A Blueprint for a mind

Zippora Arzi–Gonczarowski
Typographics Ltd, Jerusalem 96222, Israel

2000 draft

Abstract

The category of artificial perceptions has been conceived as an infrastructure for the mathematical modelling of AI processes, applying a unified rigorous ontology to various intelligent capabilities. It is shown how this theoretical standard provides an account of perceptual cognitive and affective processes, and integrates them in a high level design proposal for a functioning mind. The schema is presented here through examples, and the formal design is viewed from the perspective of pretheoretical intuitions about minds, cognition, and affect. The dialogue between these intuitions on one hand, and context free mathematical structures on the other hand, is the essence of the research, with the distal goal to approximate a general theoretical account, as well as particular models, of 'minds'.

1 Introduction

A unified ontology that is applied to various intelligent capabilities and skills may help combine them in a fully functioning mind. That ontology should be able to capture some essence of mind processes, yet it should avoid over determinism and be general enough for its eclectic purpose. Mathematical category theory typically provides formal tools to capture a structural essence without being over deterministic. Based on these tools, the category of artificial perceptions has been conceived and proposed as an infrastructure for a theory of AI processes, and it is further proposed to design a high level AI architecture on the basis of the ontology provided by that formalism.

The basic objects of the category are snapshots of perceptual states. Each consists of associating between an environment and some internal structure, producing responses, and recording the experience in a tuned perceptual state, which serves as further basis for more processes, thoughts and deliberation. Streams of perceptual states are formed through transitions that are formalized by morphisms (and other categorical constructs). Any one of the elements that make a perceptual state (i.e. the internal structure, the environment, or the responses) could be modified along paths to other perceptual states. A significant family of transitions involves the formation of complex internal structures, such as acutely perceptive mental representations that could layer on top of basic observations. These complex structures provide a bridge for scaling up to higher-level, rational and emotional, capabilities (e.g. reasoning, creative planning, integrated behaviour management, and autonomous regulatory control).

This extended abstract provides a summary of how the schema models the perceptual states themselves, and how it captures various cognitive and affective processes. It is then shown how the unified theoretical standard underlying the various processes enables a rigorous interweaving and integration of all of them in one 'formal mind'. The 'fragments' enhance one another rather than interfere with one another, making a whole that is more than the sum of its parts. The essence of the perceptual-cognitive 'circuits' will be presented by means of examples (the limited length of this paper does not permit a totally self contained digest of all the formal issues, but references are provided to published works).

In mathematical theories, generalizations and principles are typically described by equations. If the concepts and measurement units of several equations match, then they may be embedded in one another, forming an integrated whole. In place of equations, the proposed formalism employs commutative diagrams, that are '*the categorist's way of expressing equations*' (Barr and Wells, 1995, p.83). Like equations, the diagrams can be composed into an integrated compound whole because they are all based on the same categorical premises. The commutative diagram provides a tentative high level 'blueprint'[1] for the eventual programmed design of an artificial 'mind', highlighting the engineering objectives of the formalism. Autonomous action tendencies (urges, emotions) are fomalized as the natural engines of mind vitality: they impel actual performance of transitions between perceptual states. If the diagram provides a 'blueprint of the circuits', then this is the actual 'current'.

The definitions, constructions and results were all operated within the formal mathematical framework, ensuring a tidy treatment that introduces to the related domains tools of mathematical rigor and results that are meticulously stated. On the other hand, the results may be examined relative to the grounding pretheoretical intuitions and existing theories about minds and cognition. After the construction of the diagram, a study of its mathematical properties provides further systematizations of intuitions about the boundaries of minds and intelligence.

2 The Working Example

To illustrate the ideas of the proposed formalism, we take off from a perception of a market stand. It consists of a display of fruits, vegetables, flowers, etc. Depending on its sensory motor neural capabilities, an agent may be able to perceive (some of) the colors, tastes, and odors of the produce, feel their touch, listen to the seller. Depending on its current interests and goals the agent may attend to the price of items, to their nutritional value, to their potential use for gastronomic dishes, to how pretty they are going to look on the dinner table, etc. The perceiver may have been trained or programmed to associate each produce with its botanical classification, or with its country of origin. Various perceptual states and attitudes may consist of different apects from the above.

Distinct perceptual states may pertain to different agents, or to distinct states within the same agent. This calls for mind processes that steer between them. A pretheoretical intuition of this study is that adequate steering between perceptual states is an essence of a functioning mind. Mind processes that

[1] The terminology is borrowed from Magnan and Reyes (1994), who suggest that categorical constructs provide *blueprints* for the design of mind activities.

depend on perception include, among others: urges for direct interaction, that may be either satisfied or not (e.g. get closer, touch, smell, taste), analysis (e.g. "the price of fruit is higher than the price of vegetables"), planning (e.g. plan a nutritious salad, or a novel cross-breeding of fruits.), analogy making (e.g. compare the quality/prices with another stand). Various forms of behaviour (buy, eat, etc.) may follow these processes. Although this toy example is simple, it raises quite a few paradigm issues that a functioning mind needs to tackle. The idea is not original: from the forbidden fruits of the Garden of Eden to sour grapes, our interaction with our natural food has often been a paradigm of other interactions with the world around us.

3 Basic Perceptual Circuits

3.1 Perceptions

The mathematical premises for the proposed formalism have been presented in (Arzi-Gonczarowski and Lehmann, 1998b), and they are briefly summarized now. 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 $\varrho : \mathcal{E} \times \mathcal{I} \to \{t, f, u\}$.

The set $\mathcal{E}$ represents the perceived environment, *world elements* (*w-elements*) that could perhaps be discerned by a perceiving agent. In the example, each separate produce could be one element of $\mathcal{E}$, or maybe every case of produce would be a w-element, or maybe each stand in the market would be a single w-element: even if the environment exists independent of the perceiving agent, its chunking into individual w-elements varies with perception. To communicate between author and reader, a w-element that stands for a certain object will be designated by w_{bla}, for instance w_{apple}, however the perception under consideration may not necessarily relate to it as an apple.

