From: AAAI Technical Report SS-01-05. Compilation copyright © 2001, AAAI (www.aaai.org). All rights reserved.

Generating Grounded Representations: A Mathematical Schema

Zippora Arzi-Gonczarowski
Typographics, Ltd. Jerusalem 96222 Israel
zippie@actcom.co.il
http://www.actcom.co.il/typographics/zippie

Abstract

This extended abstract digests a formal mathematical schema for the generation of representations over basic sensory motor neural perceptions. This representation can then be applied for a broad spectrum of cognitive and behavioral mechanisms. Category theoretical constructs provide tools of rigour that capture a structural essence but avoid over determinism. The mathematical setting systematizes the grounding of taxonomic hierarchies, compound concepts, and organization for both access and inference. It also provides computational basis for the adjustment of the representation in view of perceptual changes.

Introduction

When an agent is provided with a method to generate representations of eventual ulterior environments, one does not enjoy the advantage of having exact knowledge of these environments. A domain general methodology for the generation of representations could not even enjoy the advantage of knowing the details of the actual agent: the nature of its sensing abilities, the phenomena and the discriminations that will be consequential to its survival and to the activities that it will be expected to engage in, etc. Hence, this methodology should be general enough to be applied to a broad spectrum of environments, agents, and applications. The prescription for representation generation should hence be schematic, leaving open many details, until actual substitution instances eventually provide an express description of specific environments, agents, etc. On the other hand, to be actually implemented in programmed systems, one needs a methodology that is clear, precise, and testable. The versatility could not be achieved through vagueness and obscurity. The answer that is proposed here consists of a formal schema, where *formal* stands for *general+rigorous*.

Mathematical category theory typically provides meticulous tools of rigour to capture a structural essence without being over deterministic. It has already been successfully applied to various issues in computer science, like programming language semantics and the design of programs using abstract data types. It has come to provide a standard ontology and language of discourse for these areas of research, that also require that rigour be applied across a broad spectrum of domains. The proposed categorical formalism has been conceived to apply mathematical tools for the generation of grounded representations, and for cognitive operations on represented information. The theoretical standard can guide the design and implementation of systems that construct their own representations from raw sensory-motor-neural information, and use them for intelligent, cognitive and behavioral, processes.

This is a summary of papers that have been published in the last two years (Arzi-Gonczarowski & Lehmann 1998a; 1998b; Arzi-Gonczarowski 1998; 1999a; 1999b; 2000a; 2000b). It is submitted to the symposium on learning grounded representations to present ISAAC[1] as part of the state of the art in this area of research. The versatility of ISAAC has allowed a natural integration of examples from an independent reasearch about grounded perception (Oates, Schmill, & Cohen 2000), that has not been known to this author until recently.

Pretheory

Philosophical discussions of mental representation date back at least to the 17th century, to Malebranche and Locke. Their view, that mental entities could represent features of the world that is perceived, was bitterly criticized by philosophers who doubted whether representations could ever acquire genuine content, pointing beyond themselves (Arnauld, Foucher, Berkeley, and more recently, (Searle 1984)). In modern research, mental representations are central to AI and cognitive science. The work described here shares some prevalent views, (Palmer 1978; Pierce 1931 1958) among others, that mental representations acquire their content by being grounded meaningfully in direct experience and in motor-neural sensation. Representations should have a realization in the agent that is useful as a basis for cognitive activities. Representations are *about* things in the environment, without confusing the representation with the thing that is represented.

From the methodological point of view, the proposed formalism starts from perception. High level perception is, loosely, the end result of sensation and (some degree of) comprehension. It is suggestive that *comprehension* stands

Copyright © 2000, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

[1] ISAAC: Integrated Schema for Affective Artificial Cognition.

for both *understanding* and *inclusion*. It is the inclusion of the (encoding of a) sensation within some internal container, as opposed to sensations that are not *comprehended* internally. The pretheoretical intuition is that representations emerge from such 'comprehended containers' by means of sorting and organizing them into meaningful structures.

