Conceptual Fingerprints: Lexical Decomposition by Means of Frames – a Neuro-cognitive Model

Wiebke Petersen and Markus Werning

Heinrich-Heine University Düsseldorf
Universitätsstr. 1, 40225 Düsseldorf, Germany
{petersen,werning}@phil.uni-duesseldorf.de

Abstract. Frames, i.e., recursive attribute-value structures, are a general format for the decomposition of lexical concepts. Attributes assign unique values to objects and thus describe functional relations. Concepts can be classified into four groups: sortal, individual, relational and functional concepts. The classification is reflected by different grammatical roles of the corresponding nouns. The paper aims at a cognitively adequate decomposition, particularly, of sortal concepts by means of frames. Using typed feature structures, an explicit formalism for the characterization of cognitive frames is developed. The frame model can be extended to account for typicality effects. Applying the paradigm of object-related neural synchronization, furthermore, a biologically motivated model for the cortical implementation of frames is developed. Cortically distributed synchronization patterns may be regarded as the fingerprints of concepts.

1 Introduction

If one does not want to assume lexical atomism – the view that the possession of any concept expressible by a simple word is completely independent of the possession of any other concept – the question arises in which particular way the possession of some lexical concepts depends on the possession of other concepts. An explicit answer to that question should ideally be (i) in accordance with linguistic data, (ii) formally explicit, (iii) cognitively plausible, and (iv) neurobiologically realistic.

In this paper we will outline a theory of lexical decomposition that attempts to fulfil the four desiderata. Driven by linguistic considerations on the grammatical role of nouns, we will begin with a classification of lexical concepts into four groups. For our account of lexical decomposition, we will use Barsalou's (1992) cognitive frame theory as a point of departure. We will show how frames can be rendered by labeled graphs and how this graphical structure is transformed into a formally explicit typed-feature structure. Concentrating on frames for concepts which linguistically are expressed by nouns, our project aligns with well-established graph-based knowledge representation formalisms that focus on *situations* as in frame semantics (Fillmore, 1982) and *propositions* as with conceptual graph theory (Sowa, 1984). Our formalism is guided by Guarino's (1992) considerations on the ontological status of attributes in frames. To match our

U. Priss, S. Polovina, and R. Hill (Eds.): ICCS 2007, LNAI 4604, pp. 415–428, 2007.

approach with psychological results on categorization (J. D. Smith & Minda, 2000, for review), we digress from a classical Aristotelian interpretation of lexical decomposition, which proposes a definitional relation between concepts and their constituents. Instead, our theory will accommodate cognitive typicality effects regarding concept satisfaction. In contrast to decompositional approaches in prototype theory (E. E. Smith & Medin, 1981), which render concepts by flat feature lists, our frame-theoretic approach allows for a much deeper hierarchical structure. The last part builds on neurobiological evidence that in earlier work has already been proposed to support semantic structure (Werning, 2005). Using oscillatory neural networks as a model, we will show how frames might be implemented in the cortex.

2 Nouns and the Classification of Concepts

Concepts can be distinguished with respect to both arity and referential uniqueness (Löbner, 1985). Sortal and individual concepts are of arity 1 and thus typically have no possessor argument. Sortal concepts (e.g., 'apple') denote classical categories and fail to have unique referents. Individual concepts (e.g., 'Mary'), in contrast, have unique referents. Concepts with arity greater than 1 comprise all relational concepts including functional concepts. It is characteristic for relational concepts (e.g., 'brother') that their referents are given by a relation to a possessor ('brother of Tom'), while unique reference is not generally warranted. Functional concepts (e.g., 'mother') form a special case of uniquely referring relational concepts. They establish a right-unique mapping from possessors to referents (Fig. 1).

	non-unique reference	**unique reference**
arity:1	**SC**: sortal concepts: *person*, *house*, *verb*, *wood*	**IC**: individual concepts: *Mary*, *pope*, *sun*
arity:>1	**RC**: (proper) relational concepts: *brother*, *argument*, *entrance*	**FC**: functional concepts: *mother*, *meaning*, *distance*, *spouse*

Fig. 1. The classification of concepts

The classification of concepts typically corresponds to specific grammatical properties of the expressing noun itself or its context. In English, nouns expressing concepts without unique reference (SCs and RCs) are typically used without definite article. Nouns expressing concepts of higher arity (RCs and FCs) are typically used in possessive constructions, where the possessor is specified by a genitive (*the cat's pow*) or prepositional phrase (*the pow of the cat*).