The set $\mathcal{I}$ stands for the discriminations of w-elements, *connotations* that are typically internal to the agent. The market example could feature color connotations (e.g. *red, green* ...), tactile connotations (e.g. *soft, hard* ...), personal taste preferences (e.g. *savoury, unpalatable* ...), nutritional markers (e.g. *vitamin_E, pectin* ...), botanical classifications (e.g. *cucurbitacea, cruciferae* ...), price classes (e.g. *expensive, reasonable* ...) etc. Connotations will be designated in italics. Anything which may be stored and manipulated internally (words, symbols, icons, pictures, diagrams etc.) could be a legitimate connotation.

The 3-valued *Perception Predicate* (*p-predicate*) ϱ relates w-elements with connotations. Let $w_{orange} \in \mathcal{E}$, then a plausible perception $\mathcal{P}$ could have $\varrho(w_{orange}, \mathit{blue}) = f$, $\varrho(w_{orange}, \mathit{vitamin\_C}) = t$, and $\varrho(w_{orange}, \mathit{Jaffa}) = u$ (meaning that, for some reason, perception does not attend to whether this is a Jaffa orange). The p-predicate captures grounded perception. Without going into issues of symbol grounding (Searle, 1984; Harnad, 1990, and others), $\mathcal{P}$ may perhaps associate between lemons and an internal notion of sour taste, namely $\varrho(w_{lemon}, \mathit{sour}) = t$, on the basis of a variety of bottom-up and top-down processes: direct sensation, past experience, the internalization of acquired knowledge, etc.

Actual sets $\mathcal{E}$ and $\mathcal{I}$, and the values of ϱ, once given, provide particular substitution instances. They vary with the embodiment of agents, their envi-

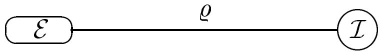

Figure 1: A perception schema

ronment, their history, their goals, etc.

Hence, the $\mathcal{P}$'s stand for high-level perceptual states that happen at, and above, the level of the basic sensory motor neural apparatus and the recognition of cohesive wholes, where conscious cognizance and symbols begin to play a role. The diagrammatic description of a perception will be based on fig.1: An oval designates a set of w-elements, a circle designates a set of connotations, and the connecting line represents some predicative connection ϱ between the two.

Behaviour at the level of this definition was introduced in (Arzi-Gonczarowski, 1998). It consists of reactions that are motivated by perception, providing for agents that can not only passively perceive, but also respond and interact with the environment. In programming terminology, consider, for example, a w-element $w_{apple} \in \mathcal{E}$, a connotation $juicy \in \mathcal{I}$, and assume that $\varrho(w_{apple}, juicy) = t$. The combination of w, α, and $\varrho(w, \alpha)$ could send a message to an object. Methods that can be activated by these messages are the reactions that are associated with perception. A perceived combination of *juicy*, w_{apple}, and $\varrho(w_{apple}, juicy) = t$, for example, could trigger a gland response and/or a motor reaction: 'bite!'. These responses are part of the definition of a perception $\mathcal{P}$. They may consist of whatever mental or physical actions that the agent is capable of performing.

3.2 Transitions

The flow between 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, then a p-morphism $h : \mathcal{P}_1 \to \mathcal{P}_2$ defines the set mappings: $h : \mathcal{E}_1 \to \mathcal{E}_2$, $h : \mathcal{I}_1 \to \mathcal{I}_2$, and *No–Blur* is the structure preservation condition: for all $w \in \mathcal{E}$ and $\alpha \in \mathcal{I}$, whenever $\varrho_1(w, \alpha) \neq u$ then $\varrho_2(h(w), h(\alpha)) = \varrho_1(w, \alpha)$.

Arrows are a fundamental categorical tool that serves us by capturing a broad spectrum of inter- and intra- agent transitions. The mapping of connotations, $h : \mathcal{I}_1 \to \mathcal{I}_2$, could capture $(\imath)$ Simple 'translative' interpretations between perceptions that apply different connotations to the same environment. For example, if $\mathcal{P}_1$ is about pigments, and $\mathcal{P}_2$ is about nutritional substances, then a p-morphism could 'translate' from every pigment to the substance that is most typically associated with it: *h(yellow)* = β*-carotene*, *h(red)* = *lycopene*[2], etc. The meaning of 'most typically associated', for that matter, is as captured by the 'no-blur' structure preservation condition. $(\imath\imath)$ Many-to-one mappings could merge similar connotations, for example *h(foliage)* = *h(leafage)* = *verdure*, *h(seed)* = *h(germ)* = *kernel*, etc. The meaning of 'similar', for that matter, is as captured by the 'no-blur' structure preservation condition. $(\imath\imath\imath)$ The internal representation can be expanded by an inclusive map that is not 'onto'. For example, vitamin connotations could be added to broaden the perspective of perception of produce, capturing the learning of new discriminations. The mapping of w-elements, $h : \mathcal{E}_1 \to \mathcal{E}_2$, could capture $(\imath\imath\imath\imath)$ Simple, literal, analogies

[2]Higher level constructs that will be considered later will enable the mapping of a color also to *a disjunction* of substances, if a pigment is associated with more than one substance.

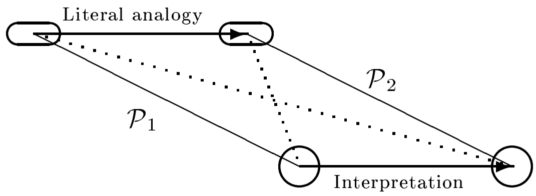

Figure 2: A transition between two perceptions

between perceptions that apply the same connotations to distinct environments. If $w_{nuts1} \in \mathcal{E}_1$ stands for all nuts in stand 1, and $w_{nuts2} \in \mathcal{E}_2$ stands for all nuts in stand 2, than a p-morphism could map $h(w_{nuts1}) = w_{nuts2}$, and the same may be done for bananas, etc. (*ııı*) A p-morphism could merge w-elements into more general environmental chunks: if A is a subset of indistinguishable w-elements (e.g. all of them are flowers), then a p-morphism could map, for all $w \in A$, $h(w) = w_{flower}$, where w_{flower} is a single, generalized, w-element. (*ıııı*) A p-morphism could also expand the perceived environment via an inclusive map that is not 'onto', adding new w-elements to the perception.