According to (Billman 1998) a representation is primarily characterized by its content, organization, and format. These issues will be addressed in the context of ISAAC. Once a generation of a representation over basic perception is established, its properties may be studied. For example, one may ask about perceptual change that could cause a change in representation: does the formalism provide tools for the adjustment of the representation in that case? It would be nice if the extent of the adjustment were proportional to the extent of the change that caused it. Ultimately, the representation is needed as a basis for various activities, and the generated representation has been shown to be theoretically capable of serving its purpose.

Formal Premises

In this section the basic concepts of the formalism are briefly sketched. Extensive discussions are provided in the publications that were cited in the introduction.

A *Perception* is defined as a 3-tuple $\mathcal{P} = \langle \mathcal{E}, \mathcal{I}, \varrho \rangle$ where $\mathcal{E}$ and $\mathcal{I}$ are finite, disjoint sets, and ϱ is a 3-valued predicate $\varrho : \mathcal{E} \times \mathcal{I} \rightarrow \{t, f, u\}$.

The set $\mathcal{E}$ represents a 'snapshot' of a perceived environment, a collection of environmental chunks, most typically objects or events, *world elements* (*w-elements*) that could perhaps be discerned by an agent. Even if the environment exists independent of its perception, its carving up into individuated w-elements typically depends on the agent. Where one perceives one forest, another might perceive many trees. For example, in (Oates, Schmill, & Cohen 2000), some w-elements are *experiences* that involve the moving robot and its interaction with other objects. Other w-elements are visible objects and objects that imped or obstruct the robot's path.

The set $\mathcal{I}$ stands for the building blocks of discriminations, *connotations* that are internally afforded by the perceiving agent. They span the possible *content* of the representation and the ontological distinctions that can eventually be made. In a written paper it is easiest to provide examples where the format for expressing the content of the representation are alpha-numeric strings. For example, an agent that has the ability to sense colors would typically afford internal symbols that encode these colors. However, connotations could consist of other data structures as well: icons, diagrams, pointers, subprograms, etc. For example, in (Oates, Schmill, & Cohen 2000), prototyped graph forms are offered as qualitative discriminations that an agent may make about various events. They are based on its sensory motor apparatus, and provide a basis for an ontology. For humans, that ontology would consist of items such as bumping, pushing, being blocked, moving unbostructed, etc.

The 3-valued *Perception Predicate* (*p-predicate*) ϱ relates w-elements and connotations, where t stands for definitely 'yes', f stands for definitely 'no', and u indicates that perception, for some reason, does not specify whether a certain w-element has a certain connotation. For example, if w is a visual object, then a perception as in (Oates, Schmill, & Cohen 2000) could perhaps indicate that $\varrho(w, approach_on_right) = t$, $\varrho(w, gets_big) = f$, $\varrho(w, vanish_off_left) = u$.

Actual sets $\mathcal{E}$ and $\mathcal{I}$, and values of ϱ, provide a concrete perception. The mathematical objects $\mathcal{P} = \langle \mathcal{E}, \mathcal{I}, \varrho \rangle$ hence stand for basic embodied precognitions. They are high-level in the sense that they are presumed to layer on top, and be grounded by a more basic apparatus. In (Oates, Schmill, & Cohen 2000), for instance, this apparatus consists of the agent's sensors, and a clustering process that categorizes experiences that are recorded by these sensors. ISAAC is about representations and processes that happen at, and above, the level of (some) recognition of individuated wholes or events, where (some) basic perception is grounded, and cognizance begins to play a role.

Environments, agents, and perceptions are dynamic and ever changing. A stream of perceptual states, or paths between them, can be combined into a 'movie' that describes perceptual transitions, either inter- or intra- agent. A flow of perceptions is formalized by *perception morphisms (p-morphisms, arrows)*:

Let $\mathcal{P}_1 = \langle \mathcal{E}_1, \mathcal{I}_1, \varrho_1 \rangle$ and $\mathcal{P}_2 = \langle \mathcal{E}_1, \mathcal{I}_2, \varrho_2 \rangle$ be perceptions. A p-morphism $h : \mathcal{P}_1 \rightarrow \mathcal{P}_2$ consists of the set mappings: $h : \mathcal{E}_1 \rightarrow \mathcal{E}_2$, $h : \mathcal{I}_1 \rightarrow \mathcal{I}_2$, with the structure preservation *No–Blur* condition: For all w in $\mathcal{E}$, and for all α in $\mathcal{I}$, whenever $\varrho_1(w, \alpha) \neq u$ then $\varrho_2(h(w), h(\alpha)) = \varrho_1(w, \alpha)$.

