Artificial Intelligence for Engineering Design, Analysis and Manufacturing (2012), *26*, 63–83.
© Cambridge University Press 2012 0890-0604/12 $25.00
doi:10.1017/S0890060411000023

A complexity theory of design intentionality

THEODORE ZAMENOPOULOS

Department of Design, Development, Environment and Materials, Open University, Milton Keynes, United Kingdom

(Received July 26, 2008; Accepted January 7, 2011)

Abstract

The subject of this paper is design intentionality. The paper is concerned with the property of the mind to hold intentional states (its capacity to represent or reflect existing and nonexisting realities) and with the way that these mental states are constructed during design tasks. The aim is to develop a mathematical theory of design intentionality, capturing the structures and processes that characterize an intentional system with the mental ability to address design tasks. The philosophical notion of intentionality is approached methodologically from a complexity theoretic perspective. More specifically, the focus is placed on the mathematical characterization of the organizational complexity of intentional states and the type of phase transitions that occur on the mental states of an intentional system during design tasks. The paper uses category theory in order to build a framework that is able to mathematically capture the meaning of these notions.

Keywords: Category Theory; Complexity; Design Intentionality; Philosophy of Mind; Semantics

1. INTRODUCTION

This paper is concerned with the mental capacity of certain organisms, like humans, to address design tasks. The ability to address design tasks has been approached and explained in many different ways. For example, design has been predominantly approached as a cognitive ability that is associated with a special form of information processing capacity (e.g., Stiny & March, 1981; Akin, 1986; Goel & Pirolli, 1989), or with a special form of logical reasoning (e.g., March, 1976; Coyne, 1988; Takeda et al., 1990; Roozenburg, 1993). In other studies, the ability to design has been approached as a knowledge-level competence that requires certain types of knowledge (Brazier et al., 2001; Smithers, 1996, 1998, 2002) or certain knowledge structures (Yoshikawa, 1981).

Although design is often perceived as a mental skill or capacity, there is to date no theory that explicitly associates our understanding of design ability with theory of mind and the philosophical notion of intentionality. In order to address this gap, the paper focuses on the property of the mind to hold intentional states (i.e., its capacity of representing or reflecting existing and nonexisting realities), and the way that these mental states are constructed during design tasks. Despite the existence of many theories that contribute to understanding design at some level of observation (e.g., information processing level, logical reasoning level, knowledge level), there is no mathematical account able to express how intentional states are constructed during design tasks. This account would be useful as a basis for looking at the relationship between the mental activity of design and the structures and processes that characterize its realization at potentially different levels of observation. For example, although there is some empirical evidence that identifies certain neurological correlates of the ability to design (Goel & Grafman, 2000; Vartanian & Goel, 2005; Alexiou et al., 2009; Gilbert et al., 2010), there is no theoretical and mathematical framework able to support hypotheses on how neurological structures and processes are linked with the mental capacity of constructing intentional states during design tasks. The same difficulty may apply to mapping the relationship between the construction of intentional states during design and the social structures and processes that underlie design activity (Alexiou, 2007; Alexiou & Zamenopoulos, 2008). Thus, overall, the paper aims to develop a mathematical account on the type of structures and processes that are involved in the formation of intentional states during design tasks.

In this paper the philosophical notion of intentionality is approached from a complexity theoretic perspective. In general, complexity science aims to develop a set of methods and theories for understanding the organizational principles that underlie the creation of higher level functions or structures (see, e.g., Haken, 1983; Badii & Politi, 1997; Schuster, 2001; Boccara, 2004). Complexity science strives to identify whether there are certain common principles that govern how components as diverse as atoms, cells, animals, or humans organize themselves, and in doing so lead to the formation of macroscopic

Reprint requests to: Theodore Zamenopoulos, Department of Design, Development, Environment and Materials, Open University, Milton Keynes MK7 6AA, UK. E-mail: t.zamenopoulos@open.ac.uk

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

phenomena like chemical patterns, living structures, cognitive functions, social, or economic constructs. Part of this enquiry is the identification of the organizational structures and processes that lead to intentionality (Atlan, 1998). However, the potential of complexity science in understanding design as a mental capacity has not been explored fully, despite the increasing import of complexity in design studies. Complexity has been perceived and employed as a scientific approach for conducting design research, as a theoretical and methodological toolkit for addressing design tasks, and as a subject/problem of design practice (for a review, see Zamenopoulos & Alexiou, 2005*a*). Yet, complexity science also presents us with an interesting hypothesis about the nature of design: that the (mental) capacity to design can be perceived and studied by looking at the organizational structures and processes that govern the formation of intentional states during design tasks. Currently there are only tenuous links between design intentionality and organizational level notions usually encountered in complexity theory, such as self-organization, bifurcations, coordination, or phase transition. The aim of the paper is to develop a mathematical theory of the organizational complexity that characterizes the intentional states of a design-capable system during design tasks.

The paper is organized as follows. Section 2 discusses the nature of design phenomena and introduces the core premise of this study about design intentionality. Section 3 explicates the scope and meaning of intentionality theoretically, but also in relation to some fundamental mathematical expressions. The mathematical treatment is based on category theory. Section 4 develops some new mathematical structures that aim to capture variations in the complexity of intentional states. Section 5 brings together the proposed mathematical expressions with the theoretical discussion about design intentionality. Section 6 summarizes and discusses the contribution of the proposed mathematical and theoretical framework.

2. THE NATURE OF DESIGN PHENOMENA AND DESIGN INTENTIONALITY

In the most general sense, the phenomenon of design arises with the formation of organisms whose survival depends on their capacity to construct or adapt their own environment. For instance, the capacity of birds to build nests, beavers to dam ponds, or humans to construct hunting tools, can be perceived as primitive examples of design tasks. This capacity may be equally distributed within a society of individuals. For instance, the capacity of ants to build nests or people to create cities or virtual communities in the Internet may be perceived as examples where the ability to design is the product of the distributed activity of individuals. The capacity of an organism to change or adapt its environment may be contrasted to other logically distinct abilities or strategies: for example, the capacity of an organism to adapt itself to environmental changes, or the capacity to migrate to a new environment (Kirsh, 1996). Overall, the phenomenon of design can be understood as the product of an evolutionary pressure that leads to the formation of organisms with the individual or collective capacity to create artifacts and as a result adapt their environments for their own benefit.

Although evolutionary pressure may explain the presence of design abilities in certain organisms, or the formation of species of design artefacts as a product of exosomatic adaptation (e.g., Steadman 1979), evolutionary theory in itself cannot describe what makes certain organisms, such as humans, capable of designing. A typical response to this quest is to assume that design requires cognition; a mind with the capacity to recognize the possibility of alternative environments and generate instantiations that fulfill the properties of these (imagined) environments. Design activity then arises in response to a *problematic situation* where there is a desire, need, or idea to construct a change in a certain environment, but the precise *means* and *ends* of this construction are not given. Although, this premise is commonly held in design research (i.e., see, Archer, 1965; Mitchell, 1990; Smithers, 2002), it can nevertheless take a number of different interpretations.

The above situation is often interpreted in relation to an information processing system (i.e., a cognitive designer) that faces a special type of problem solving task (e.g., Goel & Pirolli, 1989). According to this perspective, a problematic situation arises with a problem statement or task environment that an information processing system needs to address. The hypothesis implies a distinction between an external environment that sets the problem, and an internal environment that represents the task environment (the problem space). The problem space is then a representation of a set of possible states, a set of legal operations, as well as an evaluation function or stopping criteria for the problem solving task (e.g., Ernst & Newell, 1969; Newell & Simon, 1972). The peculiarity of the design task is that the means (i.e., the representation of the problem space and the possible operations over the problem space), as well as the ends (i.e., the evaluation function or the stopping criteria) are not given in the task environment but are part of the design process (Simon, 1973; Goel & Pirolli, 1989, 1992).

Another interpretation postulates that a problematic situation in design arises from the reflective activity of a designer over an objective reality, that is, a professional, social, or operational environment that includes humans, tools, external representations, artifacts, and so forth (Schön, 1983). According to this paradigm, design is not a characteristic of the task environment (i.e., of a problem statement), but a characteristic of the coupling or interplay between a subject (i.e., a designer) and an objective reality (i.e., the environment). A design situation then arises not as a problem, but as a mental state of a "reflective" agent. This view essentially conveys the idea that cognitive functions are not simply the product of an information processing system (an isolated mind); they are instead formed from the coupling between the mind, the body, and its environment. In design research, a number of different approaches have been proposed following this perspective: for example, seeing design as a hermeneutic act (Snodgrass & Coyne, 2006), as a constructive act of a situated cognitive agent (e.g., Gero, 1998), or as a self-organization process that involves the interaction between internal and external representations (Portugali & Casa-

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

kin, 2002). A more abstract approach sees design as a coevolutionary process between a problem and a solution space (e.g., Maher et al., 1996; Dorst & Cross, 2001).

In the following, the premise that design arises in response to a particular situation is approached in relation to the capacity of an organism to have intentionality or intentional states. Before we explain this premise, we briefly introduce the notion of intentionality.

Although the term intentionality has its origins in the writings of medieval Scholastic philosophers, the contemporary "technical" meaning of the term was reintroduced in philosophy of mind by Brentano in 1874 (republished in 1995). According to Brentano, intentionality aims to describe the very essence of the mind: the capacity to represent or reflect objects or states of affairs in the world, either existing or nonexisting. The capacity of a mind to represent an object or state of affairs in the world is intentional in the sense that mental states are "semantically evaluable" and therefore "refer to something," or they are "about something" (Fodor, 1995). In philosophy of mind it is generally assumed that there are two archetypical, logically distinct intentional states: *beliefs* about "what the properties of a certain reality are," and *desires* about "what the properties of a certain reality ought to be." Although it is analytically convenient to distinguish different types of intentional states, intentionality is generally perceived as a holistic or emergent property of the mind that is derived by a network of causally related intentional states; that is, a network of beliefs, desires or intentions.[1] It is worth noting that although intentionality is often approached as a unique characteristic property of the mind, mental states are not always intentional. Mental states such as pain, general anxiety, or elation are not intentional or semantically evaluable because they are not directed to an object or state of affairs in the world.