The example transitions above are elementary. The idea is that successive basic transitions can be composed into elaborate ones, like a movement of a cartoon character that is made of a series of basic movements of every joint. The 'mind' could perform complex transitions from, say, a perception of the produce in the market to a perception of a nearby exhibit of gems and minerals, creatively soaring from the cherries to a rubi, from the grapes to an emerald, and from the blueberries to a sapphire, making an elaborate interpretive analogy.

The diagrammatic description of p-morphism transitions consists of arrows between sets of w-elements and between sets of connotations as in fig.2. Every such transition can be factorized into an *interpretation*, which consists of the mapping of connotations, and a *literal analogy*, which consists of the mapping of environments. They can be composed in any order. That is why they are shown as parallels in the figure. Whether the interpretation (or the literal analogy) is the first or the second factor effects the *metaphorical perception* that is generated in between. The dotted diagonals in fig.2 designate the metaphorical perceptions that blend connotations from one perception with w-elements from another. This was studied in (Arzi-Gonczarowski, 1999b).

Emotive reactions are part of the definition of perceptions, as was just described, hence perceptual states are also affective states. A transition from $\mathcal{P}_1$ to $\mathcal{P}_2$ may involve a change in some, or all, reactions, featuring a change of mood or attitude. If $\varrho_1(w_{grapes}, sweet) = t$, that could trigger the emotive reaction 'take it', but if a change of perception is based on the map *h(sweet)=sour*, and hence $\varrho_2(w_{grapes}, sour) = t$, then the sour grapes would probably conjure a different reaction.

Technically, composition and the identity are defined by those of set mappings, and perceptions with p-morphisms make a mathematical category, designated $\mathcal{P}rc$. This provides a well developed infrastructure for a mathematical theory. The theoretical standard affords constructs that capture perceptual cognitive transitions in a technically rigorous manner. Examples: (*ı*) 'Blurring' transitions are formalized by traversal of arrows in the reverse direction (in the mirror categroy). This may be applied to cognitive abstraction from details, and to transitions that intentionally 'ignore trifles' that are irrelevant and may interfere with further transitions. For example, the above mentioned transition from w_{cherry} to w_{rubi} needs to first ignore that $\varrho(w_{cherry}, \mathit{edible}) = t$. $(\imath\imath)$ Categorical products and pullback transitions capture joining perceptions into an abstractive schema that highlights the similarities between them and neatly 'blurs' the differences. For example, to cognitively join all the specific stands in the market into an abstractive perception of a 'schematic market stand', one would probably have to ignore the exact layout of displays, that vary from one stand to another. $(\imath\imath\imath)$ Categorical coproducts (direct sums) capture an expansion of several perceptions into a shared and broader perspective (e.g. a market perception that attends to all the possible connotations offered in section 2). Further pushout transitions formalize 'commonsense' meanings that are shared by all perspectives.

A connecting thread of $(\imath - \imath\imath\imath)$ above is that perception is fluid and it changes all the time, continuously deleting, replacing, and adding 'facts' and constituents in an ad hoc manner. For example, the edibility of produce could be a crucial discrimination in one context, and a neglectable detail in the transition described above. However, the pretheoretical intuition is that an over permissive account of mind versatililty could deteriorate to inconsistencies that even a 'flexible mind' would have sanctioned. To be implemented in programmed systems, one needs a methodology that is clear, precise, and testable. It should grasp the evasive invariable aspect of meaning with a loose and flexible, yet durable, harness. Structure preservation, in the form of the 'no-blur' condition on p-morphisms, both forces the artificial mind to take a rigorous 'mental note' of the meanings that are being toyed with, and at the same time the formalism also provides flexible tools to neatly play this game.

4 Higher Level Circuits

4.1 Analysis: Representation Generation

Scaling up from basic direct perception to higher-level habilitations is a significant task of the mind. This includes the generation of perceptive and useful representations for reasoning, creative planning, etc. Boolean constructs are proposed for these purposes.

In the example, assume that sets of connotations are closed under Boolean operations. Quite a few features of complemented and distributive lattices, namely Boolean algebras, seem to be capable of serving knowledge representation purposes and related procedural objectives: $(\imath)$ Boolean lattices feature a partial order. This may enable the organization of connotations in *taxonomic hierarchies*, with inheritance of information. For example, *citrus* would probably lie below *vitamin_C*, meaning that 'if it is a citrus then it has vitamin C, and if it does not have vitamin C then it is not a citrus'. Patterns could sometimes be specific to a perception: for example, *orange* could lie below *local*, or *inexpensive*, in one context, but not in another. (A more formal treatment of subsumption of connotations will be provided in section 7.) $(\imath\imath)$ Boolean lattices feature the two binary operations $\vee$ and $\wedge$, and the unary operation $\neg$, allowing the formation of *compound concepts* as Boolean combinations of basic connotations (e.g. $\mathit{lemon}\vee\mathit{orange}\vee\mathit{grapefruit}=\mathit{citrus}$). $(\imath\imath\imath)$ The lattice aspect of Boolean algebras provides links for *ease of access*: access the connotations

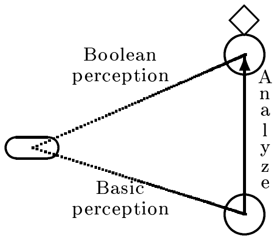

Figure 3: Boolean Representation Generation

through their links to other connotations (e.g. links from *lemon*, to *citrus*, to *vitamin_C*). (*iii*) 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* (e.g. 'if it is a citrus and it is not a lemon then it must be either an orange or a grapefruit').