The mapping $\mathcal{E}_1 \rightarrow \mathcal{E}_2$ is a formal tool to describe transitions in the perceived environment, that may be initiated either by external or by internal catalysts. For example, $\mathcal{E}_2$ may feature a new w-element, that has not been perceived in $\mathcal{E}_1$, either because it has just been brought into the scene, or because the former perceptual state did not attend to it, but now it does. The mapping $\mathcal{I}_1 \rightarrow \mathcal{I}_2$ is a formal tool to describe transitions in the internal scene. For example, $\mathcal{I}_2$ may feature a new connotation, indicating a newly learned distinction. In other cases, two connotations from $\mathcal{I}_1$ may be merged into one generalized discrimination (that is what the clustering of (Oates, Schmill, & Cohen 2000) does). Constituents (either in $\mathcal{E}$ or in $\mathcal{I}$) may be added, replaced, merged, etc. The values of the p-predicate may also be modified along arrows, as confined by the *No–Blur* structure preservation condition. (Category theory provides the *mirror category* to traverse arrows also in the opposite direction, blurring perceptual values, deleting or splitting constituents, etc.) An example of an inter-agent perceptual transition can be given from (Oates, Schmill, & Cohen 2000), where $\mathcal{I}_1 \rightarrow \mathcal{I}_2$ maps one-to-one between the robot's connotations, that are prototyped graph forms, and human connotations, that are natural language words and phrases such as *push, squash, can't move,* etc. P-morphisms thus provide a versatile tool for the formal description of many types of perceptual cognitive processes. Quite a few additional types of transitions are described in the cited papers.

Arrows may be *factorized* into simpler ones. Simple factors may consist, for example, of modifying or translating a single constituent, or unblurring a specific value of the p-

predicate, etc. Pursuing the 'movie' metaphor from above, this is analogous to the movement of a cartoon character, a movement that may be broken into smaller movements of every joint. The operation complementary to factorization is *composition* of p-morphisms, a fundamental categorical tool.

Technically, composition and the identity p-morphism are defined by composition and identity of set mappings, and perceptions with p-morphisms make a mathematical category, designated $\mathcal{P}rc$. This provides a well developed mathematical infrastructure for a theory of artificial perceptual transitions. Categorical constructions that are based on the concepts of this section schematize perceptually based processes. In the next section we concentrate on a schema for the generation of high level representations over the raw perceptions as just described.

Self Generation of a Representation

Organization

Given a basic perception $\mathcal{P} = \langle \mathcal{E}, \mathcal{I}, \varrho \rangle$ as defined above, an intelligent agent could be expected to *organize* the elements of the set $\mathcal{I}$ into a meaningful representation. An organized representation would typically relate the connotations to each other, form compound concepts, and provide infrastructure for access and inference. Loosely, the intuition is that if the details of the organization are bootstrapped from the basic framework $\mathcal{P} = \langle \mathcal{E}, \mathcal{I}, \varrho \rangle$ itself, then the organization may be considered *grounded*.

It is proposed to let the agent organize its connotations in a complemented distributive lattice, namely a Boolean algebra (Sikorski 1964). Quite a few features of Boolean lattices may serve representational purposes and related procedural objectives:

- They feature a partial order. This may enable the organization of connotations in *taxonomic hierarchies*, with inheritance of information.
- They feature the two binary operations $\vee$ and $\wedge$, and the unary operation $\neg$, allowing the formation of *compound concepts* as combinations of more basic concepts.
- The lattice aspect of Boolean algebras provides links for *ease of access*.
- The propositional aspect of Boolean algebras, where $\wedge$ stands for 'and', $\vee$ stands for 'or', and $\neg$ stands for 'not' may underlie an intepretation of the representation in logical formulas, and be applied for *ease of inference*.

The basic categorical infrastructure from the former section allows the application of well developed mathematical tools to transform the set $\mathcal{I}$ into a Boolean lattice. An outline of the constructs and results from (Arzi-Gonczarowski & Lehmann 1998a) follows.