Based on these terms, the core premise of this study is that design activity arises in response to the appearance of conflicting or incomplete intentional states, which more specifically express inconsistencies between beliefs held about the past, current, and future states of the world, and desires regarding the state of the world. In response to this situation, the *ability to design* requires both the mental capacity to formulate intentional states about a desired reality and the mental capacity to formulate intentional objects that resolve inconsistencies or conflicts between intentional states (e.g., conflicting beliefs and desires). However, in order to explicate this general premise it is important to introduce a common theoretical background on intentionality.

3. THE STRUCTURE OF INTENTIONALITY: A CATEGORY THEORETIC APPROACH

As briefly discussed, beliefs, desires, hopes, or intentions are different types of intentionality with the common property that they are all directed to something. For instance, the belief that "it is raining" or the desire "to stay dry" refer to a specific state of affairs in the world. It is often suggested that intentionality is determined by two components (Searle, 1983): the *representational content* or *intentional object* that describes the properties of an object or state of affairs in the world (e.g., "it is raining" or "stay dry"), and the *attitude* or *psychological mode* that determines the type of intentional state (e.g., belief, desire, or intention). This structure is often denoted as $\boldsymbol{F}(s)$, where $\boldsymbol{F}$ is the attitude and s is the representational content. Alternatively, Fodor (1975, 1987) defines the attitude of an intentional state as a binary relation $\boldsymbol{F}$ between an organism O and its mental states/mental representations s; that is $\boldsymbol{F}(O, s)$. In this notation, s is a token, a syntactical physical entity, associated with an intentional object (e.g., "it is raining").

According to Searle (1983), an attitude $\boldsymbol{F}$ (e.g., beliefs or desires) expresses the underlying assumptions regarding the relative independence of the world from the mind. These assumptions determine a *direction of fit*. More specifically, an attitude $\boldsymbol{F}$ expresses a direction of fit in the sense that any mismatch between the representational content (e.g., "it is raining," or "stay dry") and the world is resolved by either changing the mind in relation to the world, or changing the world in relation to the mind. Beliefs and desires are thus two archetypical intentional states that express two opposite directions of fit. A *mind-to-world* direction of fit assumes that an intentional object represents an independently existing world and therefore any mismatch between representation and the world must be followed by an adaptation of the representational content. A belief is either true or false in the sense that the correspondence between representational content and the world is evaluated against an independently existing world. A *world-to-mind* direction of fit assumes that an intentional object exists independently from the world and any mismatch between the representational content and the world must be followed by an adaptation of the world. A desire cannot be said to be true or false. A desire is satisfied or fulfilled in the sense that the correspondence between representational content and the world is realized when there is a change in the world in relation to the mind. In both cases, the representational content s of an intentional state determines the *conditions of satisfaction* of a certain intentional state. Based on these terms, intentional states are therefore defined as mental states that have certain conditions of satisfaction s with a certain direction of fit $\boldsymbol{F}$ in relation to an object or state of affairs in the world.

The core notion of intentionality refers to the capacity of the mind to hold *representations* of an existing or nonexisting reality. From this perspective, mental states have a *semantic content*. Moreover, intentional states are *semantically evaluable*: beliefs may be true or false, and desires may be satisfied or unsatisfied. However, existing theories of intentionality take different positions about the nature of mental representations, their semantic content, and their realization. Thus, in the following we review various approaches on (a) how mental representations can be understood and modeled, (b) what is the semantic content of mental representations and how it

[1] Note that the term intentionality (the capacity to hold intentional states) should not be confused with the term intention, which is only one type of intentional state (i.e., the intention to do something).

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

can be evaluated, and (c) how is intentionality realized (particularly the physical realization of mental states). These are outstanding questions also in domains like artificial intelligence, cognitive science and linguistics, but the focus here will be on philosophy of mind. In parallel to presenting the various approaches, the paper will aim to introduce some category theoretic notions and expressions that can capture basic aspects of intentionality. This will establish the mathematical language based on which the proposed hypothesis about design intentionality will be formally expressed.

3.1. Mental representations

How mental representations can be understood and modeled? There are three very broad traditions regarding the way the representational content s of an intentional state $\boldsymbol{F}$ is understood and ultimately modeled: as a *linguistic* (or symbolic), as a *dynamical* (or behavioral). and as a *structural* (or topological) entity. Let us briefly consider these traditions before we move on to a category theoretic (algebraic) formalization that to some extent is able to offer a unifying way of considering these different traditions.

3.1.1. Approaches regarding mental representations

The first approach starts with the hypothesis that mental representations can be understood and modeled as *symbolic expressions* that are generated by a mental language (Fodor, 1975; Pylyshyn, 1984). Symbolic expressions of mental states are often referred to as *propositional attitudes*. According to this view, mental representations have a *semantic content* but also a *syntax* that constitutes the domain over which mental processes are applied. The underlying model is that the mind performs computations (i.e., logical operations) over mental representations that transform one mental state to another. These computations are defined as truth preserving operations: that is, operations that transform propositional attitudes (e.g., beliefs) that are true, into propositions that are also true under certain conditions (i.e., interpretations). In this sense, the mind is realized by a computational machine that *evaluates* the truth value of symbolic expressions. In other words, according to this approach, the semantic content of mental representations is specified by the relation between a language and the machine that realizes or evaluates this language.

The second approach starts with the hypothesis that mental representations can be understood and modeled as particular states of a dynamical system (e.g., Van Gelder & Port, 1995). According to this tradition the attractors or bifurcations that characterize the dynamical behavior of an intentional system can be perceived as core entities that exemplify the meaning of mental representations. Mental processes are then modeled by differential equations and mental representations are reduced to their dynamical properties. The very existence of mental representations as a foundation of intentionality has been challenged in many different contexts (e.g., Dennett, 1987; Freeman & Skarda, 1990). For instance, Dennett (1987) argues that intentionality is a behavioral disposition, so an organism has intentionality because someone takes an "intentional stance" that certain behavioral patterns are related to certain attitudes (e.g., beliefs or desires). In this sense, attitudes may exist without representations (for a critique, see Fodor, 1987).

Finally, the third approach starts with the hypothesis that mental representations can be understood and modeled as structures that are distinguished by their topological or geometrical features. One possible interpretation of this tradition can be found in Harnad (1987, 1990) or more explicitly in Gärdenfor (2004). Gärdenfor (2004) posits the hypothesis that mental representations are regions within an *n*-dimensional Euclidian space. The shapes, spatial, and topological relationships of these regions are the main representational tools for expressing properties and concepts. The view of mental representations as topological or geometrical properties is more formally interpreted as a set of operations that preserve these properties invariant. These operations express the representational content of an intentional state. To put it differently, the representational content of a mental state is understood and modeled as a structure preserving operation that captures the invariant properties of an existing or nonexisting reality.

3.1.2. Category theoretic account of mental representations

For the purpose of this paper, the hypothesis that mental representations are structure-preserving operations is more formally interpreted within the context of category theory. Category theory is a type of algebra that studies the behavior of structure preserving transformations or operations between different species of mathematical structures (called categories). One of the interesting features of category theory stems from the fact that the notion of category is a metastructure that can be used to express different types of mathematical structures: symbolic, dynamical, or topological. Although the use of category theory may imply the hypothesis that mental representations are structural properties of an intentional system, it also seems plausible to interpret the more abstract notions of category theoretic constructions as a unifying framework between the different approaches/levels. A detailed exposition of category theoretic concepts, methods and results can be found in Goldblatt (1984), Barr and Wells (1985, 1990), Lambek and Scott (1986), Lawrere and Schanuel (1997), and Mac Lane (1998). One of the first applications of category theory in design and planning can be found in Ho (1982).

The notion of category. As a mathematical entity, a *category* is defined by a graph structure G (i.e., a set of objects and arrows between objects) together with an operation that allows every different path of arrows within the graph G to be composed and as a result create new structures. More formally, a category C is defined as a *graph* structure that consists of a family of *objects* {a, b, c, . . .}, a family of *arrows* $\{f, g, h, \ldots\}$ and two additional operations *identity* and *composition*. For each object *a* the *identity operation* assigns an identity arrow 1_a: a $\rightarrow$ a and for each pair of arrows $<g, f>$

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

that form a path a → b → c the *composition operation* assigns a new arrow $g \circ f$: a → c that is called the composite. The composition of arrows must satisfy two general axioms. The first (axiom of *associativity*) posits that the composition of arrows is associative, namely, that $h \circ (g \circ f) = (h \circ g) \circ f$ and the second (axiom of *identity*) posits that for every arrow f: a → b, $f \circ 1_a = f$ and $1_b \circ f = f$.

Because of the underlying graph structure, the notion of category has been extensively used to model structural properties of a system and their invariant characteristics. However, the categorical structure can be also seen as an algebraic expression of a formal language or theory. A common interpretation is to see the notion of a category as a formal theory or formal language where a set of *formulas* are expressed by objects {a, b, c, . . .}, *deductions* are expressed by arrows {*f*, *g*, *h*, . . .}, and *rules of inference* are expressed by compositions of arrows (e.g., Goldblatt, 1984; Lambek & Scott, 1986). The composition of arrows over an underlying graph structure is essentially the category theoretic way to express and study how the recursive application of rules of inference (i.e., compositions of arrows) generates new formal expressions and structures. For instance, according to this interpretation, a category C may be seen as an algebra of shapes or shape grammar (Stiny, 2006) where shapes are the objects of the category C and the shape rules are its arrows.