Analytic organizations of grounded representations were formalized in (Arzi-Gonczarowski and Lehmann, 1998a) by *Boolean generations*, that close sets of connotations under Boolean operations, transforming the $\mathcal{I}$'s into Boolean algebras (with an adequate embedding of the 3 valued p-predicate in these perceptions). P-morphism are then based on Boolean homomorphisms between connotations, capturing structure aligning transitions. Category theoretical natural transformations systematized the transitions into perceptions that feature the Boolean property. The transition is schematized in fig.3, where the Boolean set of connotations is topped with a diamond. The arrow marked *analyze* designates the natural transformation. Two different canonical Boolean closures will be described in Section 7.

The import of the Boolean construct to behaviour is the option to control and regulate conflicts. A complex combination of perceptual constituents may eventually be wired to a complex combination of conflicting reactions. The lattice structure of Boolean closures provides natural junction collocations for the integration of simultaneous action tendencies. For example, consider a w-element $w_{prickly\ pear}$[3], where $\varrho(w_{prickly\ pear}, \mathit{sweet}) = t$ is likely to conjure the reaction 'eat it', while $\varrho(w_{prickly\ pear}, \mathit{thorny}) = t$ is likely to conjure the reaction 'avoid it'. Perception of the conjunction $\varrho(w_{prickly\ pear}, \mathit{sweet} \wedge \mathit{thorny}) = t$ could be wired to integrative regulatory control of the conflict, such as 'hold the appetitive urge, wear gloves, peel carefully, then eat'. Since the Boolean representation may have a lasting existence inside the agent, integrated reactions can also be planned 'off line', in a deliberative manner (e.g. 'what would I do if I was offered a prickly pear'). This is shared with aspects of design processes that will be discussed in the next subsection.

Reactive control could be wired to anything that the agent is capable of doing, and hence also to *the activation of a p-morphism*, capturing an internal transition to a new affective state to perhaps avoid the conflict by a change of attitude. $\varrho(w_{grapes}, \mathit{sweet}) = t$ may conjure the reaction 'take it', whereas $\varrho(w_{grapes}, \mathit{too\ high}) = t$ may inhibit that reaction. Perception of the conjunction $\varrho(w_{grapes}, \mathit{sweet} \wedge \mathit{too\ high}) = t$ could be wired to a transition that maps *h(sweet)=sour*, and hence the conflict is eliminated. In the biological context, the need to deal with conflicting action tendencies could have been a significant

[3]A prickly pear, *sabra*, is the edible fruit of certain species of cacti.

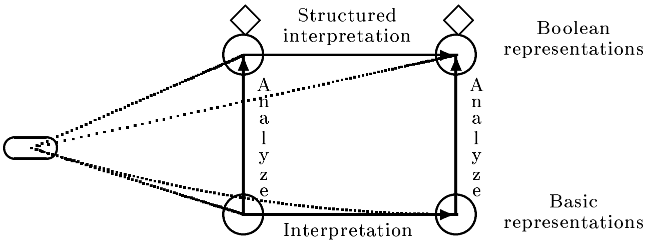

Figure 4: Analytic representations with interpretation

pressure behind the evolution of an interwined net. It is likely that social agents needed to regulate their behaviour and impulses well before mazy internal connections developed into representations underlying what George Boole entitled 'The Laws of Thought'.

The Boolean construct provides infrastructure, but it does not warrant that integration and control of conflicting action tendencies are always achievable. Possible obstacles: ($\imath$) Boolean closures have an exponential computational complexity, and minds may be unable to cope computationally with an overwhelming simultaneous combination of too many rousing stimuli. ($\imath\imath$) Not all conflicts have solutions, and conflicting behaviours can not always be integrated or prioritized. ($\imath\imath\imath$) Even in view of a plausible compromise or a rational decision, some action tendencies have the nature of not lending themselves to regulatory control, perhaps like the tertiary emotions from (Sloman, 2000). Recall, from Genesis 3:6, the original paradigm of all conflicts: *"And when the woman saw that the tree was good for food, and that it was a delight to the eyes, and a tree to be desired to make one wise, she took of its fruit, and did eat...".* It may be formalized as $\varrho(\mathrm{w}_{fruit}, \mathit{desired} \wedge \mathit{forbidden}) = t$. Minds sometimes need to function in spite of action tendencies that have not been consumed. Until their vigilance is somehow calmed, demanding unsatisfied urges *"crouch at the door"*, and may cause a (partial or total) derailing of mind function, attention and control. In spite of that, it would not be a good idea to do without action tendencies, as they are the essence of vitality, the 'current in the circuits'.

Diagramatically, to scale up the affective-cognitive performance of the agent, the 'plane' that is shown in fig.2 is going to serve as a 'base' for a diagram that looks like a 'box'. The generating arrow of fig.3 is the basic 'corner support'. The functorial construction provides an entire 'wall' along with that 'corner support', in the form of a commutative diagram that is shown in fig.4. The commutative diagram is an *equation* that warrants: A transition from the lower left circle to the upper right diamond can be effected in either one of two *equivalent* ways. In the example: let the lower arrow in fig.4 be a p-morphism h that interprets from a 'vitamin minded' perception $\mathcal{P}_1$, to a 'color' perception of the market $\mathcal{P}_2$, mapping from every vitamin to the color of its reachest source in the stand: $h(\mathit{vitamin\_A}) = \mathit{red}$, $h(\mathit{vitamin\_C}) = \mathit{yellow}$, etc. Let ${\mathcal{P}_i}^\diamond, i = 1, 2$, be the 'diamond' perceptions at the top of the diagram wall, featuring analytic representations that are effected by Boolean closures of the $\mathcal{I}_i$'s as explained before. ${\mathcal{P}_1}^\diamond$ represents, among others, the combined connotation $\mathit{multivitamin} = \mathit{vitamin\_A} \wedge \mathit{vitamin\_B} \wedge \ldots$, and ${\mathcal{P}_2}^\diamond$ represents, among others, the combined connotation $\mathit{colorful} = \mathit{red} \wedge \mathit{yellow} \wedge \ldots$. The upper arrow in fig.4 stands for a natural extension of h into a structure preserving interpretation between the two higher level representations, an extension