1. As a basis for further constructions, the simplest and most general way to close $\mathcal{I}$ under Boolean operations is to take the connotations in $\mathcal{I}$ as *free generators* of a Boolean algebra over $\mathcal{I}$, designated ${\mathcal{B}_\mathcal{I}}^{\text{free}}$.
2. The cognitive transition to a perception where the set of connotations is ${\mathcal{B}_\mathcal{I}}^{\text{free}}$, is captured by a p-morphism:

$$\xi^{\text{free}} : \langle \mathcal{E}, \mathcal{I}, \varrho \rangle \to \langle \mathcal{E}, {\mathcal{B}_\mathcal{I}}^{\text{free}}, \varrho^{\text{free}} \rangle$$

   The mapping of w-elements $\xi^{\text{free}} : \mathcal{E} \to \mathcal{E}$ is the identity (i.e. this transition manipulates only the internal apparatus). The mapping of connotations $\xi^{\text{free}} : \mathcal{I} \to {\mathcal{B}_\mathcal{I}}^{\text{free}}$ is simply the inclusion of generators, namely raw sensory perception is naturally extended into a structured closure. This introduces the features anticipated before, like combined concepts such as (for generating connotations α and β) $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$, and $(\neg\alpha)$, appropriate lattice links between them, etc.
3. The p-predicate ϱ^{free} on $\mathcal{E} \times {\mathcal{B}_\mathcal{I}}^{\text{free}}$, calls for adequate dependence between the p-predicate values of different connotations that are dependent in the Boolean sense. For example, if a w-element w is perceived as having each one of the two connotations α, β then perceiving w as *not* having $(\alpha \wedge \beta)$ 'does not make sense'. Technically, an adequate definition of a 3-valued predicate on a Boolean aglebra is provided. The context free construct systematizes the expected features. For example, $(\alpha \wedge \beta)$ inherits the information from both α and β. Reasoning inferences are enabled, for example 'if w has the connotations $(\alpha \vee \beta)$ and $(\neg\alpha)$, then it must have the connotation β'.
4. The *subcategory of Boolean perceptions (b-perceptions)*, $\mathcal{P}rc^{\text{bl}-\mathcal{I}}$, is defined to comprise all perceptions where the sets of connotations are Boolean algebras, and the p-predicates are restricted accordingly.

The categorical construction of the p-predicate provides truth tables, that may be applied algorithmically for the computation of the Boolean p-predicate (b-p-predicate). For the generators in $\mathcal{I}$, the values of ϱ^{free} are provided (and hence grounded) by ϱ from $\mathcal{P} = \langle \mathcal{E}, \mathcal{I}, \varrho \rangle$. Perception of their Boolean combinations (that are the added elements of ${\mathcal{B}_\mathcal{I}}^{\text{free}}$) is computed on the basis of the truth tables, and it is also grounded, because the computation 'bottoms out' at the generating connotations. For example, the value of $\varrho^{\text{free}}(w, \alpha \wedge \beta)$ is computed from (grounded by) the raw sensory values $\varrho(w, \alpha)$ and $\varrho(w, \beta)$.

In categorical terminology, a *free functor* from the category of perceptions into the subcategory of b-perceptions has just been described:

$$\mathcal{G}^{\text{free}} : \mathcal{P}rc \to \mathcal{P}rc^{\text{bl}-\mathcal{I}}, \langle \mathcal{E}, \mathcal{I}, \varrho \rangle \mapsto \langle \mathcal{E}, {\mathcal{B}_\mathcal{I}}^{\text{free}}, \varrho^{\text{free}} \rangle$$

and the p-morphism $\xi^{\text{free}} : \mathcal{P} \to \mathcal{G}^{\text{free}}(\mathcal{P})$, mentioned earlier, is a *natural transformation*. The categorical concepts *free functor, natural transformation* warrant precise and testable properties that may be applied to serve representational objectives. For example, a commutative diagram (fig. 1) provides basis for a rigorous computational handling of perceptual change. Categorically, the process preserves morphisms, hence for every modification in basic perception the schema prescribes how to adjust the representation. If there exists a p-morphism (top arrow of fig. 1) between two simple perceptions, describing the passage from one to the other, then it can be methodically extended to prescribe a