There is considerable variation in the expression of definiteness and possession across languages. Languages that lack definite articles often employ other strategies to indicate definiteness. In Russian, e.g., word order can be used to signal that a noun refers unambiguously. Here, the preverbal position of a noun phrase hints at a definite interpretation. Hence a question such as *What is on the table?* is likely to be answered by

(1) *Na stol'-e l'ež-it knig-a.*
on table-PREP lie-3SG.PRES book-NOM.SG
'There is a book on the table'

In contrast, *Where is the book?* is likely to be answered by

(2) *Knig-a l'ež-it na stol'-e.*
book-NOM.SG lie-3SG.PRES on table-PREP
'The book is on the table'

To express possession, a manifold of strategies is used as well. Hungarian, e.g., displays morphological agreement of the possessum with the possessor (quoted from Ortmann, 2006):

(3) a. *a te kalop-od*
DF PRON.2SG hat-P'OR.2SG
'your hat'

b. *a Péter kalop-ja*
DF PRON.1SG hat-P'OR.3SG
'Peter's hat'

A suffix (here, *-od* and *-ja*) is attached to the possessed noun, thus specifying agreement with the possessor with respect to the features number and person.

Languages with alienability splits such as the Hokan language Eastern Pomo distinguish overtly whether the concept of the possessed object is conceptualized as being of arity equal or greater one (quoted from Ortmann, 2006):

(4) a. *wí-bayle*
1SG-husband
'my husband'

b. *wáx ša?ri*
PRON.1SG.GEN BASKET
'my basket'

If the noun is conceptualized as being relational ('husband'), it will enter the inalienable possessive construction: the possessor is simply realized by a prefix attached to the possessed noun. In case of alienable possession by contrast the noun ('basket') is not conceptualized as being relational. Possession cannot be expressed on the word level, but rather on the phrase level, by means of a free pronoun.

The type of a given concept may be shifted according to context: The noun *father*, which, in its normal use, has unique reference and arity 2 and thus expresses a functional concept, can be used in contexts like *Fathers don't like cooking* or *The fathers of the constitution were wise*, in a way expressing a sortal or a relational concept, respectively. In some languages (e.g., Yucatec Maya) those type shifts are even overtly realized (Ortmann, 2006).

3 Non-relational Frames

Following Minsky (1975) and Barsalou (1992), frames as recursive attribute-value structures are a general format to account for mental concepts. Guided by the

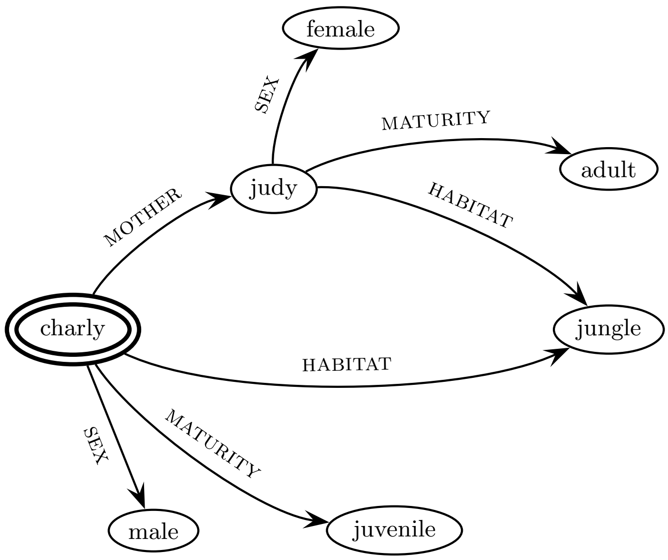

Fig. 2. A knowledge representing frame drawn as a labeled graph

above mentioned demands that our concept-decomposition framework should be formally explicit and cognitively adequate, we aim at keeping our frame model as simple and rigid as possible. We do not want to introduce any elements in our model language solely due to technical or computational reasons or for the sake of generality and expressibility. In section 4 and section 5, rather, we will point to cognitive and neuro-biological evidence for our model language. As our aim is to decompose *concepts* we rest our frame model on the restrictive theory of typed feature structures (Carpenter, 1992) and not on the much wider framework of conceptual graphs (Sowa, 1984). We are aware, though, that all our frames can be easily transformed into conceptual graphs (but not *vice versa*). In accordance with Barsalou (1992) we, for our frame model, assume that attributes assign unique values to objects and thus describe functional relations. The values can be structured frames themselves. Attributes in frames are therefore functional concepts and embody the concept type on which the categorization is based.

We model non-relational frames as connected directed acyclic rooted graphs with labeled nodes (value types) and arcs (attributes).[1] Fig. 2 shows the graph of an example frame representing knowledge about a young male gorilla Charly living with his mother Judy in the same jungle. The double-encircled node ‘charly’ points out that the graph represents a frame about Charly. The outgoing arcs of the ‘charly’-node stand for the attributes of Charly and point to their

[1] The definition of frames as directed rooted graphs enables us to adopt the theory of typed feature structures, which is well-established in computational linguistics. Our definitions follow Carpenter (1992) as closely as possible, except for definition 5, which digresses in one point fundamentally.

values. Hence, the sex of Charly is male and the maturity of Charly is juvenile. The value of the attribute 'mother' is a complex frame itself, describing that Judy, the mother of Charly, is female and adult. The fact that Judy and Charly live in the same jungle is indicated by the single 'jungle'-node to which the two 'habitat'-arcs from 'charly' and 'judy' point.