More generally, a category is an algebraic expression of a mathematical structure: for example, a formal language, a dynamical system, or a topological structure. In this sense, categories are often used to describe types, or species, of structures and processes. More specifically, it is a common practice to consider the structure of categories as a "metastructure" whose objects are mathematical structures (linguistic, dynamical, or topological structures) and whose arrows are operations that preserve certain structural qualities invariant.

Mental representations as functors. The arrows that preserve certain structural properties invariant are called *functors*, and they are defined as structure-preserving operations between two categories. More specifically, a functor $\boldsymbol{F}$: B → W from a category B to a category W is a graph homomorphism that preserves the identity arrows and the compositions of arrows. The definition suggests that the notion of a functor has two aspects: first, the functor preserves properties of the underlying graph structure of the category B invariant in W, and second, the functor expresses how structures and compositions of arrows in category B can be used in order to model structural aspects of the category W.

Following this definition and observations, mental representations can be formally expressed in relation to functors. More specifically, a mental representation is a linguistic, dynamical, or geometrical property expressed in category B (e.g., an organism's brain) that remains invariant in a certain world W. The functor $\boldsymbol{F}$:B → W explicates the mathematical meaning and conditions of the invariance of B in W. In this setting, category B can be seen as a mental structure, whereas category W can be seen either as an external perceived reality or as a possible reality that corresponds to another mental structure. As a simple example, we can consider that the category B models the structure of an intentional system that expresses spatial relations. In Figure 1, category B has a graph structure that expresses spatial relations between activities or rooms. In the same figure, category W is the category of shape configurations whose objects are rectangular shapes and whose arrows are adjacency relations between shapes. For the purpose of this paper, category W can be equally thought as an inner (cognitive) representation of a shape or alternatively as an external representation (e.g., a drawing). The functor $\boldsymbol{F}$:B → W is a mathematical notion that explicates the meaning of the structural invariance between spatial or topological relations and shape configurations.

An important aspect of mental representations is that they construct abstractions. This aspect of representation is referred to as *universality*. In complexity science, a *universal construction* is the equivalent of an "order parameter" that enslaves the properties (structure and behaviors) of a family of objects (e.g., Haken, 1983). Similarly in philosophy, a universal construction is a representation of properties that characterize a family of objects. The represented property is said to be a *universal property* and the objects are said to *instantiate* (or *participate in*) the universal property. In set theory, a universal property is captured by the notion of a set; a set is an abstract entity constructed by a collection of objects that have the represented property (they satisfy a membership relation). In contrast to set theory, in category theory the representation of abstract properties is given by concrete objects/structures. Given a family of entities, a universal construction is an archetypical object/structure that exemplifies a property in a universal way. In category theory the notion of universal construction is probably best explained by Ellerman (1988) as a concrete universal: a concrete universal is an archetypical object that represents a property in such a way that all other objects that hold this property resemble or participate in the archetypical object. More formally, universality is identified with the existence of a particular structure in category B. This structure is constructed by a unique arrow *u* in the same way that a set is constructed by a membership relation. The arrow *u* optimally represents the properties b of a family of objects *d* in W as follows:

DEFINITION (universality). Given a functor $\boldsymbol{U}$:W → B and an object b in B, a universal arrow <r, *u*:b → $\boldsymbol{U}$(_)> is an arrow of the form *u*:b → $\boldsymbol{U}$(_) with an object r in W such that the following universal property is satisfied: for

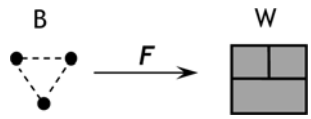
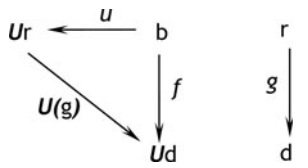


Fig. 1. An illustration of a functor $\boldsymbol{F}$ between topological relations and shape configurations. The notion of functor constitutes the basis of mental representations.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

every d in W and f:b → $\boldsymbol{U}$d in B there is a unique arrow g:r → d such that the following diagram commutes:



■

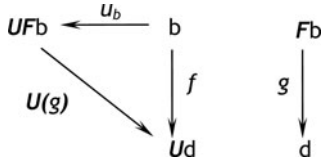
The statement that the above diagram "commutes" essentially means that the equation $f = \boldsymbol{U}(_) \circ u$ is solvable for a unique g (i.e., $f = \boldsymbol{U}(g) \circ u$). The universal construction shaped by the arrow u:b → $\boldsymbol{U}$(r) is the "order parameter" for any equation $f = \boldsymbol{U}(_) \circ u$, in the sense that it determines the properties of structures g:r → d that participate in it. It also means that the arrow u is a "minimum representation" in the sense that any other representation (i.e., any other arrow f in B) can be uniquely constructed using the arrow u. Moreover, the object r is *unique* (or more precisely *unique up to isomorphism*) in the sense that any other object in W that has this universal property is isomorphic to r.

In category theory, the notion of universality leads to the idea of *complementary representations* (or complementary functors), which is mathematically expressed by the notion of *adjunction* between two categories. An example of adjunction is the relation between a computing machine and a language. The concept of machine alludes to an effective procedure that is specified by a language or program fed into the machine, whereas the concept of language alludes to a set of instructions that specify the behavior of the machine. The two constructions are complementary entities. It is a well-known result that this complementary relation can be formally expressed by the notion of adjunction (Goguen, 1973). A more concrete example can be given by looking at the functors between spatial rules (or spatial relations) and shape configurations (Fig. 2). Let us assume that B is the category of topological relations and W is the category of shape configurations. The notion of adjunction suggests the existence of two functors as follows. Let us take a particular graph structure G of topological relations in B and create a set $\boldsymbol{F}$(G) of all possible shape configurations that satisfy the spatial relations expressed by G. This is achieved by a functor $\boldsymbol{F}$:B → W (termed *free functor*). Similarly, let us take a particular shape configuration d in W and create a graph structure $\boldsymbol{U}$(d) that represents the underlying topological relations of d. This is achieved by a functor $\boldsymbol{U}$:W → B (termed *underlying functor*). The notion of adjunction in Figure 2 implies that the category $\boldsymbol{F}$(G) generates a shape configuration d if and only if there is a graph homomorphism from G to $\boldsymbol{U}$(d).

Definition (complementary representations or adjoint functors). Given two functors $\boldsymbol{F}$:B → W and $\boldsymbol{U}$:W → B, an adjunction between B and W is the tuple $<\boldsymbol{F}, \boldsymbol{U}, \varphi>$ where φ is a natural bijection φ:$\mathbf{W}(\boldsymbol{F}\text{b}, \text{d}) \cong \mathbf{B}(\text{b}, \boldsymbol{U}\text{d})$ for *every* b in B and *every* d in W. ■

This definition entails that the notion of complementary representations or adjoint functors gives a universal arrow for every b in B and every d in W. More specifically, given two functors $\boldsymbol{F}$:B → W and $\boldsymbol{U}$:W → B:

- For every f:b → $\boldsymbol{U}$d in B there is a *unique* arrow g:$\boldsymbol{F}$b → d in W such that $f = \boldsymbol{U}(g) \circ u_{\text{b}}$ (i.e., for every b there is a universal arrow u_{b})



- For every g:$\boldsymbol{F}$b → d in W there is a unique arrow f:b → $\boldsymbol{U}$d in B such that $g = \varepsilon_{\text{d}} \circ \boldsymbol{F}(f)$ (i.e., for every d there is a universal arrow ε_{d}).

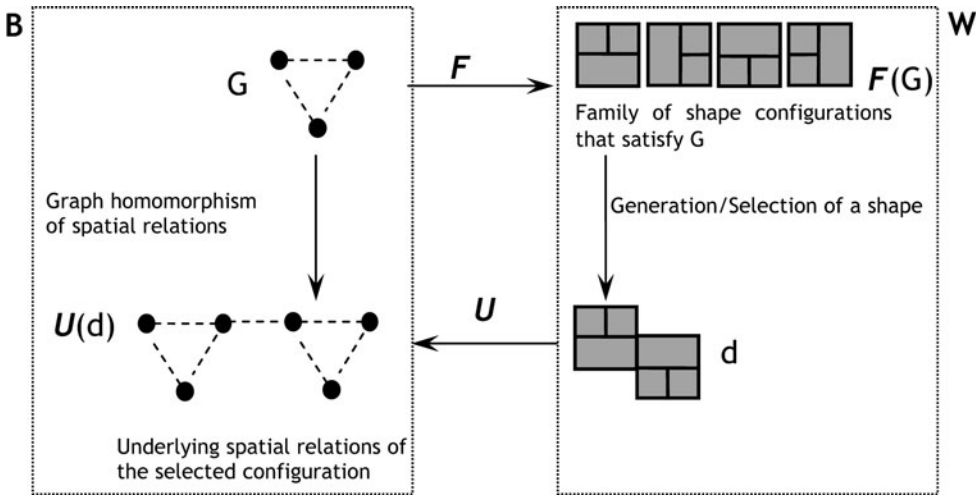
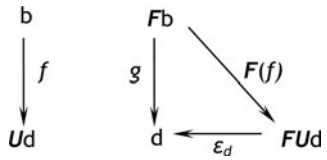


Fig. 2. An illustration of the notion of adjunction. The $\boldsymbol{F}$ functor generates a family of shape configurations that satisfy the topological relations in G. The $\boldsymbol{U}$ functor generates topological relations of shape configurations that can be deduced from G.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press



3.2. Semantic aspects of intentionality

The notion of intentionality and the notion of semantics are inextricably related. As already discussed, the notion of intentionality alludes to mental states that are directed to some existing or possible reality. Intentional states are therefore mental representations that are characterized by their content, their meaning. According to Searle (1983), semantic phenomena must be understood in the context of intentionality and theory of mind. For Fodor (1987), the relation between intentionality and semantics is even more intrinsic: if semantics is an indispensable component of a theory of language, then it should also be an indispensable component of a theory of "inner" language or "language of thought." In this sense, a theory of mind should eventually converge into a theory of language and ultimately semantics. This section aims to lay out some basic issues regarding semantics as they relate to intentionality, in preparation of a more formal treatment in the next section.