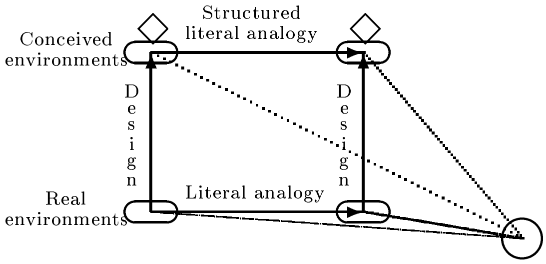

Figure 5: Conceived environments with analogies

that is part of the functorial categorical construction. It captures a structure aligning transition, where the concept *multivitamin* maps to *colorful*, (indeed, Boolean homomorphisms preserve conjunctions). The arrow path that goes first upwards, and then to the right, stands for a transition that first analyzes and represents 'multivitamin', and then follows with a structured interpretation to 'colorful'. The arrow path that goes first to the right, then upwards, stands for a transition that first follows a simple interpretation from vitamins to colors, and then analyzes and represents 'colorful'. This systematizes the interweaving of analytical and interpretive capabilities in one 'mind', where each capability enhances and supports the other. The result that the two optional transition paths are the same is a way of saying that generations of grounded analytic representations are commensurate, because the schema is methodical.

4.2 Synthesis: Design & Plan Generation

A salient property of the premises is the symmetry between $\mathcal{E}$, the environment, and $\mathcal{I}$, the representation. From a purely technical, context free, point of view, the roles that a w-element and a connotation play in the definitions are interchangeable. This *duality* has the technical consequence that any construct or theorem that is established for connotations (w-elements) can automatically be applied to w-elements (connotations), mutatis mutandis. The duality was applied to erect a second wall that faces the wall from fig.4, formalizing creative–imaginative processes. This was studied in (Arzi-Gonczarowski, 1999a). It is summarized in fig.5, which is dual to fig.4, being technically based on mathematical results that were achieved by sweeping the roles of $\mathcal{E}$ and of $\mathcal{I}$. However, the cognitive processes that are formalized here are different. (The structural similarity between fig.4 and fig.5 is more than a technical convenience. It provides insights into similarities of cognitive processes such as reasoning and design.)

In perceptions with *conceived Boolean environments* the sets of w-elements are Boolean algebras, providing a formalism for an adequate internal conception of combinations of similes and examples from the actual environment. (Boolean environments are designated here by an oval topped with a diamond.) This sets a formal basis for the creative imagination of plans and designs. Transitions between perceptions of conceived Boolean environments are based on p-morphisms where the maps between w-elements are Boolean homomorphisms, systematizing structure aligning analogies. Natural transformations formalized methodical cognitive transitions from perceptions of authentic environments to conceived environments. A Boolean combination of w-elements is interpretable as a logical formula that can be further applied for a rigorous effective plan to realize the conceived design.

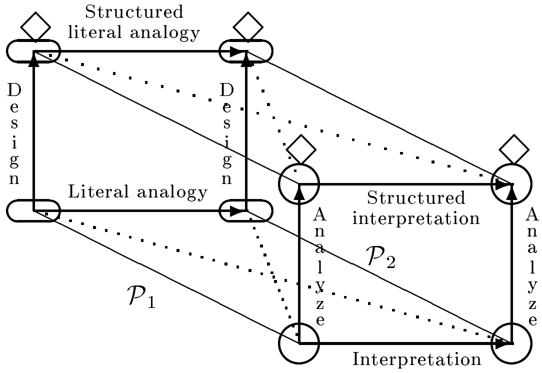

Figure 6: The synthesis

The functorial construction warrants that, if there exists a simple analogy path between two environments, then this path is preserved by the respective Boolean generations, and can be extended to a Boolean structure preserving path between the conceived environments. This is the import of the diagram in fig.5, that interweaves analogy making and creative design in one 'circuit'. A transition from the lower left oval to the upper right oval (with diamond) can be effected in either one of two ways: One could first conceive of a design and then follow with a structure aligning analogy to another design, or, alternatively, one could first follow a simple analogy between existing environments, and then conceive of a design that is already based on the analogical environment. In the example, assume that the perceiving agent conceives of a basket of mixed fruit. The conceived w-element $w_{cherries} \vee w_{grapes} \vee w_{blueberries}$ is an element in a conceived Boolean environment, and its connotations can be perceived with the 'inner eye', on the basis of what is now offered in the market stand, with authentic w-elements that serve as 'raw material' for the plan. Assume now that there is an analogy from the perception of that authentic environment to another environment, say a literal analogy to another stand in the market. The formalism provides computational tools to extend the bottom arrow to the top arrow between conceived plans, for example a transition to an analogous, planned, basket with fruits bought from the other stand, comparing the overall quality and price.

Action tendencies that are conjured by perceptions of conceived environments systematize 'what if' emotions. An agent that perceives an ulterior environment with its inner eye may have emotive reactions to the possibility that the imagined situation could perhaps become real (e.g. excited anticipation, anxiety). An example could be a motivation to actually effectuate the plan and materialize the design: buy the fruits and arrange them in a real basket. The schema for the generation of conceived environments also systematizes the intuition that both the design and its effectuation are easier if an available basket happens to be perceived in the actual environment.