Definition 1. *Given a set* TYPE *of types and a finite set* ATTR *of attributes. A* non-relational frame *is a tuple $F = (Q, \bar{q}, \theta, \Phi)$ where:*

- *Q is a finite set of nodes,*
- *$\bar{q} \in Q$ is the root node,*
- *$\theta : Q \to$ TYPE is the total* node typing function*,*
- *Φ : ATTR $\times Q \to Q$ is the partial* transition function*.*

Furthermore, for each $q \in Q$ there be a finite sequence of attributes $A_1 \dots A_n \in$ ATTR* with $\Phi(A_n, \dots, \Phi(A_2, \Phi(A_1, \bar{q})) \dots) = q$, i.e., q and $\bar{q}$ are connected by a finite path; and for no $q \in Q$ there be a finite sequence of attributes $A_1 \dots A_n \in$ ATTR* with $\Phi(A_n, \dots, \Phi(A_2, \Phi(A_1, q)) \dots) = q$, i.e., the graph is acyclic.

The root node of a non-relational frame is its referring node. If $\theta(\bar{q}) = t$, we say that the frame is of type t. A node with no outgoing arcs is called an *end node* of the frame. To be able to speak of the paths of a frame, we need the following definition:

Definition 2. *Given a set* TYPE *of types, a finite set* ATTR *of attributes, and a non-relational frame $F = (Q, \bar{q}, \theta, \Phi)$. A sequence of attributes $A_1 \dots A_n \in$ ATTR* is a* path *of F if $\Phi(A_n, \dots, \Phi(A_2, \Phi(A_1, \bar{q})) \dots)$ is defined. The set of all paths of a frame F is denoted by Π_F. A path $\pi \in \Pi_F$ is said to be* maximal *in F if $\pi A \notin \Pi_F$ for all attributes A.* MAXPATH$_F$ *denotes the set of maximal paths in F. The node typing function θ can be extended to the* path typing function *$\Theta : \Pi_F \to$ TYPE in a natural way:*

$$\Theta(A_1 \dots A_n) = \theta(\Phi(A_n, \dots, \Phi(A_2, \Phi(A_1, \bar{q})) \dots)).$$

Since the information represented by a frame does not depend on the concrete set from which the nodes are drawn, we can abstract away from this set and focus on how the nodes are connected by labeled arcs. Fig. 3 shows the frame of Fig. 2 represented as an recursive attribute-value-matrix (AVM). The AVMs are constructed as follows: Frames are enclosed in square brackets with an index denoting the type of the root node. Each first-level attribute is stated in the brackets followed by a colon and followed by the value of the attribute. The values are either complex frames themselves or unstructured (i.e., not specified by further attributes). In the case of an unstructured value, we write 'ATTRIBUTE:type' instead of 'ATTRIBUTE:[ ]$_{type}$'. The symbol $\boxed{1}$ indicates that the path [MOTHER:HABITAT:] starting from the root node points to the same node as the path [HABITAT:], i.e., the two paths share the same value.

The types are ordered in a type hierarchy, which induces a subsumption order on frames. Cognitively the types correspond to categories and the type hierarchy to an IS-A-hierarchy.

$$\begin{bmatrix} \text{MOTHER}: & \begin{bmatrix} \text{HABITAT}: & \boxed{1} \\ \text{SEX}: & female \\ \text{MATURITY}: & adult \end{bmatrix}_{judy} \\ \text{HABITAT}: & \boxed{1}\ jungle \\ \text{MATURITY}: & juvenile \\ \text{SEX}: & male \end{bmatrix}_{charly}$$

Fig. 3. AVM-abstraction of the frame-graph of Fig. 2

Definition 3. *A* type hierarchy $(\text{TYPE}, \sqsupseteq)$ *is a partial ordered set which forms a join semilattice, i.e., for any two types there exists a least upper bound.*

A type t_1 is a subtype *of a type t_2 if $t_1 \sqsupseteq t_2$. A type t is said to be* minimal *if it has no subtypes. The set of minimal types is denoted by* MinType.