3.2.1. Approaches regarding semantics

It is generally agreed that semantics allude to the capacity of an entity, whether a symbolic expression, dynamical, or structural entity in the mind, to hold and manipulate meaning. This is a view that is commonly held in philosophy of mind, linguistics, and logic. However, the identification of the "thing" that we call meaning (i.e., the ontology of meaning), but also the specification of the relation between expression and meaning (the specification of semantic relations) is a notoriously contentious problem.

The predominant view is to perceive semantics as a *referential* relation. According to this view, semantic relations allude to the capacity of an expression, including a natural sign, a linguistic sign, or a mental state, to refer to an object or state of affairs in the world. There are two interrelated strategies for understanding *meaning* (Davis & Gillon, 2004). One strategy is to associate the term meaning with the specification of the properties of the objects or state of affairs in the world to which meaning is attributed. This specification is variously referred to as "sense," "intension," "connotation," "universal property," or "representational content." Another strategy is to associate the term meaning with the very object(s) or state of affair(s) in the world that is attributed to an expression. The set of objects that satisfy the specification is variously referred to as "extension," "denotation," or "reference." The specification of the properties of an object (i.e., the representation content) determines the *truth conditions* or *conditions of satisfaction* that evaluate the relation between a reference and an expression (Fig. 3). Thus, an expression refers to an object or state of affairs in the world because that object or state of affairs in the world has properties that satisfy the expressed truth conditions or conditions of satisfaction. Historically, this tradition has its origins in the work of Frege (e.g., Beaney, 1997) and in the work of Tarski (1956) on model theory.

In philosophy of mind and language these terms are discussed in relation to certain existential and ontological questions. For example, where is the locus of the conditions of satisfaction of an object? Are these conditions expressions of a mental state, or of an external world? Moreover, is the reference an entity that is located in the mind or in the external world?

In response to these questions, a *conceptualist* starts with the hypothesis that meaning is a mental structure. The view that a mental entity is associated with meaning implies that the reference of an expression is located in the brain (e.g., Jackendoff, 1976, 1990; Gärdenfors, 2004) or that the conditions of satisfaction are expressions of a mental state (e.g., Searle, 1983). In contrast, a *realist* rejects the idea that meaning is determined by a mental entity and defines the notion of meaning in relation to an object that is outside the mental state of an agent (e.g., Putnam, 1975). Within this tradition one approach is to postulate that the reference of an expression is determined by certain conditions located in the interactions and structure of the environment of a cognitive agent (Barwise & Perry, 1983). This approach gave rise to the notion of *ecological realism* where the "meaning arises out of recurring relations between situations."

The referential approach to semantics, either realist or conceptual, is not the only way to define semantic relations. The fact that an expression means something, or refers to something, can be perceived as having to do with a purpose or the function of representing something (Dretske, 1995; Millikan, 2004). According to this approach, meaning is a functional entity (for a review, see Harder, 1996). Hence, the *functionalist* or *teleological* view starts with the assumption that the formation of meaning is not derived from the representational content or from a reference in the external

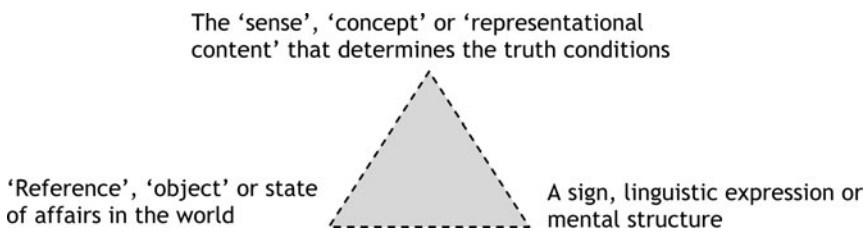


Fig. 3. The semantic triangle. Different theories take different views on the relation and existence of these three basic terms.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

world, but from the function of representing a certain object or state of affairs in the world. More specifically, the meaning of an expression is determined by the intention to produce certain effects and by the purpose and principles of communication (e.g., Grice, 1957). An expression is "false" or "not satisfied" because of its failure to represent or convey a particular meaning.

For the purpose of this study, it is assumed that the functionalist view of meaning does not necessarily reject the existence of referential relations. Semantic relations can be ultimately perceived as referential relations that are selected for their function (Millikan, 2004, p. 67). Therefore, the focus will remain on the referential structure of meaning. Moreover, because the objective of this section is to explicate the logical structure of semantic relations the locus of meaning will be treated both as a mental (subjective) and as an external (objective) entity.

3.2.2. Category theoretic account of the semantic content of intentionality

The formal specification of semantic relations has been part of mathematical logic and in particular *model theory* (Barwise, 1977). Model theory is concerned with the relation between a set of logical statements expressed in a language L and the mathematical structures that satisfy the postulated statements. The set of logical statements form a *theory* and the algebraic structures that satisfy the statements of the theory are called *models*. A theory is consistent if it has at least one model. Assuming the existence of a model, a theory is a set of truth conditions for that model. The distinction between *theories* and *models* is therefore the model theoretic way to explicate the formation of semantic relations as the interplay between "the specification of the properties of a family of objects" and "the set of objects that satisfy the properties of a specification." Note that the terms theory and model are used in a different way than how they are used in science. For an application of model theory to design theory, see Mitchell (1990).

Another tradition to formal semantics is originated in Kripke's semantics and modal logic (Kripke, 1963; Goldblatt, 2006). Modal logic is concerned with formal expressions that assert the mode of truth: for instance, that something is "possibly" or "necessary" true in a given context. Semantics are defined in relation to a relational structure that captures interconnections between *possible worlds* within which an expression is true or false. In this context, a model alludes to relational or algebraic structure that represents the interconnectivity between possible worlds. Other similar forms of formal semantics have been defined by Hintikka (1961) and also Barwise and Perry (1983).

An alternative way to explicate the formation of semantic relations can be found in category theory. Instead of employing symbolic expressions in order to specify semantic content (i.e., the truth conditions or conditions of satisfaction for a certain mental state), category theory uses diagrams of objects, arrows, and compositions of arrows. These diagrams are formally termed *sketches*. Sketches are therefore the category theoretic way of expressing the conditions of satisfaction for a certain intentional state: for instance, sketches may express aspects of a desired artefact such as proximity of rooms in a building layout or strategic knowledge on the appropriate interaction between different design tasks.[2] One of the advantages of the mathematical notion of sketch is that "of being independent of any particular presentation" (Wells, 1994, p. 7). It thus makes it possible to capture the semantic content of intentional states at different levels of analysis: as dynamical, symbolic, or topological structures. The paper will use examples that emphasize the link with symbolic level representations, but the mathematical treatment applies generally to different types of structures.

A sketch generates a mathematical category in the same way that recursive rules or formal grammars generate a language. Based on this analogy, a category can be seen as an algebraic way to express the notion of a formal language. In category theory, the semantic evaluation of sketches is mathematically specified using the notion of functor between categories. A more detailed account of the correspondence between category theoretic formalizations and other traditions in formal semantics can found in Barr and Wells (1985, 1990). For examples and literature on how sketches can capture dynamical processes, symbolic processes or topological transformations see also Lawrere and Schanuel (1997). Let us explore the formal characteristics of sketches in more detail.

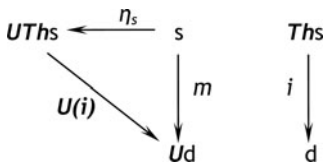
The notion of sketch. A sketch generates a category by specifying some core syntactical aspects of the category theoretic structure. The core aspects of a category theoretical structure include the following: first, a graph structure that depicts structural properties of a category; second, diagrams that specify possible compositions or constraints in the compositions of arrows; and third, diagrams that specify universal properties. More formally, a sketch $s = <G_s, D_s, C_s, Co_s>$ is a graph G_s with a set of diagrams (graph homomorphisms) D_s that determine constraints on the composition of arrows over G_s; as well as a set C_s of cones and a set Co_s of co-cones. A cone, as well as its dual concept of co-cone, is a special type of diagram (graph homomorphism) in G_s that defines the universal properties of the generated category. For instance, a cone may be a diagram that depicts how new objects may be created from the composition (product) of two other objects within a given category. Cones and co-cones are dual concepts: namely, they are diagrams with arrows in opposite directions. For the purpose of this paper, a formal definition of cones (and co-cones) is not necessary but it can be found in any textbook in category theory. For full details about the definition of sketches as a collection of graphs, graph homomorphisms and cones/co-cones the interested reader can look at Barr and Wells (1985, 1990).