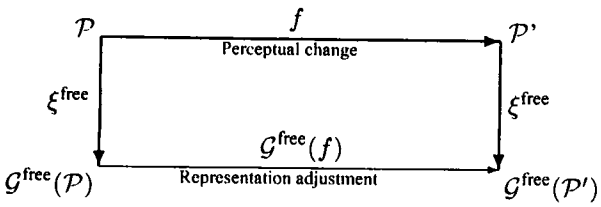

Figure 1: A Commutative Diagram: Generation, modification, adjustment

passage from the representation generated over one to the representation generated over the other. The extended p-morphism (bottom arrow of fig. 1) is based on a Boolean homomorphism, featuring a structural alignment of the organized representations

By (Barr & Wells 1995, p.83), commutative diagrams are the categorist's way of expressing equations, and the commutative diagram in fig. 1 indeed grasps the above in a precise form. From the computational point of view, one may choose to follow either one of two arrow routes:

- Right then down: generate a new representation that is already based on the modified perception.
- Down then right: perform the transition directly on the representation that was generated over the former perception.

The diagram warrants that the results of both routes would be the same.

Remark: 'Perceptual change', in fig. 1, is any p-morphism, allowing for dynamic changes both in $\mathcal{I}$ and in $\mathcal{E}$, although the construction in (Arzi-Gonczarowski & Lehmann 1998a) deals only with changes in $\mathcal{I}$. The formalism can be readily extended.

More Grounding

The free representation generation schema from the last subsection is grounded in the sense that the values of the b-p-predicate are computed from the raw sensory perception that generated it. However, the other aspects of the representation are guided by more abstract speculations. The lattice partial order which organizes connotations in a taxonomic hierarchy (e.g. $\alpha \wedge \beta$ subsumes α) is not grounded in perception but in logical symbol manipulation. Likewise, the formation of compound concepts as combinations of other concepts 'makes sense', but there is nothing directly perceptive about it. Our pretheoretical intuition is that a representation should marry logical sensibility with sensitivity to the perceived situation. This could contribute significantly to an intelligent representation and, consequently, to the performance of the agent.

The formal tool that is proposed for capturing observation of lawlike patterns is a *subsumption relation* in $\mathcal{I}$:

A connotation α *subsumes* another connotation β, designated $\alpha \trianglelefteq \beta$, if, for all w-elements $w \in \mathcal{E}$, $\varrho(w,\alpha) = t$ implies that $\varrho(w,\beta) = t$, and $\varrho(w,\beta) = f$ implies that $\varrho(w,\alpha) = f$.

Synonyms subsume one another, and hence for all w, $\varrho(w,\alpha) = \varrho(w,\beta)$, and this is designated $\alpha \simeq \beta$.

Synonymity is an equivalence relation. Subsumption is a quasi ordering on $\mathcal{I}$, and a partial ordering on the equivalence classes of synonymity.

These relations can also be defined in the Boolean closure $\mathcal{B}_{\mathcal{I}}{}^{\text{free}}$, *between Boolean combinations of connotations*, opening the possibility for an integration of abstract logical speculations with direct observations.

For example, assume that, in a given situation, all objects w that show *yellow_lights* are also *approaching*, while all those that show *red_lights* are also *receding*. This may translate to lawlike patterns like *red_lights*$\trianglelefteq$*receding*, and *yellow_lights*$\trianglelefteq$*approaching* (provided, of course, that the underlying apparatus can phenomenally make these discriminations). It might be useful to embed the observations in the representation by introducing the subsumptions into the lattice structure, namely to place the node *red_lights* below *receding*, etc.

Synonymity and subsumption are not meant to express universal facts. They are specifically grounded in situations and perceptions. It is desirable that an observant agent should be able to incorporate them, in a methodical manner, into its representation. The mathematical categorical infrastructure provides tools to model that. A representation that is freely generated as in the former subsection is *valid* in the sense that its lattice links hold perceptually: if α subsumes β *in the lattice* then the perceptual subsumption $\alpha \trianglelefteq \beta$ holds as well. In other words, the abstract speculations from the former subsection are safely grounded in (verified by) perception. However, in an even better grounded representation specific *observed patterns* should also be reflected in the lattice. The generation of the representation is *complete* if *all* of them are so reflected. Formally, for any two connotations α and β, if, *perceptually*, $\alpha \trianglelefteq \beta$, then α should subsume β in the Boolean algebra as well.