Definition 4. *Given a type hierarchy* $(\text{TYPE}, \sqsupseteq)$ *and a finite set* ATTR *of attributes. A frame $F = (Q, \bar{q}, \theta, \Phi)$* subsumes *a frame $F' = (Q', \bar{q}', \theta', \Phi')$, notated as $F \sqsubseteq F'$, iff there exists a total function $h : Q \to Q'$ such that:*

- *$h(\bar{q}) = \bar{q}'$,*
- *for each $q \in Q$: $\theta(q) \sqsubseteq \theta'(h(q))$,*
- *if $q \in Q$, $a \in$ ATTR, and if $\Phi(a, q)$ is defined, then $h(\Phi(a, q)) = \Phi'(a, h(q))$.*

The following example shows an unspecified 'ape'-frame subsuming an unspecified 'gorilla'-frame, which subsumes the fully specified 'charly'-frame (see the type hierarchy in Fig. 4):

$$\begin{bmatrix} \text{HABITAT}: & habitat \\ \text{SEX}: & sex \end{bmatrix}_{ape} \sqsubseteq \begin{bmatrix} \text{HABITAT}: & jungle \\ \text{SEX}: & sex \end{bmatrix}_{gorilla} \sqsubseteq \begin{bmatrix} \text{HABITAT}: & jungle \\ \text{SEX}: & male \end{bmatrix}_{charly}.$$

As Guarino (1992) points out, frame-based knowledge engineering systems as well as feature-structure-based linguistic formalisms normally force a radical choice between attributes and types. As a consequence, generic frames like

$$\begin{bmatrix} \text{MATURITY}: & maturity \\ \text{HABITAT}: & jungle \end{bmatrix}_{gorilla}$$

occur frequently in which the unspecified value 'maturity' is assigned to the attribute 'MATURITY'. The parallel naming of the attribute 'MATURITY' and the type 'maturity' pretends a systematic relationship between the attribute and the type which is not intended by the formalism.

A second problem addressed in Guarino (1992) concerns the question which binary relations should be expressed by attributes. If one allows attributes to be unrestricted arbitrary binary relations, this leads to frames like the following one, which was first discussed in Woods (1975):

$$\begin{bmatrix} \text{HEIGHT}: & 6 feet \\ \text{HIT}: & mary \end{bmatrix}_{john}.$$

Although 'HEIGHT' and 'HIT' can be represented by binary predicates, the ontological status of the link they establish between 'john' and '6 feet' and between 'john' and 'mary' respectively differs fundamentally.

Our main thesis on frames is that non-relational frames decompose non-relational concepts into functional concepts. But our definition of non-relational frames only uses attributes for the decomposition. Hence, the question arises how attributes and functional concepts are connected. All sample attributes we have used so far (MOTHER, SEX, ...) correspond to functional concepts. Guarino (1992) distinguishes between the *denotational* and the *relational* interpretation of a relational concept. This distinction can be used to explain how functional concepts can act as concepts and as attributes: Let there be a universe $\mathcal{U}$ and a set of functional concepts $\mathcal{F}$. A functional concept (like any concept) denotes a set of entities:

$$\delta : \mathcal{F} \rightarrow 2^{\mathcal{U}}$$
$$\text{(e.g., } \delta(\text{mother}) = \{m \,|\, m \text{ is the mother of someone}\}).$$

A functional concept also has a relational interpretation:

$$\varrho : \mathcal{F} \rightarrow 2^{\mathcal{U} \times \mathcal{U}}$$
$$\text{(e.g., } \varrho(\text{mother}) = \{(p, m) \,|\, m \text{ is the mother of } p\}).$$

The denotational and the relational interpretation of a functional attribute have to respect the following *consistency postulate*: Any value of a relationally interpreted functional concept is also an instance of the denotation of that concept. (If $(p, m) \in \varrho(\text{mother})$, then $m \in \delta(\text{mother})$). Furthermore, the relational interpretation of a functional concept f is a function, i.e., if $(a, b), (a, c) \in \varrho(f)$, then $b = c$.

These considerations allow us, to clarify the ontological status of attributes in frames: Attributes in frames are relationally interpreted functional concepts! Hence, attributes are not frames themselves and are therefore unstructured. Frames of non-relational concepts decompose into relationally interpreted functional concepts.

In order to restrict the class of admissible frames, the plain type hierarchy can be enriched by an appropriateness specification. It regulates which attributes are appropriate for frames of a special type and restricts the values of the appropriate attributes.[2] Our definition of type signatures consequently dismisses the artificial distinction between attributes and types in contrast to the standard definition (Carpenter, 1992): the attribute set is merely a subset of the type set. Hence, attributes occur in two different roles: as names of binary functional relations between types and as types themselves.

[2] Type signatures can be automatically induced from sets of untyped non-relational frames, i.e. frames in which only the maximal paths are typed. With FCAType an implemented system for such inductions is available, which uses formal concept analysis (Kilbury, Petersen, & Rumpf, 2006; Petersen, 2006, 2007).