[2] Sketches are treated here as mental structures. Borrowing Goel's (1995) book title, sketches are for this paper "sketches of thought." Even though we can easily draw some analogies between the category theoretic notion of sketch with the sketches (drawings) that designers use in their practice, any reference to sketches in this paper is strictly to the formal mathematic entity.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

For each sketch determined by s = <G_s, D_s, C_s, Co_s> it is possible to construct a category that has as an underlying graph the graph G_s; as commutative diagrams the set D_s; and as universal properties the types of cones and co-cones defined by C_s, Co_s. The category generated by a sketch s = <G_s, D_s, C_s, Co_s> is denoted by ***Th***(s) or ***Th***s, and is referred to as a *theory* or *language* generated by the sketch s. The generated theory ***Th***s is a category of entities that satisfy the properties (structure, constraints, and abstractions) specified by a sketch. Equally, each category d has an underlying sketch, denoted by ***U***d = <G_s, D_s, C_s, Co_s>, that specifies these properties. For a sketch s and for a theory of a sketch ***Th***s there is a functor *i*:***Th***s → d that preserves the aforementioned properties of the theory, and as a result generates a possible instantiation of the theory in d. This functor is defined as an *interpretation of the theory* ***Th***s. In the same way, a *model of a sketch* s is defined as a sketch morphism *m*:s → ***U***d from a sketch s to the underlying sketch ***U***d of the category d. A *sketch morphism* is a graph homomorphism that takes the set of diagrams D_s, cones C_s and co-cones Co_s to a set of diagrams, cones and co-cones in ***U***d.

Based on this notation, semantic relations are defined in relation to the following universal property: for every sketch s and every model of a sketch *m*:s → ***U***d there is unique functor *i*:***Th***s → d for which the following diagram commutes (i.e., the two alternative paths of arrows are equal so, $m = U(i)\ \eta_s$):



This property states that there is a natural bijection between models of a sketch *m*:s → ***U***d and interpretations of a theory *i*:***Th***s → d (i.e., an adjunction between a theory functor ***Th*** and the underlying functor ***U***), which is denoted by Mod(s, ***U***Cd) ≅ Int(***Th***s, d).

In sum, the category theoretic concept of sketch offers an alternative way to formalize semantics: the notion of theory or language is expressed by the category ***Th***s, whereas the notion of model is expressed by the special type of arrow *m*:s → ***U***d.

How sketches express the semantic content of intentionality. These category theoretic concepts can be now used to explicate how an intentional state leads to the formation of semantic relations. To do this we make reference to Searle's terms. For Searle (1983), the formation of semantics involves two layers of intentionality: one that corresponds to a "sincerity condition" and one to "meaning intentions." First, there is an intentional state to be expressed by an object (e.g., the desire to stay dry may be expressed by a drawing or sentence); second, there is an intention to express that intentional state. Meaning is therefore acquired because of the intention of the mind to associate an intentional state with a nonmental or physical expression. This association is realized because the intention to express an intentional state holds the same conditions of satisfaction as the expressed intentional state. In category theoretic terms Searle's ideas can be expressed as follows.

First, there is an intentional state expressed in an intentional mind B. This intentional state is explicated by the model ***m***:s → ***U***d and has certain conditions of satisfaction s. This state essentially corresponds to the first layer of Searle's intentionality (Searle, 1983, pp. 160–179). Second, there is an intentional state (in particular, an *intention*) that is realized by the functor ***F***:B → W between two categories B and W. Here, B corresponds to a subjective reality and W to a language that expresses an objective (possible or observed) reality. The functor ***F*** explicates the intention to express something in a language W given certain conditions of satisfaction expressed by the sketch s in B. This essentially corresponds to the second layer of Searle's intentionality (Searle, 1983). A representation of the semantic properties of intentional states is given in Figure 4.

3.3. Realizations of intentionality

This last subsection on intentionality is concerned with the way intentional states as mental entities are related to a physical reality (e.g., the brain). Up to this point, the term intentionality has been discussed as a special mental capacity:

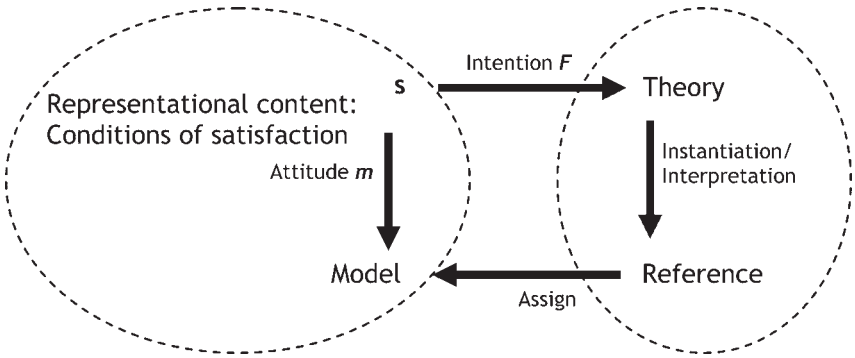


Fig. 4. A schematic representation of the semantic properties of intentional states. The category theoretic notion of sketch s is the representational content (or conditions of satisfaction) of an intentional state. The category generated by a sketch is the inner or outer language (a theory) used in order to express a specific intentional state m.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

a capacity that is inextricably related with the formation of semantic relations. However, what is the locus, body, or universe of intentionality? In other words, how is intentionality realized? Is the brain the only physical realization of an intentional mind? Do social entities have intentionality? In order to respond to such questions, the "problem" of intentionality must be placed in relation to the "mind–body problem."

The mind–body problem concerns the relationship between mental states or processes and physical events. In contemporary philosophy and science, it is commonly held that mental phenomena are somehow linked to physical phenomena, although the exact nature of this relation is not clear. There are four main approaches regarding this relation (e.g., Harman, 1989; Chalmers, 2002).

One approach, *behaviorism*, associates mental states and processes with the behavioral dispositions or behavioral tendencies generated by the underlying physical processes of an organism. According to this view, the mind is a special aspect of the behavior of a physical system (e.g., Ryle, 1949; Dennett, 1987).

The second approach, *identity theory*, associates mental states and processes with physical states and processes. More specifically, identity theory claims that mental states are essentially physical states of the brain of an organism (e.g., Place, 1956; Smart, 1959).

The third approach, *functionalist theory*, attributes mental states or processes to physical states or processes whose behavioral dispositions are distinguished for their functional role. A predominant interpretation of this approach postulates that mental states are functions of a computational machine (e.g., Putnam, 1973). The view of mental states as functional entities of a computational machine helped explain how the same mental states might take multiple physical realizations. Thus, for instance, animals and humans may both have a similar intentional state, but different neurological structures.

Fourth, and finally, a number of alternative studies on the relation between mind and body have brought to the fore the *emergentist* hypothesis that mental states are higher order/emergent properties of lower level (typically physical) states and process (e.g., McLaughlin, 1992, 1997; Horgan, 1993). According to this approach, intentional states are emergent qualities that cannot be deduced from the principles of the components found at a lower level of abstraction. This relation (between higher order emergent properties and lower level states and processes) is specified in different ways: as ontological, logical, or epistemological relation (see Zamenopoulos, 2008). The emergentist view of the mind has been closely linked to the concept of "supervenience," first introduced in the context of philosophy of mind by Davidson (1970). According to this view, the mind supervenes the physical states or processes of a brain: that is, although, the same mental states may have a different physical realization, identical physical states or processes specify identical mental phenomena.

The functionalist and the emergentist approaches in theory of mind both brought to the fore the hypothesis that intentional states can have many different realizations (e.g., Kim, 1992). According to this hypothesis, intentionality can be said to be embodied in certain organizational structures and processes that are not necessarily realized in the physical structure of a brain. The hypothesis suggests the possibility to understand and study the notion of intentionality as a phenomenon that may arise across different levels: biological, cognitive, or social (Atlan, 1974, 1998). For the purpose of this study, the term intentionality or intentional state will be used for any organism or system that has the capacity to hold representations of an objective (existing or nonexisting/possible) reality. In this sense, this distinction between representation and the representing object marks the distinction between a subjective and an objective reality. The ontoloical status of these realities is not specified, but the logical properties of this coupling will be the main focus of the next sections.

4. THE ORGANIZATIONAL COMPLEXITY OF INTENTIONALITY

In Section 2, the phenomenon of design was approached as a mental capacity that arises in relation to a particular intentional state: when there are inconsistencies between beliefs held about the past, current, and future states of the world, and desires regarding the state of the world. Following this idea, Section 3 introduced some well-known category theoretic tools in order to capture the meaning and structure of intentionality. In particular, the notion of adjunction was introduced in order to explicate the semantic content and evaluation of intentional states. However, the introduced treatment precludes the expression of incomplete representations or the expression of conflicting intentional states. For this reason in this section the paper introduces some new mathematical entities in order to capture qualitative variations in the representational properties of intentional states, allowing the formal expression of inconsistencies.

These variations are essentially expressions of the organizational complexity of intentional states. As mentioned in the introduction, complexity science aims to understand the organizational processes and structures that underlie the creation of higher level functions or structures. In the context of this paper, the focus is placed on the organizational processes and structures that underlie the formation of intentional states. More specifically, the objective is to distinguish different qualities of mental representations independently from the actual content and attitude of intentional states. For that purpose, it is necessary to introduce a semantic/intentional theory of organizational complexity: a mathematical framework for expressing semantic aspects of complexity. A more detailed account of the proposed mathematical framework can be found in Zamenopoulos (2008) and Zamenopoulos and Alexiou (2007*a*).

We first briefly review some existing approaches to the development of mathematical descriptions of organizational complexity. One approach is to focus on *ontological properties* of organization, such as, for instance, the degrees of freedom and mutual information in the description of a system (e.g., Nicolis & Prigogine, 1967; Atlan, 1974; Von Foerster, 1984). A second approach is to focus on the *logical or computational properties* of organization, such as the computa-

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

tional resources (the time or logical effort) needed in order to describe/compute a system (e.g., Kolmogorov, 1965; Chaitin, 1966, 1997; Bennett, 1985), or the computability of a system (e.g., Langton, 1990; Wolfram, 1994). A third approach is to focus on the *epistemological aspects* of organization, such as, for instance, the characterization of the number of independent descriptions or models of a system (e.g., Rosen, 1991; Cariani, 1991; Casti, 1992).