The second canonical construction, the one that will be sketched now, differs from the construction of the former subsection in perceptual acuity as just explained. In addition to sensible, yet abstract, logical speculations, its lattice of connotations, $\mathcal{B}_{\mathcal{I}}{}^{\text{cmp}}$, delineates the full, *observed*, subsumption relation that consists of perceived lawlike patterns[2]. This generation is technically captured by another free functor, which takes the free closure $\mathcal{B}_{\mathcal{I}}{}^{\text{free}}$ and 'moves things around' to refine the representation. Connotations that were originally perceived independent of one another in the free representation may now be dependent and placed 'one below the other', and synonyms are merged. This introduces more meaning into the representation. The main steps of the construction from (Arzi-Gonczarowski & Lehmann 1998a) are, roughly:

1. The basis for the construct is the free Boolean algebra $\mathcal{B}_{\mathcal{I}}{}^{\text{free}}$ that was described before.
2. A delineation of patterns is effected by taking this Boolean algebra modulu a suitable ideal (loosely: the ideal generated by elements of the form $\alpha \wedge \neg\beta$ whenever

[2]The exhaustive article explains how one may also choose to represent only selected lawlike patterns.

it is observed that $\alpha \trianglelefteq \beta$), yielding a more constrained Boolean algebra, ${\mathcal{B}_{\mathcal{I}}}^{\text{cmp}}$, where the generating connotations still ground the representation, but not necessarily as free generators. There may now be Boolean dependencies among them, capturing internalization and representation of grounded observations. The transition is captured by a p-morphism:

$$\xi^{\text{cmp}} : \langle \mathcal{E}, \mathcal{I}, \varrho \rangle \rightarrow \langle \mathcal{E}, {\mathcal{B}_{\mathcal{I}}}^{\text{cmp}}, \varrho^{\text{cmp}} \rangle$$

with the inclusion of generators $\xi^{\text{cmp}} : \mathcal{I} \rightarrow {\mathcal{B}_{\mathcal{I}}}^{\text{cmp}}$.

3. The b-p-predicate ϱ^{cmp} for the modified closure is computed, as for all b-perceptions, according to the 3-valued truth tables mentioned earlier, and the computation eventually 'bottoms out' at the generating, grounding, connotations.

4. In categorical terminology, a *free functor* is defined from the *subcategory* $\mathcal{P}rc^{\text{strct}-\mathcal{I}}$ of perceptions into the subcategory of *complete* b-perceptions:

$$\mathcal{G}^{\text{cmp}} : \mathcal{P}rc^{\text{strct}-\mathcal{I}} \rightarrow \mathcal{P}rc^{\text{bl}-\mathcal{I}}, \langle \mathcal{E}, \mathcal{I}, \varrho \rangle \mapsto \langle \mathcal{E}, {\mathcal{B}_{\mathcal{I}}}^{\text{cmp}}, \varrho^{\text{cmp}} \rangle$$

and the p-morphism, $\xi^{\text{cmp}} : \mathcal{P} \rightarrow \mathcal{G}^{\text{cmp}}(\mathcal{P})$, mentioned earlier, is a *natural transformation*.

The categorical concepts provide, as before, a precise and testable form of stating the computational properties of the process. A commutative diagram similar to the one in fig. 1 (with 'complete' replacing 'free'), provides an equation about the interrelation between perceptual change and the corresponding adjustment of the representation.