Definition 5. *Given a type hierarchy* $(\text{TYPE}, \sqsupseteq)$ *and a set of attributes* $\text{ATTR} \subseteq \text{TYPE}$*. An* appropriateness specification *on* $(\text{TYPE}, \sqsupseteq)$ *is a partial function* $\text{Approp} : \text{ATTR} \times \text{TYPE} \to \text{TYPE}$ *such that for each* $a \in \text{ATTR}$ *the following holds:*

– ***attribute introduction:*** *There is a type* $\text{Intro}(a) \in \text{TYPE}$ *with:*
  - $\text{Approp}(a, \text{Intro}(a)) = a$ *and*
  - *for every* $t \in \text{TYPE}$*: if* $\text{Approp}(a, t)$ *is defined, then* $\text{Intro}(a) \sqsubseteq t$*.*

– ***specification closure:*** *If* $\text{Approp}(a, s)$ *is defined and* $s \sqsubseteq t$*, then* $\text{Approp}(a, t)$ *is defined and* $\text{Approp}(a, s) \sqsubseteq \text{Approp}(a, t)$*.*

– ***attribute consistency:*** *If* $\text{Approp}(a, s) = t$*, then* $a \sqsubseteq t$*.*

A type signature *is a tuple* $(\text{TYPE}, \sqsupseteq, \text{ATTR}, \text{Approp})$*, where* $(\text{TYPE}, \sqsupseteq)$ *is a type hierarchy,* $\text{ATTR} \subseteq \text{TYPE}$ *is a set of attributes, and* $\text{Approp} : \text{ATTR} \times \text{TYPE} \to \text{TYPE}$ *is an appropriateness specification. A type* t *is said to be* atomic *if* $\text{Approp}(a, t)$ *is undefined for any* $a \in \text{ATTR}$*.*

The first two conditions on an appropriateness specification are standard in the theory of type signatures (Carpenter, 1992), except that we tighten up the attribute introduction condition by claiming that the introductory type of an attribute a carries the appropriateness condition '$a : a$'. By the attribute consistency condition we ensure that Guarino's consistency postulate holds and that Barsalou's view on frames, attributes, and values is modeled appropriately:

> At their core, frames contain attribute-value sets. Attributes are concepts that represent aspects of a category's members, and values are subordinate concepts of attributes, (Barsalou, 1992).

Hence, the possible values of an attribute are subconcepts of the denotationally interpreted functional concept. This is reflected in the type signature by the condition that the possible values of an attribute are restricted to subtypes of the type corresponding to the attribute.

A small example type signature is given in Fig. 4. The appropriateness specification is split-up into single appropriateness conditions:[3] The expression 'SEX:sex' at type 'ape' means that the attribute 'SEX' is appropriate for frames of type 'ape' and its value is restricted to 'sex', hence, $\text{Approp}(\text{SEX}, ape) = sex$. The attribute conditions are passed on downwards. Hence, the type 'gorilla' inherits the appropriateness condition 'SEX:sex' from its upper neighbor 'ape'. It also inherits the appropriateness condition 'HABITAT:habitat', but tightens it up to 'HABITAT:jungle', which is permissible by the specification closure condition. The definition of the type signature makes sure that the permissible values of an attribute are subtypes of the attribute type. Hence, the possible values of SEX, i.e., 'female' and 'male', are subtypes of the type 'sex'. Notice that the subtypes of an attribute type are not generally attribute types themselves.

[3] To improve readability we mark the two roles of attributes in our frame notation: attributes used as types are written in small letters and attributes used as attributes in capitals.

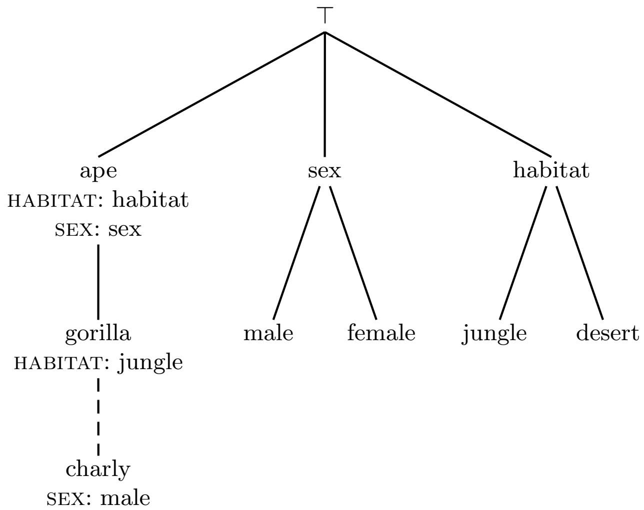

Fig. 4. Example type signature

A frame whose end nodes are all labeled by atomic minimal types is said to be a *fully-specified frame*. We call a non-relational frame *well-typed* with respect to a type signature if all attributes of the frame are licensed by the type signature and if additionally the attribute values are consistent with the appropriateness specification. The definition of the appropriateness specification guarantees that every arc in a well-typed frame points to a node which is typed by a subtype of the type corresponding to the attribute labeling the arc. The decomposition of concepts into frames requires that the frame in question be well-typed.