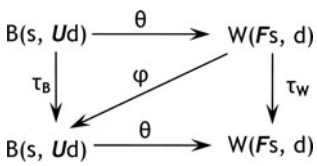
For the purpose of this study, the focus is placed on the *semantic aspects* of organization. The proposition put forward is that variations in the notion of sketch can be used in order to describe the organizational complexity of a system. More precisely, the universal property of semantic relations (that is the adjunction that arises in the semantic content of intentional states) is employed as a tool for the characterization of organizational complexity. According to this perspective, an *ordered* organization is an organization that is described by a *well-formed theory*; a *random* organization is an organization that is characterized by the *absence of a theory*; and *complex* organization is an organization that is described by a notion of a *weak theory*. A weak theory can then be thought as the most general class of theories, which allows us to consider well-formed and random theories as special cases. On that basis, the organizational complexity of an intentional state can be thought to vary according to the type of sketches, theories, and models involved: from *well-ordered intentional states* that generate strong (or well-formed) theories, to *complex intentional states* that generate weak theories.

For the purpose of defining the notion of weak theory it is important to specify the notion of *weak adjunction*. More formally, the concept of weak adjunction between two categories B and W requires the existence of two functors $\boldsymbol{F}$ and $\boldsymbol{U}$ with the following properties.

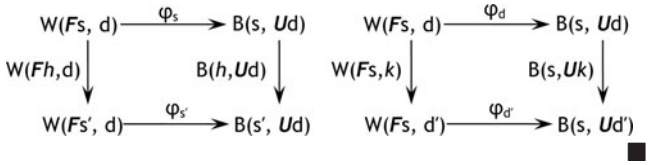
DEFINITION (weak adjunction). A weak adjunction between two categories B and W is defined by a tuple $<\boldsymbol{F}, \boldsymbol{U}, \varphi, \theta, \tau_{\mathbf{B}}, \tau_{\mathbf{W}}>$, where $\boldsymbol{F}$:B → W and $\boldsymbol{U}$:W → B are functors; the arrow $\tau_{\mathbf{B}}$ is a (natural) transformation between any two arrows (functors) m:s → $\boldsymbol{U}$d and m':s → $\boldsymbol{U}$d of the set B(s, $\boldsymbol{U}$d); the arrow $\tau_{\mathbf{W}}$ is a (natural) transformation between any two arrows (functors) i:$\boldsymbol{F}$s → d and i':$\boldsymbol{F}$s → d of the set W($\boldsymbol{F}$s, d); and the arrows φ and θ make the following diagram commute naturally in s and d (i.e., $\varphi \cdot \theta = \tau_{\mathbf{B}}$ and $\theta \cdot \varphi = \tau_{\mathbf{W}}$):



The two diagrams can be equally depicted as follows:



As usual, the condition of naturality for arrow φ (and similarly for θ) means that the arrow φ preserves the categorical structure as s and d vary. More formally, naturality suggests that the following diagrams commute for every arrow h:s′ → s in B and k:d → d′ in W:



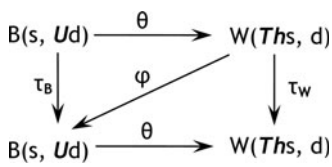
Based on this definition the following special cases can be defined:

- If there is an object s in B or d in W with $\tau_{\mathbf{B}} = 1_{\mathrm{B(s,Ud)}}$ or $\tau_{\mathbf{W}} = 1_{\mathrm{W(Fs,d)}}$, where $1_{\mathrm{B(s,Ud)}}$ and $1_{\mathrm{W(Fs,d)}}$ are identity arrows over B(s, $\boldsymbol{U}$d) and W($\boldsymbol{F}$s, d), respectively, then for the objects s in B and d in W there is a universal arrow η_{s}:s → $\boldsymbol{UF}$s and ε_{d}:$\boldsymbol{FU}$d → d, respectively.
- If for every object s in B or d in W the arrows $\tau_{\mathbf{B}} = 1_{\mathrm{B(s,Ud)}}$ or $\tau_{\mathbf{W}} = 1_{\mathrm{W(Fs,d)}}$, that is, when the arrows $\tau_{\mathbf{B}}$ and $\tau_{\mathbf{W}}$ are identity arrows for every object in B or W, then θ and φ form a bijection that is natural in s and d (hence, the tuple $<\boldsymbol{F}, \boldsymbol{U}, \varphi, \theta>$ is an adjunction).

The above definition of weak adjunction is related with two important ideas in category theoretic algebra. First, the concept of weak adjunction implies that universal properties are applicable *for some* objects s in B or d in W, whereas the concept of strong adjunction implies that universal properties are applicable *for every object* s in B and d in W. Second, the concept of weak adjunction implies that the very meaning of universal construction is also weakened, in the sense that the arrows i and m are not required to be unique. The second aspect of weak adjunction has appeared in the literature in different forms (e.g., Maranda, 1964; Kainen, 1971; Mac Lane, 1998) but also in relation to higher dimensional or 2-categories (e.g., Seely, 1979). In this paper, the proposed definition of weak adjunction covers both dimensions of "weakness," but more importantly, it is formulated in a way that will naturally lead to the subsequent formal notions of weak theory and the complexity theoretic notion of phase transition in the intentional state of a system.

More specifically, based on this construction, a *weak theory* is defined as follows.

DEFINITION (weak theory). A weak theory is a category $\boldsymbol{Th}$s that is constructed by a sketch s in B, and a functor $\boldsymbol{Th}$ such that a weak adjunction $<\boldsymbol{Th}, \boldsymbol{U}, \varphi, \theta, \tau_{\mathbf{B}}, \tau_{\mathbf{W}}>$ is defined; that is, the relation φ and θ between interpretations i of theories $\boldsymbol{Th}$s in W and models m of a sketch s in B are determined by the following diagram:



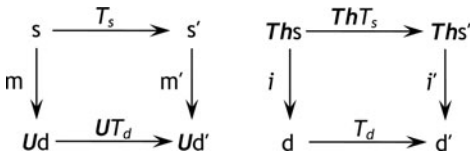
https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

Well-formed and *random* theories can be thought as special cases of weak theories in the following sense:

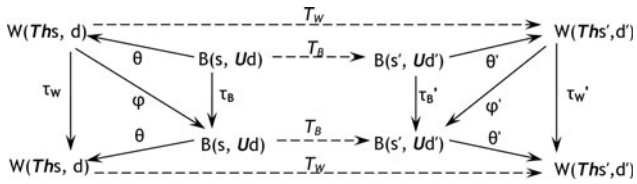
- A well-formed theory is constructed when for every object s in B or d in W the arrows $\tau_{\mathbf{B}} = 1_{\mathrm{B}(s,\boldsymbol{U}d)}$ and $\tau_{\mathbf{W}} = 1_{\mathrm{W}(\boldsymbol{Th}s,d)}$, that is, when the arrows $\tau_{\mathbf{B}}$ and $\tau_{\mathbf{W}}$ are identity arrows for every object in B or W. In this case, the arrows θ and φ form a bijection that is natural in s and d, and the tuple $<\boldsymbol{Th}, \boldsymbol{U}, \varphi, \theta>$ is an adjunction.
- A random theory is constructed when there is no object s in B or d in W such that $\tau_{\mathbf{B}} = 1_{\mathrm{B}(s,\boldsymbol{U}d)}$ or $\tau_{\mathbf{W}} = 1_{\mathrm{W}(\boldsymbol{Th}s,d)}$. In this case. the adjunction $<\boldsymbol{Th}, \boldsymbol{U}, \varphi, \theta>$ is broken. ■

The notion of weak theory can now be used to build a mathematical construction that describes a qualitative change (a *phase transition*) in the organizational complexity of the mathematical structures that characterize an intentional state. A *phase transition* is perceived as a transformation of the properties and degree of complementarity between descriptions (theories) and their interpretations (models).

DEFINITION (phase transition). Given the arrows T_s:s → s′ and T_d:d → d′ shown below,



a phase transition is defined by the transformations $T_B = B(T_s, \boldsymbol{U}T_d)$:B(s, **U**d) → B(s′, **U**d′) and $T_W = W(\boldsymbol{Th}T_s, T_d)$:W(***Th***s, d) → W(***Th***s′, d′) that make the following diagram commute:



When the arrows $\tau_{\mathbf{B}}'$ and $\tau_{\mathbf{W}}'$ in the above diagram are identity arrows then the transition is a *phase transition to universality.* Let us now bring together all the introduced mathematical constructions in order to formalize an organizational level description of design intentionality.

5. ORGANIZATIONAL LEVEL DESCRIPTION OF DESIGN INTENTIONALITY

As discussed in Section 2, the core hypothesis of this study is that design activity arises in response to a situation where there are conflicting, inconsistent, or unsatisfied intentional states expressed in the brain, a social group or any other intentional system. This situation is now more precisely approached as an intentional state where desires about the world generate expressions of theories and/or models that do not follow the correspondence between theories and models as this is established by one's belief system. A design task requires forming an interpretation of this situation (and therefore constructing a vision about the world) together with an instantiation able to fulfill this vision. This section aims to mathematically characterize the organizational structures and processes of an intentional system that is capable to address such design tasks.

5.1. Core elements of design intentionality: An example

Let us first explicate the core elements of design intentionality using an example. The objective of this example is to provide one plausible interpretation of the mathematical structures and processes that characterize the formation of intentional states during design tasks. Note that the example is used in order to instantiate and clarify the main mathematical ideas rather than to demonstrate the scope of their applicability. As explained in Section 3.2.2, there are other possible ways to apply the notion of sketch that may focus on the underlying dynamical processes or topological structures. In any case the core elements and descriptions of the properties that underlie design intentionality would remain the same.