Two different representation generations have been formalized. A fixed point formalism (Arzi-Gonczarowski 2000a) shows that the latter is 'most grounded' in observant perception. However, there is a tradeoff price for the refinement. Perceptive readers may have noticed that the second functor is not defined on the entire category $\mathcal{P}rc$ but rather on some subcategory, $\mathcal{P}rc^{\text{strct}-\mathcal{I}}$. This is the key to the difference between the two constructs, that will be explained now. Both constructs ensure preservation of transitions as shown in the commutative diagram, and the arrow between the generated representations features a structural alignment. However, recall that in the second construct, the representation is organized not only by general logical considerations, but also by observed lawlike patterns that are situation specific. It follows that *two representations cannot be aligned if they feature conflicting lawlike patterns*. Technically, the second functor has to be defined on a subcategory of $\mathcal{P}rc$ that consists of *all basic perceptions, but arrows are restricted to "$\trianglelefteq$ preservers"*. In other words, every perception can generate its most acutely grounded representation, at the price of sanctioning transitions to other perceptions that feature conflicting patterns. Following a former example, if the agent perceives a w-element w with the connotation *red_lights*, that is NOT *receding*, then the representation has to be reorganized, breaking some links in the lattice. This may not be such a bad thing after all, as one gets a means of distinction between 'minor' modifications and modifications that require some reorganization of the representation. Technically, if the transition can not be captured by a 'straight forward' p-morphism, then a categorical pullback construct (that consists of a structure of p-morphisms) should be able to capture the similarities between the two representations, 'erase' the broken links, and preserve the others. This was studied in (Arzi-Gonczarowski 1999b). It turns out that the formalism captures a certain natural notion of 'mental effort': substantial modifications require heavier tools to adjust the representation.

Sensible representations should avoid over complicated descriptions (recall *Ockham's razor* principle). Another feature of the more structured representation generation is that the merging of synonyms replaces complex concepts with simpler ones, wherever possible. If an agent observes, for example, the synonyms:

$$\textit{move_fast} \wedge \textit{red_rear_lights} \wedge \textit{yellow_front_lights} \simeq \textit{resound_crrr}$$

Namely, objects w that move fast and have red rear lights and yellow front lights are exactly those objects that feature a specific recognizable sound. The basic connotation *resound_crrr* is merged with the conjunction connotation *move_fast* $\wedge$ *red_rear_lights* $\wedge$ *yellow_front_lights*. As a result, perception of the compound concept, that requires computation before it bottoms out at the three ground conjuncts, may be replaced by immediate grounded perception of a single (auditory) connotation. The more grounded representation generation thus captures, among other things, an *internalization* of compound descriptions. Transmission of an example message *w is a 'resound_crrr'* employs the internalization. If the message is misunderstood, there is always the option to paraphrase with a combination of other basic concepts, providing a clarification.

Applying the Representation

Representations should be useful for intelligent activities. The generated representation has been shown to provide a theoretical infrastructure for a variety of cognitive and behavioral mechanisms. For example, an agent may use multiple connotation sets, switch among them, and supplement (or constrain) the weaknesses of one with the strengths of the other by applying various forms of joint perceptions that are formalized by products, coproducts (direct sums), pullbacks and pushouts, as studied in (Arzi-Gonczarowski & Lehmann 1998b). Additional processes that are supported are reasoning, interpretations and communications (Arzi-Gonczarowski & Lehmann 1998a), planning and design (Arzi-Gonczarowski 1999a), analogy making (Arzi-Gonczarowski 1999b), behavior management (Arzi-Gonczarowski 1998) and autonomy control (Arzi-Gonczarowski 2000b). The theoretical integration of all these mechanisms in one 'mind' is described in (Arzi-Gonczarowski 2000a), along with an analysis of the boundaries that the formalism sets on the 'mind'. The current research frontier is to deploy the same mathematical infrastructure to systematize and evaluate requirements for some aspects of self reflection (Arzi-Gonczarowski 2001).

A challenge for the future is to implement the formalism and its applications in a programmed system. Mathemati-

cal constructs usually lend themselves to algorithmic interpretations. Detection of lawlike patterns between Boolean classifications can be based, for example, on a programmed implementation like LAD (Boros *et al.* 1996). Prospects for a neat and flexible implementation are based on the concise formal ontology. Like a reduced instruction set for a computer, it conflates the types of building blocks that are required for a high level architecture: w-elements, connotations, p-predicate, categorical primitives, and Boolean primitives. The spectrum of perceptual cognitive and behavioral processes that are supported seems, however, to be rather broad.