4 Frames and Typicality

One of the main virtue of frames is that they allow the decomposition of sortal and individual concepts by means of functional concepts. This decomposition now enables us to explain how a subject may subsume a perceived or otherwise given object under a sortal or individual concept. The degree, between 0 and 1, to which an object of the universe $\mathcal{U}$ instantiates a certain type is given by the function:

$$d : \text{TYPE} \times \mathcal{U} \rightarrow [0, 1].$$

In every frame the root node corresponds to the decomposed concept ('charly', 'cherry'). The set of maximal paths MAXPATH $\subseteq \Pi$ is well-defined for every frame. In a fully specified frame, end nodes, e.g, 'red', are atomic minimal types and are identified by maximal paths, i.e., [COLOR:HUE:] beginning at the root node, i.e., 'cherry' (Fig. 5).

It is natural to assume that the cognitive subject is endowed with a detector system that for all atomic minimal types renders the degree d to which it is instantiated by a given object. These might be hue detectors, sex detectors etc.

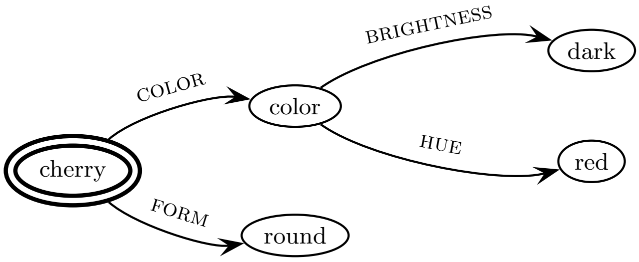

Fig. 5. Cherry frame. Example for a frame without reference shift.

It is important to notice that many attributes when applied to an object shift the referential object. One may say that the hue of the color of a cherry is still a property of the cherry and a hue detector may well be directed to the cherry in order to assign a value. However, the sex of the mother of Charly is not a property of Charly, and detecting the value of the sex of the mother of Charly, requires a potential sex detector to be directed to the mother. It is hence useful to introduce a reference-shifting function

$$\sigma : \mathcal{U} \times \Pi \rightarrow \mathcal{U}$$

that maps every object of the universe relative to the path in question onto the same or another object of the universe.

In the classical bi-valued case, the values of d are restricted to 0 and 1. Here, for a fully specified frame, where $\Theta(\textsc{MaxPath}) \subseteq \textsc{MinType}$, we can conclude that an object x is to be subsumed under the decomposed concept C if and only if all the types of the end nodes are properly instantiated:

$$d(C, x) = \min_{m \in \textsc{MaxPath}} d(\Theta(m), \sigma(x, m)).$$

A cognitively more realistic picture, however, is attained if we specify how typical a certain minimal type is for instances of the concept. Red may, e.g., be more typical than green as the hue of the color of cherries. Nevertheless the hue of the color of some cherries still is green. We can achieve this by considering alternative types for each maximal path. For each maximal path m we then have a set $\textsc{alt}(m)$ containing the minimal type $\Theta(m)$ and all its alternative types with regard to the path. I.e., provided that $m \in \textsc{attr}^*$ is a maximal path of a frame with $m = A_1 \ldots A_n$, $\textsc{alt}(m)$ is the set of all atomic minimal subtypes of the type A_n. Our definition of the type signature guarantees that $\textsc{alt}(m)$ covers all possible maximally specified values of A_n. For each of the types $t \in \textsc{alt}(m)$ we can then specify a typicality value relative to the maximal path m of the fully specified frame for a decomposed concept C. The typicality value $\tau(C, m, t)$ tells how typical the type t is for the object $\sigma(x, m)$ given that x instantiates C. With these conventions we can apply previous results of Werning and Maye (2005, 2007) and, on the basis of the detector outputs, estimate to which degree an object x instantiates the decomposed concept C:

$$d(C,x) \geq \min_{m \in \textsc{MaxPath}} \max_{t \in \textsc{alt}(m)} \tau(C,m,t)\, d(t, \sigma(x,m)).$$

The proof of the theorem, firstly, requires that the types of $\textsc{alt}(m)$, for each maximal path m, are a quasi partitioning of the referentially appropriately shifted universe $\mathcal{U}$. I.e., for all $x \in \mathcal{U}$

$$\sum_{t \in \textsc{alt}(m)} d(t, \sigma(x,m)) = 1.$$

In the equation the $\geq$-direction reflects exhaustivity and the $\leq$-direction reflects exclusivity. A cognitively appropriate type signature guarantees that each subtype of an attribute type is a reasonable value of the attribute. Furthermore, the attribute consistency condition on appropriateness warrants that all values of an attribute are subtypes of the attribute type. However, it does not follow that the minimal subtypes of an attribute type exhaust the values of the attribute ('red' could have two subtypes 'light red' and 'dark red' such that something red could be neither light red nor dark red). However, such a situation is excluded if we only consider type signatures automatically induced from untyped object frames by the system of Petersen (2007). Exclusivity is warranted by the fact that all types in $\textsc{alt}(m)$ are minimal. A case where 'light red', 'dark red', and 'red' all occur as types in $\textsc{alt}(m)$ is thus excluded.