5 The Integrated Circuit

A composite diagram emerges from the fragments: a base with two walls define a box, a whole that features more than the some of its parts. By fig.6, a 'top cover', two 'side walls', and two 'diagonal walls' are gained, representing more perceptions and composite transitions, all of which can be integrated in a single architecture. The category theoretical equational reasoning affirms that the composite box commutes. Various AI cognitive habilitations are interrelated in a wider theoretical framework, with a high-level prescriptive blueprint for an integrated computational framework. Each one of the new walls describes a transition that takes a basic perception ($\mathcal{P}_1$ and $\mathcal{P}_2$, respectively) and scales it up to a cognitive perception with ($\imath$) Analytic mental representation, ($\imath\imath$) A perceptive inner eye that conceives of potential designs and plans, ($\imath\imath\imath$) Integration of behaviours with autonomous regulatory control. The top cover describes an interpretive and analogical transition that applies Boolean homomorphisms to align the high level capabilities ($\imath - \imath\imath\imath$)) that were just described. An example transition of this kind could be based on the interpretive analogy that was mentioned before, between the market display to an exhibit of gems and minerals. Applying the functorial categorical construction, the mind could design, for example, an ornament that would look like a basket of fruits, with rubies for cherries, emeralds for grapes, and sapphires for blueberries. The mathematical construct warrants that one could first conceive of a basket with real fruit, then follow (along the 'top cover' of the box) with a structured interpretive analogy to a conceived ornament or, alternatively, one could first follow a simpler interpretive analogy (along the 'base' of the box) from fruits to gems, and then generate a design on the basis of gems. Further transitions could modify the design by replacement of materials, colors or forms, sometimes to a point where it would not be easy, even to the functioning mind itself, to trace the design back to its original inspiration in the market.

Diagonals and diagonal walls of the diagram have to do with metaphorical perceptions (not all diagonals are shown in the figure). Action tendencies that are conjured by metaphorical perceptions may feature interesting discrepancies: A perceptual state that associates between an environment from one perception and discriminations from another perception, could bring about behaviour that has developed relative to a different, literal, context. It could be quite unexpected in the borrowed context. For example, perceiving an ornament (or reading a research paper) that alludes to fruit may conjure one's appetite, although there is nothing edible there.

6 Inspiriting the circuits

An agent could be initialized to a 'genetic inherited' perceptual state that features essential constituents: it attends to environmental chunks ($\mathcal{E}$) and to discriminations ($\mathcal{I}$) that are vitally consequential to its survival, and its urges and impulses are those that will make it endure. In the biological context these constituents are, of course, naturally selected by evolution. They vary with environments and embodiments. The abundance of natural species, even in specific environments, shows that there is typically more than one rudimentary embodied perceptual state that copes with a situation.

When the initial perceptual state lends itself to contingent transitions, adapts and matures, then perhaps it came with a certain 'mentality'. (Arzi-Gonczarowski, 2000) catalogues the various types of action tendencies that are formalized by the proposed schema, with emphasis on motivations to actually perform the transitions that the blueprint diagram affords. The 'mind' typically functions and develops by interaction with its environment, perceiving and performing the various affective and cognitive transitions that were described above. For example, an agent that perceives how the environment responds to one of its behaviours may be impelled to undergo an internal transition to a modified state that features that behaviour reinforced, or mellowed, according to the perceived response. Different kinds of impact are needed, for instance, for cracking a nut and for peeling a banana. Sensitivity to the properties of materials could be refined through interaction, as well as inter-agent and social sensitivity, and there are, of course, other ripening interactions.

Besides transitions that happen as (either rationally planned, or instinctive) reactions to perceived constituents, some 'mind vitality' could also have its roots in action tendencies that are not related to the agent's relationship with its environment. Perseverant explorer types, for example, are often motivated by persistent drives. A fallout of the formalism is an extension of the spectrum of action tendencies that it systematizes, to behaviour that is driven by internal mental agendas. Internal agendas could be captured as built-in drives towards *attractor states* (although one may never really get to the attractor state)[4]. A formalization of such states is based on 'terminal objects' in mathematical category theory, that will be presented in the next section. Very loosely, an agent with an innate 'curious and interpretive inclination' might have a built-in tendency to move along the arrows of the front wall of the circuit box, invariably analyzing and improving its internal representation. Dually, an agent with an innate 'imaginative designer inclination' might have a built-in tendency to move along the arrows of the back wall of the box, inexorably conceiving and synthesizing novel environments. Subtypes can be formalized by a subtle classifications of arrow routes that are selected.

7 Boundaries of the 'mind'

This section is slightly more technical. It employs basic mathematical tools that are afforded by the formalism to systematize more intuitions about the confines of minds and intelligence. Whether the circuit box is bounded from various directions is the category theoretical version of questions regarding boundary conditions on equations.

7.1 Combinatorial Bounds

In the general case, p-morphisms add new constituents (the exceptions are mergers of similar constituents). Hence, a simple type of bound that may be considered is on the number of different constituents. From the combinatorial point of view, the bound on the number $|\mathcal{I}|$ of connotations are $0 \leq |\mathcal{I}| \leq 2^{|\mathcal{E}|}$ for a given $\mathcal{E}$ (i.e. the possible subsets of w-elements circumscribe the discriminations that one may make). Dually, $0 \leq |\mathcal{E}| \leq 2^{|\mathcal{I}|}$ for a given $\mathcal{I}$ (i.e. the possible subsets of connotations circumscribe the distinct w-elements that one may conceive of). These are obvious bounds along the direction of the arrows.

The category theoretical version of stating that 'one cannot get any further than that' is to show that an object in a category (a perception in $\mathcal{P}rc$)

[4] A similar idea is offered by the dynamical systems stance in cognitive science.

is *terminal.* By definition, a terminal perception T would be such that for all perceptions $\mathcal{P}$, there exists a unique p-morphism $h : \mathcal{P} \to T$. It was shown in (Arzi-Gonczarowski and Lehmann, 1998b) that the *Total Universal Perception of* $\mathcal{E}$, $\mathcal{U}_{\mathcal{E}} = \langle \mathcal{E}, 2^{\mathcal{E}}, \epsilon \rangle$, with $2^{|\mathcal{E}|}$ connotations, has the existence property of (arrows leading to) a terminal object, and this lax[5] terminal object is unique up to isomorphism. This perception has the most evolved representation at the far right of the front wall of the box. Dually, it was shown in (Arzi-Gonczarowski, 1999a) that a similar construct, with the *Universal Environment of* $\mathcal{I}$ that features $2^{|\mathcal{I}|}$ w-elements, has the existence property of (arrows leading to) a terminal object, and this lax terminal object is unique up to isomorphism. This perception has the most evolved conceived environment at the far right of the back wall of the box. These boundary perceptions marry the combinatorial aspect with the categorical algebraic language. They are the *attractor states* from the former section.