In this example the object of design activity and content of design intentionality is restricted to building layouts. It will be assumed that there is a need or desire for a building layout in which different activities are organized into separate wings. The task may also specify constraints reflected in the form and function of the building, for instance, that certain activities should not be adjacent, whereas certain others should be placed in closed proximity. The task may also express requirements about the style and overall form of the building layout, for example, the requirement that the overall configuration of the building forms an interesting jagged outline.

All these properties, requirements, or constraints are essential components of an intentional state (i.e., a desire) that may take different and possibly conflicting or incompatible interpretations. In other words, the organization of activities into wings, the requirement for an "interesting" or "jagged" morphology, but also the constraints regarding the proximity between activities, may all take a number of different interpretations and as a result represent conflicting intentional states. Responding to a design task requires the ability to generate a consistent interpretation of this situation and construct a vision about the meaning of all these terms. It is also requires the ability to instantiate this vision in a specific building configuration that satisfies the formulated interpretations. Building on this example, let us now revisit the main mathematical structures introduced in this paper and see how these mathematical entities realize the core elements of design intentionality.

5.1.1. The notion of a sketch

A sketch s is a mathematical expression of the organizational structure of an intentional system (e.g., brain or social group). This structure captures the content and conditions of satisfaction of intentionality. In the specific example, a sketch s may be thought to capture the desired properties of a family of built con-

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

figurations: that is, the desired type of activities, the wing forms that compose the built configuration, as well as their spatial relations. In more formal terms, the notion of sketch in this example can be thought to capture a category of shape configurations in the same way that a shape grammar captures a language of desired designs. A sketch is specified by graphs, diagrams. and cones in the same way that a shape grammar is specified by shapes and shape rules, labels that constrain the application of rules, and an initial shape. The notion of a sketch here will be therefore explained as a mathematical formalism equivalent to grammars. However, sketches can also operate at a higher level of abstraction as a mathematical machinery for defining (shape) grammars and their operation. Let us explore this relation as a basis for exemplifying the notion of sketch.

A shape grammar is defined by a finite set V of shapes (i.e., a vocabulary of shapes), an initial shape I, and a finite set R of rules $r_1, r_2, \ldots, r_n$ from V^+, to V*, where V^+ is the set of all possible shapes made up of the shapes in V and V* is the set V^+ plus the empty shape. Shapes may also have labels L that guide or constrain the application of rules.

As discussed before, shapes and spatial relations in the sets V, V^+, and V* can be specified by graphs. For instance, for the purpose of this example let us consider a vocabulary consisting of one rectangular shape. In Figure 5, the graph specifies the two diagonal corners of the rectangle.

From a category theoretic perspective, the graph structure G_s of a sketch s is a mathematical structure that is used in order to represent not only shapes and spatial relations but also the set of shape rules $R = \{r_1, r_2, \ldots, r_n\}$ of a grammar. This graph structure has *n*-arrows of the form $r_1, r_2, \ldots, r_n: V^+ \rightarrow V^*$. For the purpose of this example we will assume a shape grammar that has only one rule of the form r_1:a → b as Figure 6 suggests. In relation to this grammar, the graph G_s of the sketch s is simply a graph with two nodes (shapes a and b) and one arrow that constitutes the rule r_1.

More generally regarding any grammar, the graph G_s of a sketch s captures the idea that a shape rule is constructed by two functions $h: R \rightarrow V^+$ and $t: R \rightarrow V^*$ such that for each arrow (i.e.. rule) in R the function *h* assigns a shape at the head of the arrow, and the function *t* assigns a shape at the tail of the arrow. In this more abstract sense, the graph specifies the main types of entities R, V^+, and V* involved in the definition of a shape rule and the basic form of the shape rule (how the rule is constructed) as follows:

$$V^+ \xleftarrow{h} R \xrightarrow{t} V^*$$

A shape grammar may also include a set of labeled points L. For instance, the end point (x_2, y_2) of the graph $(x_1, y_1) \rightarrow (x_2, y_2)$ may be labeled with a symbol ● as Figure 7 suggests.

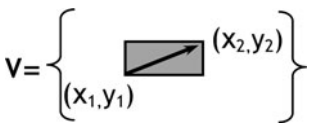


Fig. 5. A vocabulary of shapes V. It consists of a rectangle defined by a graph $(x_1, y_1) \rightarrow (x_2, y_2)$ that specifies its two diagonal corners.

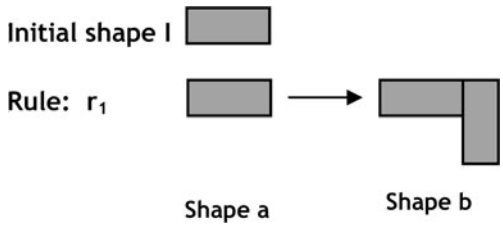
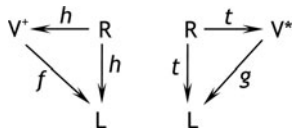


Fig. 6. A simple shape grammar with one rule.

Shape rules are then of the form $r_1, r_2, \ldots, r_n: (V, L)^+ \rightarrow (V, L)^*$ where $(V, L)^+$ denotes the set of labeled shapes made up of shapes in V and symbols in L, whereas (V, L)* also includes the empty labeled shape. The labeled points L that are assigned over shapes in V^+, and V* (as shown in Fig. 8) break the shape symmetries and as a result constrain the application of shape rules.

The use of labeled points in a shape grammar corresponds to the use of diagrams in a sketch. More generally, diagrams for a labeled shape grammar are formed by functions $f: V^+ \rightarrow L$ and $g: V^* \rightarrow L$ such that *f* assigns labeled points on the left hand shape of a rule in R and *g* assigns labeled points on the right-hand shape of the rule in R as shown below:



Similarly, there is a correspondence between the use of an initial shape in a shape grammar and the use of cones in a sketch. The initial shape is an element of the set of possible shapes V^+ with the unique property that all other shapes in the constructed language (or category) of shapes are derived from this shape. An initial shape can be perceived as the minimum shape of the language. Generally a cone determines operations such as products or sums that create new structures out of more primitive ones. In the specific example, a cone may specify an algebra of shapes (e.g., an object a ◦ b that is the product of two shapes a and b). A cone may also determine other universal constructions such as characteristic functions that determine a mechanism of choosing possible or desirable subshapes within observed spatial relations. In the specific example we will use only the initial shape as a cone of the sketch s.

5.1.2. The notion of a theory

In the same way that a shape grammar generates a language of shapes, a sketch generates a theory or category ***Th***s of shapes that satisfy the conditions expressed in s. The core idea is that the functor ***Th*** takes a sketch s and generates a category of

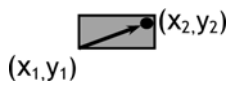


Fig. 7. A labeled shape and the corresponding labeled graph $(x_1, y_1) \rightarrow (x_2, y_2)$.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

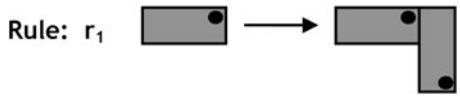


Fig. 8. A labeled shape rule.

shapes (language) ***Th***s where each arrow in the category ***Th***s is a possible path of arrows (i.e., composition of rules) defined in G_S under the constraints imposed by the diagrams (i.e., labeled points) of the sketch s. Hence, the category ***Th***s has as arrows all possible applications of the rules in R and as objects all the labeled shapes from the sets $(V, L)^+$ and (V, L)* generated by rules in R. In the specific example, the language or theory ***Th***s is essentially a category of shapes such as in Figure 9. The picture presents a particular derivation of shape configurations within a category of shapes generated by s.

5.1.3. Interpretations and models

A language of shapes ***Th***s may take different interpretations. Interpretations are expressed in a certain language that describes these configurations in terms of purpose, function, form, or organizational characteristics (for a similar treatment in the context of shape grammars, see Stiny, 1981). In shape grammars such descriptions are specified by a function i:$\boldsymbol{Th}_G \rightarrow$ D that maps configurations expressed in the language/theory of shapes $\boldsymbol{Th}_G$ generated by grammar G to descriptions in another set D. Similarly, in category theoretic terms interpretations are specified by a functor i:***Th***s → d that maps shape configurations in the category of shapes ***Th***s into descriptions of another language d, preserving the structure of the shape language. For instance, we could envisage descriptions of shape configurations in terms of form types {I-shape; L-shape, U-shape, O-Shape, . . .}, type or number of wings {a, b, c, d, . . .}, number of courts, and a matrix that specifies spatial relations between wings. Hence, a sequence of descriptions of shape configurations (interpretations) can be formed that preserve the properties of the shape language ***Th***s as Figure 10 suggests.

Any description d of a particular building configuration has an underlying sketch representation, denoted by ***U***d, which specifies the shape rules that generate the described configuration (i.e., underlying graph structure), but also the relations and constraints imposed in the application of spatial rules (i.e., diagrams). For instance the O-shape configuration in Figure 10 has an underlying sketch with one rule (the rule r_1 as expressed in Fig. 8) and one diagram, which specifies that the building cofiguration is generated from rule r_1 applied three times. Thus, the underlying sketch of an O-shape configuration has an arrow r_1^3:$(V, L)^+ \rightarrow (V, L)^*$ and a diagram that equates the arrow r_1^3 with the application of the rule r_1 three times (i.e., $r_1^3 = r_1 \circ r_1 \circ r_1$).

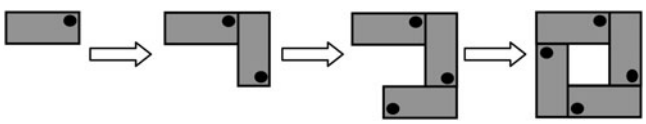


Fig. 9. An example of a language or theory ***Th***s of shapes. ***Th***s is a category of shapes generated by the sketch s described in Figure 8. Note that the arrows in this picture are not shape rules but applications of shape rule r_1.