Secondly, we have to presuppose that the extension of the decomposed concept is completely determined by the extensions of the minimal types at the end nodes of the fully specified frame. This is to say that the extensions of the types at intermediate nodes do not independently bear on the extension of the decomposed concept. However, it is not to say that the extension of the decomposed concept does not depend on the extensions of types at intermediate nodes. For, the extensions of those in turn depend on the extensions at the end nodes. The condition of complete determination again holds trivially for type signatures induced from sets of untyped frames as described by Petersen (2007).

5 Neuro-cognitive Interpretation

For many attributes (HUE, BRIGHTNESS, ORIENTATION, DIRECTION, SIZE, etc.) involved in the course of visual processing – we call them qualitative attributes – one can anatomically identify so-called neuronal feature maps (Hubel & Wiesel, 1968). These are structures of neurons that exhibit a certain topological organization. With regard to one attribute or feature dimension one finds a pinwheel-like structure for each receptive field (i.e., a specific region of the stimulus). This structure is called a hypercolumn. For each receptive field and each such attribute (e.g., HUE) we find a hypercolumn such that neurons for the entire spectrum of subtypes ('red', 'green', etc.) of that attribute fan out around a pin-wheel center. Neurons of a hypercolumn with a tuning for one and the same feature or subtype (e.g., 'red') form a so-called column. We may assume that such neurons function as detectors and thus evaluate atomic minimal types for a given stimulus object.

More than 30 cortical areas forming feature maps are experimentally known to be involved in the visual processing of the monkey (Felleman & van Essen, 1991). These findings justify the hypothesis that in the cortex there may be neural correlates of attributes and their subtypes.

The fact that subtypes of different attributes may be instantiated by the same stimulus object, but are processed in distinct regions of cortex poses the problem of how this information is integrated in an object-specific way. How can it be that the horizontality and the redness of a red horizontal bar are represented in distinct regions of cortex, but still are part of the representation of one and the same object? This is the binding problem in neuroscience (Treisman, 1996).

A prominent and experimentally well supported solution postulates oscillatory neuronal synchronization as a mechanism for binding (von der Malsburg, 1981; Gray, König, Engel, & Singer, 1989): Clusters of neurons that are indicative for different properties sometimes show synchronous oscillatory activity, but only when the properties indicated are instantiated by the same object in the perceptual field; otherwise they are firing asynchronously. Synchronous oscillation, thus, might be regarded to fulfill the task of binding together various property representations in order to form the representation of an object as having these properties (for a review see Singer, 1999).

Using oscillatory networks (Schillen & König, 1994; Maye & Werning, 2004) as models, the structure of object-related neural synchronization could be interpreted (Werning, 2005) as providing a conceptual structure expressible in a first-order predicate language. To show this, an eigenmode analysis of the network dynamics is computed. Per eigenmode, oscillation functions play the role of object representations or concepts. Clusters of feature sensitive neurons play the role of property representations or predicate concepts. Werning (2003) extends this approach from an ontology of objects to an ontology of events. Werning and Maye (2006) discuss ambiguous and illusionary representations. The following theorem (Werning & Maye, 2007) nicely links the results of this paper to previous results on the neural implementation of conceptual structure. The degree to which the object x is represented as instantiating the atomic type t by a network eigenmode is given by the equation:

$$d(t, x) = \max\{\Delta(\alpha(\mathrm{x}), f_j) | \mathbf{f} = \beta(t) c \mathbf{v}\}.$$

Here $\alpha(\mathrm{x})$ is the oscillation function representing the object x, and $\beta(t)$ is a matrix identifying the neural clusters which function as detectors for the type t. $\mathbf{v}$ and c are the results of the eigenmode analysis and account for the spatial, respectively, temporal variation of the network activity in that eigenmode. Δ is defined as the normalized inner product of two square-integrable time-dependent functions in a given temporal interval and measures the degree of synchrony between an object-related oscillation and the actual oscillatory activity in a neural cluster. $d(t, x)$ approaches 1 if the oscillation function $\alpha(\mathrm{x})$, which represents the object x, is highly synchronous with some component oscillatory activity f_j of

f – i.e., the vector containing the eigenmode-relative temporal evolution of the type-related cluster of detector neurons.