The initial object for the category is the *Empty Perception* $\mathcal{P}_{\emptyset} = \langle \emptyset, \emptyset, \varrho_{\emptyset} \rangle$. It stands for 'no environment and no representation', and puts a theoretical bound on the 'origin' of the arrows, from the left and from the bottom of the box (perhaps a theoretical *tabula rasa*).

7.2 A Fixed Point Bound

A stronger result for the top cover bound will be shown now, deploying the strengths of the proposed formalism to systematize more intuitions about intelligence. (The introduction of a fixed point formalism in $\mathcal{P}rc$ is new, although it is a direct result of the constructions from (Arzi-Gonczarowski and Lehmann, 1998a).) Figuratively, the top cover of the box could perhaps serve as a base for another box, and the question is whether it is possible to 'pile up' infinitely many boxes, one on top of the other. This would have meant that a mind could infinitely improve its high level capabilities, constantly adding more compound concepts, more plans and designs, and more integrated behaviours.

The vertical arrows of the diagrams are based on perception endofunctors of the form $\mathcal{G} : \mathcal{P}rc \to \mathcal{P}rc$, where $\mathcal{G}(\mathcal{P})$ is a Boolean perception. A vertical arrow $\xi : \mathcal{P} \to \mathcal{G}(\mathcal{P})$ is a natural transformation from the identity functor on $\mathcal{P}rc$ to the functor $\mathcal{G}$. By definition of fixed points for functors[6], a fixed point of $\mathcal{G}$ should be a pair $(\mathcal{P}, h)$ where $\mathcal{P}$ is a perception and $h : \mathcal{G}(\mathcal{P}) \to \mathcal{P}$ is a p-isomorphism. Figuratively, if $(\mathcal{P}, h)$ is a fixed point of $\mathcal{G}$, then $\mathcal{G}(\mathcal{P})$ is the same as $\mathcal{P}$, making 'a wall of no height': the piling up of walls is stopped. This would mean that ($\imath$) The cognitive transition that is systematized by $\mathcal{G}$ is unable to further scale up perception beyond that which is already featured by $\mathcal{G}(\mathcal{P})$. ($\imath\imath$) $\mathcal{G}$ is a sensible cognitive process that knows its limitations and is 'aware' of property ($\imath$).

Two canonical Boolean closures were studied in (Arzi-Gonczarowski and Lehmann, 1998a; Arzi-Gonczarowski, 1999a). Only one features a fixed point. The difference between them is related to validity and completeness in Boolean perceptions. These notions are based on relationships between the Boolean partial order $\leq$ on constituents (connotations, w-elements) on one hand, and perceived lawlike patterns on the other hand. Examples of perceived lawlike

[5]The uniqueness property of (arrows leading to) this perception does not hold in the 3-valued context.

[6](See, for example, Barr and Wells, 1995, p.272)).

patterns could be: 'Inexpensive produce is either seasonal or local', or 'Onions and shallots are the same'. Formally, the perceptual quasi order $\trianglelefteq$ is defined: ($\imath$) For $\alpha, \beta \in \mathcal{I}$, $\alpha \trianglelefteq \beta$ if $\forall w \in \mathcal{E} \;\; \varrho(w, \alpha) = t \Rightarrow \varrho(w, \beta) = t$ and also $\varrho(w, \beta) = f \Rightarrow \varrho(w, \alpha) = f$. ($\imath\imath$) For $x, y \in \mathcal{E}$, $x \trianglelefteq y$ is defined in a dual manner. Example lawlike patterns of (Boolean combinations of) constituents:
$\neg expensive \trianglelefteq seasonal \vee local$,
$w_{onion} \trianglelefteq w_{shallot}$ and $w_{shallot} \trianglelefteq w_{onion}$.

As already explained in section 4, since Boolean lattices feature a partial order, this enables the organization of connotations in hierarchies. In a *valid* Boolean perception $\leq \subseteq \trianglelefteq$, meaning that the formal Boolean hierarchy can be verified by perceptual observations. In a *complete* Boolean perception $\trianglelefteq \subseteq \leq$, meaning that all observed lawlike patterns are reflected in the Boolean structure. Boolean perceptions are always valid, but not necessarily complete. Perceptions in the *valid and complete* Boolean subcategory, $\mathcal{P}rc^{\text{bl-cmp}}$, feature total internalization of perceived lawlike patterns[7].

The simplest Boolean closure takes the constituents of basic perception as free generators, defines a free functor $\mathcal{G}^{\text{fr}} : \mathcal{P}rc \to \mathcal{P}rc^{\text{bl}}$, and systematizes a general cognitive transition from basic perceptions to Boolean perceptions. It captures methodicalness and open-mindedness, but not perceptual acuity, because ($\imath$) $\mathcal{G}^{\text{fr}}(\mathcal{P})$ is, in the general case, incomplete (freedom means that there is no dependence between constituents, which is the essence of lawlike patterns). ($\imath\imath$) $\mathcal{G}^{\text{fr}}$ has no fixed point. In particular, $\mathcal{G}^{\text{fr}}$ is unable to 'sense' a case where $\mathcal{P}$ is already a Boolean perception, and it unconditionally generates a Boolean set of 2^{2^n} constituents over any given n constituents. (A combinatorial explosion will be avoided when the 'pile' eventually hits the general combinatorial upper bound).

