC affordances, which link directly to what goal-directed activity can be accomplished, by means of the GIBSONIAN affordances.

EMBODIMENT qualitatively describes the SCALE of the object compared to an in-world agent (typically assumed to be a human) as well as whether the object is typically MOVABLE by that agent.

Plate

$$(29)\quad \left[\begin{array}{l}
\textbf{plate} \\
\text{LEX} = \left[\begin{array}{l}\text{PRED} = \textbf{plate} \\ \text{TYPE} = \textbf{physobj}\end{array}\right] \\
\text{TYPE} = \left[\begin{array}{l}\text{HEAD} = \textbf{sheet} \\ \text{COMPONENTS} = \textbf{surface, base} \\ \text{CONCAVITY} = \textbf{concave} \\ \text{ROTATSYM} = \{Y\} \\ \text{REFLECTSYM} = \{XY, YZ\}\end{array}\right] \\
\text{HABITAT} = \left[\begin{array}{l}\text{INTR} = [1]\left[\begin{array}{l}\text{UP} = align(Y, \mathcal{E}_Y) \\ \text{TOP} = top(+Y)\end{array}\right] \\ \text{EXTR} = \ldots\end{array}\right] \\
\text{AFFORD_STR} = \left[\begin{array}{l}\text{A}_1 = H[1] \to [put(x,y)]hold(y,x) \\ \text{A}_2 = \ldots \\ \text{A}_3 = \ldots\end{array}\right] \\
\text{EMBODIMENT} = \left[\begin{array}{l}\text{SCALE} = < \textbf{agent} \\ \text{MOVABLE} = \textbf{true}\end{array}\right]
\end{array}\right]$$

Plate

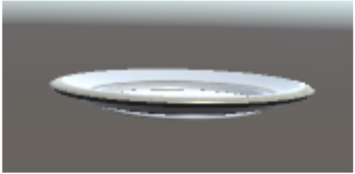

Figure: Plate voxeme instance

VoxML for cup

cup

LEX = [PRED = **cup**, TYPE = **physobj**]

TYPE = [HEAD = **cylindroid[1]**, COMPONENTS = **surface,interior**, CONCAVITY = **concave**, ROTATSYM = $\{Y\}$, REFLECTSYM = $\{XY, YZ\}$]

HABITAT = [INTR = [2] [UP = $align(Y, \mathcal{E}_Y)$, TOP = $top(+Y)$], EXTR = ...]

AFFORD_STR = [

$A_1 = H_{[2]} \rightarrow [put(x, on([1]))]support([1], x)$

$A_2 = H_{[2]} \rightarrow [put(x, in([1]))]contain([1], x)$

$A_3 = H_{[2]} \rightarrow [grasp(x, [1])]$]

EMBODIMENT = [SCALE = <**agent**, MOVABLE = **true**]

VoxML for spoon

$$
(30)\quad
\begin{bmatrix}
\textbf{spoon} \\
\text{LEX} = \begin{bmatrix} \text{PRED} = \textbf{spoon} \\ \text{TYPE} = \textbf{physobj, artifact} \end{bmatrix} \\
\text{TYPE} = \begin{bmatrix} \text{HEAD} = \textbf{sheet[1]} \\ \text{COMPONENTS} = \textbf{handle[2], bowl[3]} \\ \text{CONCAVITY} = \textbf{concave} \\ \text{ROTATSYM} = \mathit{nil} \\ \text{REFLECTSYM} = \{YZ\} \end{bmatrix} \\
\text{HABITAT} = \begin{bmatrix} \text{INTR} = [4]\begin{bmatrix} \text{CONSTR} = \{Z > X, Z \gg Y\} \\ \text{UP} = \mathit{align}(Y, \mathcal{E}_Y) \\ \text{FRONT} = \mathit{top}(+Y) \end{bmatrix} \\ \text{EXTR} = [5]\begin{bmatrix} \text{UP} = \mathit{align}(Y, \mathcal{E}_{\perp Y}) \end{bmatrix} \end{bmatrix} \\
\text{AFFORD\_STR} = \begin{bmatrix} A_1 = H_{[4]} \rightarrow [\mathit{put}(x, \mathit{in}([3]))]\mathit{contain}([3], x) \\ A_2 = H_{[4]} \rightarrow [\mathit{grasp}(x, [2])] \\ A_1 = H_{[4]} \rightarrow [\mathit{put}([1], \mathit{in}(x))]\mathit{contain}(x, [1]) \\ A_1 = H_{[5]}, \mathit{contain}(x, [1]) \rightarrow [\mathit{stir}([1], x)] \end{bmatrix} \\
\text{EMBODIMENT} = \begin{bmatrix} \text{SCALE} = <\textbf{agent} \\ \text{MOVABLE} = \textbf{true} \end{bmatrix}
\end{bmatrix}
$$