If we conjoin the estimation of $d(C, x)$ in terms of type-specific detector outputs $d(t, x)$ with the identification of the latter with particular oscillatory network activity, we may conclude with the following hypothesis: Provided that a concept is completely decomposable into a fully specified frame with detectors for each type of a maximal path, the degree to which the cortex represents an object as an instance of the concept can be estimated by a general pattern of synchronizing neural activity distributed over various feature-selective neural clusters that correspond to the atomic types of the frame. This pattern may be called the cortical fingerprint of the concept.

References

Barsalou, L.W.: Frames, concepts, and conceptual fields. In: Lehrer, A., Kittay, E.F. (eds.) Frames, fields, and contrasts, pp. 21–74. Erlbaum, Hillsdale, NJ (1992)

Carpenter, B.: The logic of typed feature structures. Cambridge University Press, Cambridge (1992)

Fillmore, C.J.: Frame semantics. In: Linguistics in the morning calm, pp. 111–137. Hanshin Publishing Co. Seoul (1982)

Gray, C., König, P., Engel, A.K., Singer, W.: Oscilliatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338, 334–337 (1989)

Gray, C., König, P., Engel, A.K., Singer, W.: Oscilliatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338, 334–337 (1989)

Guarino, N.: Concepts, attributes and arbitrary relations: some linguistic and ontological criteria for structuring knowledge bases. Data Knowl. Eng. 8(3), 249–261 (1992)

Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology 195, 215–243 (1968)

Kilbury, J., Petersen, W., Rumpf, C.: Inheritance-based models of the lexicon. In: Wunderlich, D. (ed.) Advances in the theory of the lexicon, pp. 429–477. Mouton de Gruyter, Berlin (2006)

Löbner, S.: Definites. Journal of Semantics 4(4), 279–326 (1985)

Maye, A., Werning, M.: Temporal binding of non-uniform objects. Neurocomputing 58–60, 941–948 (2004)

Minsky, M.: A framework for representing knowledge. In: Winston, P.H. (ed.) The psychology of computer vision, McGraw-Hill, New York (1975)

Ortmann, A.: (In) Alienabilitätssplits und die Typologie der adnominalen Possession (Manuscript, University of Düsseldorf) (2006)

Petersen, W.: FCAType – a system for type signature induction. In: Yahia, S.B., Nguifo, E.M. (eds.) 4th international conference on concept lattices and their applications (2006)

Petersen, W.: Inducing type signatures from decomposition lattices. International Journal of Foundations of Computer Science, 2007 (submitted)

Schillen, T.B., König, P.: Binding by temporal structure in multiple feature domains of an oscillatory neuronal network. Biological Cybernetics 70, 397–405 (1994)

Singer, W.: Neuronal synchrony: A versatile code for the definition of relations? Neuron 24, 49–65 (1999)

Smith, E.E., Medin, D.L.: Categories and concepts. Harvard University Press, Cambridge, MA (1981)

Smith, J.D., Minda, J.: Thirty categorization results in search of a model. Journal of Experimental Psychology: Learning, Memory and Cognition 26(1), 3–27 (2000)

Sowa, J.: Conceptual structures. Addison-Wesley, Reading MA (1984)

Treisman, A.: The binding problem. Current Opinion in Neurobiology 6, 171–178 (1996)

von der Malsburg, C.: The correlation theory of brain function (Internal Report Nos. 81-2). MPI for Biophysical Chemistry, Göttingen (1981)

Werning, M.: Ventral vs. dorsal pathway: the source of the semantic object/event and the syntactic noun/verb distinction. Behavioral and Brain Sciences 26(3), 299–300 (2003)

Werning, M.: The temporal dimension of thought: Cortical foundations of predicative representation. Synthese 146(1/2), 203–224 (2005)

Werning, M., Maye, A.: Frames, coherency chains and hierarchical binding: The cortical implementation of complex concepts. In: Bara, B.G., Barsalou, L., Bucciarelli, M. (eds.) Proceedings of the twenty-seventh annual Conference of the Cognitive Science Society, pp. 2347–2352. Erlbaum, Mahwah, NJ (2005)

Werning, M., Maye, A.: The neural basis of the object concept in ambiguous and illusionary perception. In: Sun, R., Miyake, N. (eds.) Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society, pp. 876–881. Erlbaum, London (2006)

Werning, M., Maye, A.: The cortical implementation of complex attribute and substance concepts: Synchrony, frames, and hierarchical binding. Chaos and Complexity Letters 2(2) 2007 (in press)

Woods, W.A.: What's in a link: Foundations for semantic networks. In: Bobrow, D.G., Collins, A.M. (eds.) Representation and understanding: Studies in cognitive science, Academic Press, New York (1975)