The sketch-structure of perceptions (Arzi-Gonczarowski and Lehmann, 1998a) answers the imperviousness of $\mathcal{G}^{\text{fr}}$. Loosely, a p-morphism in the sketch-structured subcategory, $\mathcal{P}rc^{\text{Sk}}$, preserves lawlike patterns, namely the quasi order $\trianglelefteq$ (the technical details can be found in the cited works). The endofunctor $\mathcal{G}^{\text{fr-cmp}} : \mathcal{P}rc^{\text{Sk}} \to \mathcal{P}rc^{\text{bl-cmp}}$ is a free functor. Loosely, it 'moves things around' in the Boolean lattice to reflect the perceived patterns. Consequently, the transition is perceptually acute: ($\imath$) $\mathcal{G}^{\text{fr-cmp}}(\mathcal{P})$ is valid and complete: it features total observation and internalization of all lawlike patterns that are perceptible by $\mathcal{P}$. ($\imath\imath$) For all valid and complete Boolean perceptions $\mathcal{P}$, $(\mathcal{P}, \xi^{-1})$ is a fixed point of $\mathcal{G}^{\text{fr-cmp}}$. This is a sensible cognitive process that knows its limitations, it is 'aware' of property ($\imath$), and would not modify perceptions that it is unable to amend.

The fixed point formalism tells us that $\mathcal{G}^{\text{fr-cmp}}$ is superior to $\mathcal{G}^{\text{fr}}$, not only because it is more perceptually acute, but also because it has an 'awareness' that avoids the 'unnecessary piling up of boxes'. This bound is cognitively derived from within, on the basis of own observations and own intelligence. This is different from the 'bureaucratic' combinatorial bound that has nothing to do with innate perceptual capabilities. Familiar intuitions that have just been systematized are ($\imath$) Abstract speculations are not enough for real knowledge. A perceptive agent should acutely relate to its environment to construct a truly intelligent knowledge representation. ($\imath\imath$) Sensible cognitive processes should be

[7] Detection of lawlike patterns can be based on a programmed implementation like LAD (Boros et al., 1996).

aware of their limitations.

Based on the observation that the category of $\mathcal{G}^{\text{fr-cmp}}$–algebras is, in particular, a generalized poset, one gets a hierarchy of valid and complete Boolean perceptions as fixed points $(\mathcal{P}, \xi^{-1})$ of $\mathcal{G}^{\text{fr-cmp}}$. (Figuratively: a sequence of bounded walls of ascending size.) This systematizes the intuition that among perceptions with equally advanced Boolean capabilities (namely $\mathcal{G}^{\text{fr-cmp}}$), those with the more detailed grounding apparatus will generate better cognizance. The initial, empty, perception makes the *least fixed point* (a zero size wall). This captures the intuition that even with the best speculative mind, no true cognition can emerge if there is no grounding apparatus that interacts with an authentic environment. Cognition is both enabled and circumscribed by perception.

Remark: An affective fallout of the acutely perceptive Boolean structures that observe and internalize all lawlike patterns, as described above, is a certain gain in introspection. When two observed constituents subsume one another, then they are merged by the structure. For example, if both *half full* $\trianglelefteq$ *half empty* and *half empty* $\trianglelefteq$ *half full*, then a valid and complete Boolean perception would merge the two connotations *half full* and *half empty* into one connotation. Assume now that the generating perception (perhaps that of the owner of the market stand) features a 'positive' emotive reaction when it perceives a half empty case (i.e. half of the merchandise has already been sold), but it features a 'negative' emotive reaction when it perceives a half full case (i.e. half of the merchandise has not been sold). The Boolean representation hence features the cognitive acknowledgement of a self contradicting emotion which is recorded in the representation. The 'mixed feeling' could perhaps be mellowed by a wiring to regulatory control.

8 Summary and Future Work

Lawvere (1994) wrote that categorical constructs approximate a particular model of the general which should be sufficient as a foundation for a general account of all particulars. Hence the schema can be evaluated if it eventually provides infrastructure for the approximation of particular models of minds, on the basis of the general account that has been proposed so far.

The schema still waits to be implemented in a programmed system. Like a reduced instruction set for a computer, the formal ontology conflates the types of building blocks that are required for a high level architecture (w-elements, connotations, p-predicate, categorical primitives, Boolean primitives), but not necessarily the spectrum of mind mechanisms that are modelled.

References

Z. Arzi-Gonczarowski. Wisely non rational – a categorical view of emotional cognitive artificial perceptions. In D. Cañamero, editor, *Papers from the 1998 AAAI Fall Symposium: Emotional and Intelligent*, pages 7–12, Orlando, Florida, October 1998.

Z. Arzi-Gonczarowski. Categorical tools for perceptive design: Formalizing the artificial inner eye. In J.S. Gero and M.L. Maher, editors, *Computational*

Models of Creative Design IV, pages 321–354. Key Centre of Design Computing and Cognition, University of Sydney, Australia, 1999a.

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

Z. Arzi-Gonczarowski. A categorization of autonomous action tendencies: The mathematics of emotions. In *Proceedings of the 15th European Meeting on Cybernetics and Systems Research (EMCSR) Symposium - Autonomy Control: Lessons from the Emotional*, Univ. of Vienna, April 2000.

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

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

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

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

S. Harnad. The symbol grounding problem. *Physica*, D 42:335–346, 1990.

F.W. Lawvere. Tools for the advancement of objective logic: Closed categories and toposes. In J. Macnamara and G.E. Reyes, editors, *The Logical Foundations of Cognition*, pages 43–55. Oxford University Press, 1994.

F. Magnan and G.E. Reyes. Category theory as a conceptual tool in the study of cognition. In J. Macnamara and G.E. Reyes, editors, *The Logical Foundations of Cognition*, pages 57–90. Oxford University Press, 1994.

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

A. Sloman. Architectural requirements for human-like agents, both natural and artificial (what sorts of machines can love?). In K. Dautenhahn, editor, *Human Cognition and Social Agent Technology*. John Benjamins Publishing, 2000.