VoxML for book

(31)

$$
\left[\begin{array}{l}
\textbf{book} \\
\text{LEX} = \left[\begin{array}{l} \text{PRED} = \textbf{book} \\ \text{TYPE} = \textbf{physobj, artifactj} \end{array}\right] \\
\text{TYPE} = \left[\begin{array}{l} \text{HEAD} = \textbf{rectangular\_prism[1]} \\ \text{COMPONENTS} = \textbf{cover[2]+, page[3]+} \\ \text{CONCAVITY} = \textbf{flat} \\ \text{ROTATSYM} = \textbf{nil} \\ \text{REFLECTSYM} = \{XY\} \\ \end{array}\right] \\
\text{HABITAT} = \left[\begin{array}{l} \text{INTR} = [4]\left[\begin{array}{l} \text{UP} = align(Y, \mathcal{E}_Y) \\ \text{TOP} = front(+Y) \end{array}\right] \\ \text{EXTR} = \ldots \end{array}\right] \\
\text{AFFORD\_STR} = \left[\begin{array}{l} \text{A}_1 = H \rightarrow [grasp(x, [2]), \\ \quad move(x, [2], away(from([3])))]open(x, [1]) \\ \text{A}_2 = H \rightarrow [grasp(x, [2]), \\ \quad move(x, [2], toward([3]))]close(x, [1]) \end{array}\right] \\
\text{EMBODIMENT} = \left[\begin{array}{l} \text{SCALE} = \textbf{<agent} \\ \text{MOVABLE} = \textbf{true} \end{array}\right]
\end{array}\right]
$$

VoxML Template: Program

(32)
$$\begin{bmatrix} \textbf{PROGRAM} \\ \text{LEX} = \begin{bmatrix} \text{PRED} = \textbf{...} \\ \text{TYPE} = \textbf{...} \end{bmatrix} \\ \text{TYPE} = \begin{bmatrix} \text{HEAD} = \textbf{...} \\ \text{ARGS} = \begin{bmatrix} \text{A}_1 = \textbf{x:a} \end{bmatrix} \\ \text{BODY} = \begin{bmatrix} \text{E}_n = E(a_{1..n}) \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Programs are used for modeling verbs

LEX	PROGRAM's lexical information
TYPE	PROGRAM's event typing
EMBEDDING_SPACE	PROGRAM's embodiment as a function of the participants and their changes over time

A PROGRAM's LEX attribute contains the subcomponents PRED, the lexeme predicate denoting the program, and TYPE, the program's type as given in a lexical semantic resource, e.g., its GL type.

VoxML for put

$$
(33)\quad \begin{bmatrix} \textbf{put} \\ \text{LEX} = \begin{bmatrix} \text{PRED} = \textbf{put} \\ \text{TYPE} = \textbf{transition\_event} \end{bmatrix} \\ \text{TYPE} = \begin{bmatrix} \text{HEAD} = \textbf{transition} \\ \text{ARGS} = \begin{bmatrix} \text{A}_1 = \textbf{x:agent} \\ \text{A}_2 = \textbf{y:physobj} \\ \text{A}_3 = \textbf{z:location} \end{bmatrix} \\ \text{BODY} = \begin{bmatrix} \text{E}_1 = \mathit{grasp}(x,y) \\ \text{E}_2 = [\mathit{while}(\\ \quad \mathit{hold}(x,y), \\ \quad \mathit{move}(y))] \\ \text{E}_3 = [\mathit{at}(y,z) \rightarrow \\ \quad \mathit{ungrasp}(x,y)] \end{bmatrix} \end{bmatrix} \end{bmatrix}
$$

VoxML for flip

(34)

$$\begin{bmatrix} \textbf{flip} \\ \text{LEX} = \begin{bmatrix} \text{PRED} = \textbf{flip} \\ \text{TYPE} = \textbf{transition\_event} \end{bmatrix} \\ \text{TYPE} = \begin{bmatrix} \text{HEAD} = \textbf{transition} \\ \text{ARGS} = \begin{bmatrix} \text{A}_1 = \textbf{x:agent} \\ \text{A}_2 = \textbf{y:physobj} \end{bmatrix} \\ \text{BODY} = \begin{bmatrix} \text{E}_1 = \mathit{def}(w, \mathit{as}(\mathit{orient}(y)))[\mathit{grasp}(x,y)] \\ \text{E}_2 = [\mathit{while}(\mathit{hold}(x,y), \mathit{rotate}(x,y))] \\ \text{E}_3 = [(\mathit{orient}(y) = \mathit{opp}(w)) \rightarrow \mathit{ungrasp}(x,y)] \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

VoxML for in

$$
\begin{bmatrix}
\textbf{in} \\
\text{LEX} = \begin{bmatrix} \text{PRED} = \textbf{in} \end{bmatrix} \\
\text{TYPE} = \begin{bmatrix}
\text{CLASS} = \textbf{config} \\
\text{VALUE} = \textbf{ProperPart} \parallel \textbf{PO} \\
\text{ARGS} = \begin{bmatrix} \text{A}_1 = \textbf{x:3D} \\ \text{A}_2 = \textbf{y:3D} \end{bmatrix} \\
\text{CONSTR} = \textbf{...}
\end{bmatrix}
\end{bmatrix}
$$

Modeling Action in VoxML

- Object Model: State-by-state characterization of an object as it changes or moves through time.
- Action Model: State-by-state characterization of an actor?s motion through time.
- Event Model: Composition of the object model with the action model.