Domain specific identification of relevant sensory information, and its meaningful phenomenal grounding, is to be performed by embodied substitution instances of ISAAC. In the biological context, for example, organisms are initialized, by evolutionary trial and error, to genetic inherited perceptual states that feature essential constituents: their sensory motor neural systems attend to environmental chunks ($\mathcal{E}$) and to discriminations ($\mathcal{I}$) that are vitally consequential to their survival. These constituents vary drastically across environments and embodiments. The abundance of natural species, even in well circumscribed environments, shows that there is typically more than one rudimentary grounded perceptual state that provides basis for coping with a situation. Practically any aspect of a perceived environment can be considered as a bearer of meaning. A partial example list could include animal pheromones and displays, electrical discharges, cultural symbols, postures gestures and face expressions, intangible kinesics and proxemics, vocal nonverbal means of communication such as grunts, cries and gurgles, and, of course, language.

Summary

ISAAC provides infrastructure for a variety of perceputal cognitive mechanisms, in particular for getting from sensory motor data to useful ontological representations. The dynamics of both the environment and its perception are modelled by p-morphisms. The domain general structures that are required are the minimal primitives of the formalism: w-elements, connotations, p-predicate, categorical primitives, and Boolean primitives.

References

Arzi-Gonczarowski, Z., and Lehmann, D. 1998a. From environments to representations–a mathematical theory of artificial perceptions. *Artificial Intelligence* 102(2):187–247.

Arzi-Gonczarowski, Z., and Lehmann, D. 1998b. Introducing the mathematical category of artificial perceptions. *Annals of Mathematics and Artificial Intelligence* 23(3,4):267–298.

Arzi-Gonczarowski, Z. 1998. Wisely non rational – a categorical view of emotional cognitive artificial perceptions. In Cañamero, D., ed., *Papers from the 1998 AAAI Fall Symposium: Emotional and Intelligent: The Tangled Knot of Cognition*, 7–12.

Arzi-Gonczarowski, Z. 1999a. Categorical tools for perceptive design: Formalizing the artificial inner eye. In Gero, J., and Maher, M., eds., *Computational Models of Creative Design IV*. University of Sydney, Australia: Key Centre of Design Computing and Cognition. 321–354.

Arzi-Gonczarowski, Z. 1999b. Perceive this as that - analogies, artificial perception, and category theory. *Annals of Mathematics and Artificial Intelligence* 26(1-4):215–252.

Arzi-Gonczarowski, Z. 2000a. A blueprint for a mind by a categorical commutative diagram. In *Proceedings of the AISB'00 Symposium on How to Design a Functioning Mind*, 10–18. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour, UK.

Arzi-Gonczarowski, Z. 2000b. A categorization of autonomous action tendencies: The mathematics of emotions. In *Cybernetics and Systems 2000*, volume 2, 683–688. Austrian Society for Cybernetic Studies, Vienna.

Arzi-Gonczarowski, Z. 2001. Perceptions that perceive themselves – a mathematical schema. *IJCAS: International Journal of Computing Anticipatory Systems*. Invited paper, forthcoming.

Barr, M., and Wells, C. 1995. *Category Theory for Computing Science*. Prentice Hall.

Billman, D. 1998. Representations. In *A Companion to Cognitive Science*. Blackwell. chapter 51, 649–659.

Boros, E.; Hammer, P.; Ibaraki, T.; Kogan, A.; Mayoraz, E.; and Muchnik, I. 1996. An implementation of logical analysis of data. Rutcor Research Report RRR 22-96, Rutgers University, New Brunswick, N.J.

Oates, T.; Schmill, M.; and Cohen, P. 2000. A method for clustering the experiences of a mobile robot that accords with human judgements. In *Proceedings of AAAI'2000*, 846–851.

Palmer, S. 1978. Fundamental aspects of cognitive representation. In Rosch, E., and Lloyd, B., eds., *Cognition and Categorization*. Erlbaum. 259–303.

Pierce, C. 1931–1958. *Collected Papers*. Harvard University Press.

Searle, J. 1984. *Minds, Brains, and Action*. Harvard University Press.

Sikorski, R. 1964. *Boolean Algebras*. Springer Verlag.