A model of a sketch or model of an intentional state s is a desired and/or possible specification of the properties of building configurations. This is mathematically expressed by a functor m:s → ***U***d, which maps the desired properties (expressed in s) into a design description ***U***d. A complementary relation between models and interpretations implies that a model m:s → ***U***d preserves the desired properties of the sketch s into the underlying sketch of the proposed design description ***U***d. For instance, in Figure 11, the sketch s generates a language of building configurations by applying the shape rule r_1 for any number of times. A model of a building layout may express the desire for a configuration that forms courtyards. Building configurations with courtyards (i.e., O-shape configurations) are generated by applying rule r_1 thrice (as discussed above). The underlying sketch of this configuration satisfies the model of an intentional state s because there is a functor m:s → ***U***d that can preserve the structure of s in the underlying sketch ***U***d of the proposed O-shape configuration. A building configuration with a courtyard is then created by the sketch s because there is a building configuration in the language ***Th***s and an interpretation i:***Th***s → d that satisfy this model and make the diagram in Figure 11 commute.

5.1.4. The nature of models and their interpretations

Figure 11 is an illustration of the relation between sketches, theories, models, and interpretations. It is worth noting that in this study there is no commitment on the nature of semantic relations. As explained, a sketch s and its models (i.e., m:s → ***U***d) are mathematical expressions of the conditions of satisfaction of an intentional state that are realized in an intentional system. However, the category of shapes ***Th***s and their interpretations i:***Th***s → d are mathematical expressions of an existing or possible world W that may be realized either as a cognitive structure (conceptualist view) or as an external representation (realist view). Because this paper does not make an explicit commitment to any one of these two views, it will be more abstractly said that the category B is a mathematical expression of the *subjective reality* realized in the "brain" of an intentional organism, and category W is a mathematical expression of an *objective reality (observed or possible)* realized in the brain or alternatively in the task environment of an intentional organism. A similar discussion on the distinction between subjective and objective reality can be found in Lawrere and Schanuel (1997).

5.2. The design task

Looking back at Figure 11, it is possible to envisage a number of possible conflicts or inconsistencies that may arise in the intentional states of an organism. First, a sketch may contain conflicting or incomplete information regarding an intentional object. For instance, it is possible to envisage the exis-

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

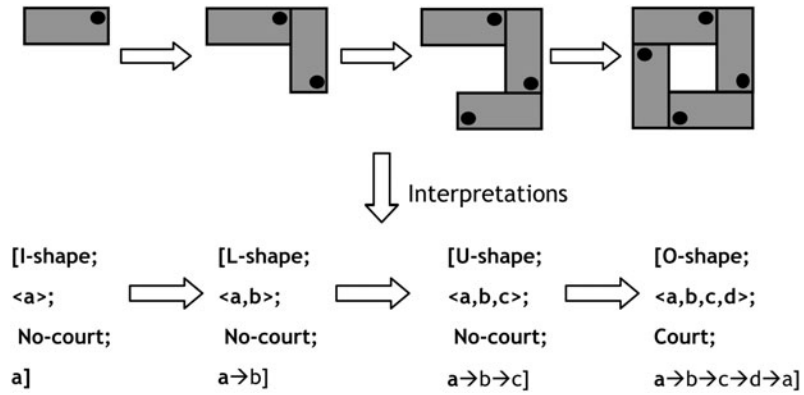


Fig. 10. A language of shapes and their interpretations. Note how the sequence of descriptions of shape configurations corresponds to the sequence of shape configurations.

tence of diagrams that represent conflicting spatial relations regarding the room adjacency. Another possible situation is to envisage the existence of an inconsistent language. An inconsistent language of shapes ***Th***s will generate shape configurations that have properties that conflict with the properties of shape configurations ***U***d described by the desired models. For instance, let us assume a language of shapes that generates Z-shape configurations while the sketch or intentional state s remains the same (Fig. 12). In this case, the sketch or intentional state s clearly has models that cannot specify the properties of this type of configurations.

It is equally possible to envisage an incomplete language of shapes or incomplete descriptions of building configurations. An incomplete language of shapes ***Th***s will be a category of shapes that cannot generate certain desired configurations. For instance, let us assume a situation where there is a desire for Z-shape configurations expressed by the underlying sketch shown in Figure 13. In this case, any interpretation of shape configurations generated in ***Th***s will not satisfy the description of the desired building configurations.

In all of these scenarios it is assumed that the elements of Figure 12 and Figure 13 are mathematical structures that are constructed or realized within an intentional system (e.g., a brain) somewhat independently from one another.

These examples illustrate the characteristic features of the situation that motivates design, as discussed in Section 2. More generally, in category theoretic terms, design intentionality appears when the developed theory ***Th***s of desired objects/processes yields interpretations in W (i.e., i:***Th***s $\rightarrow$ d) whose underlying properties ***U***d cannot be (uniquely) derived by the sketch s (i.e., m:s $\rightarrow$ ***U***d). The need to design therefore arises when the theory and model functors contain, so to speak, ambiguity, noise, or errors; hence, there is no natural bijection between B(s, ***U***d) and W(***Th***s, d).

In this situation, there is conflicting intentionality (e.g., conflicting beliefs and desires) that generates a two direc-

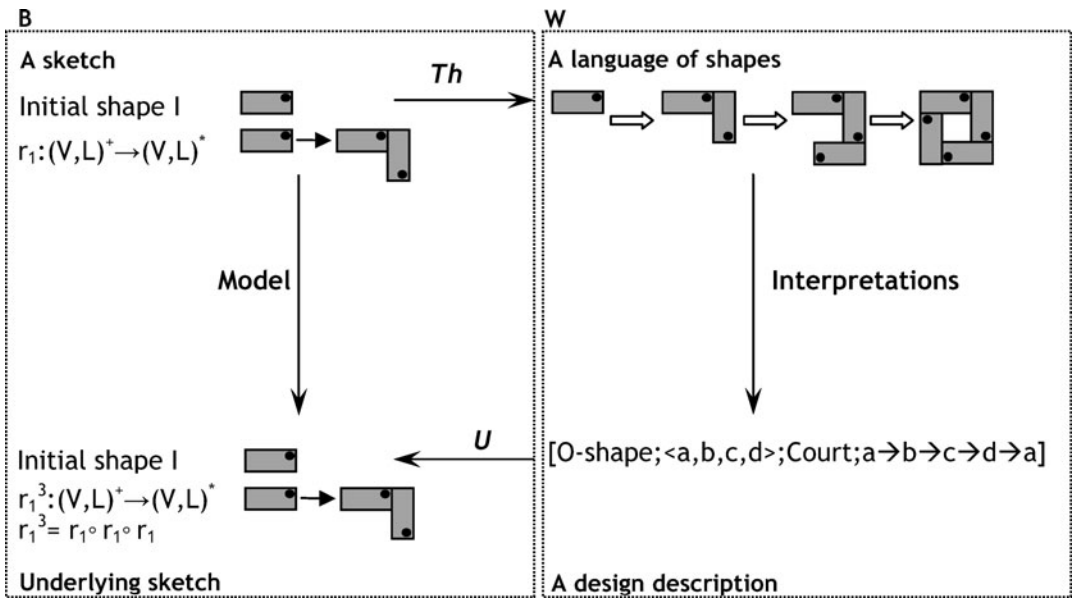


Fig. 11. An example of the adjunction between models of desired shape configurations and their interpretations.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

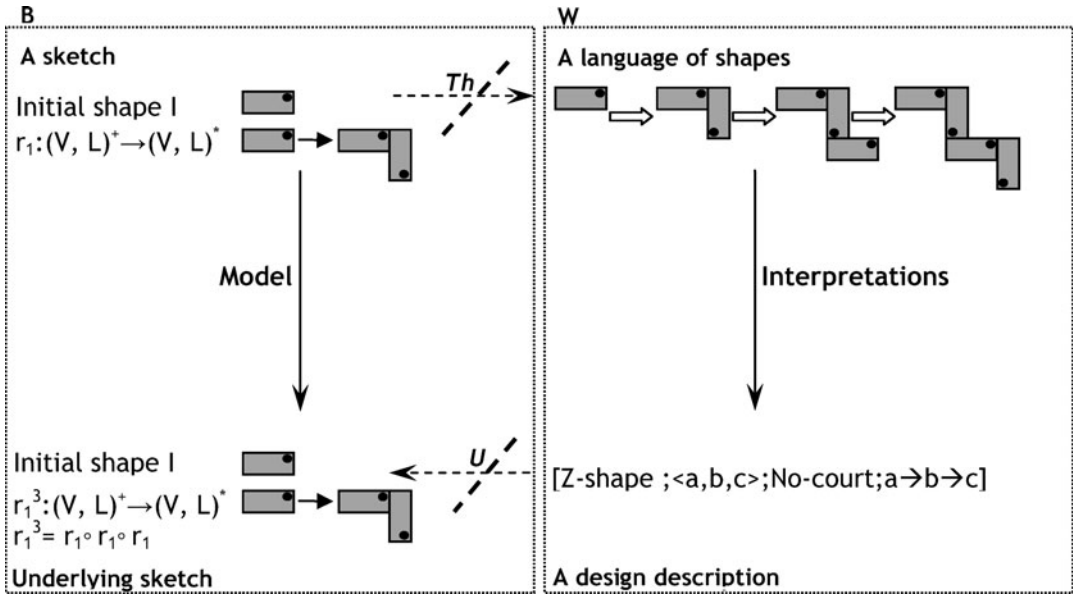


Fig. 12. The language of Z-shape configurations in W has properties that cannot be described by the models of the sketch.

tional freedom of fit: from world-to-mind and from mind-to-world (recall after Section 3.2.2). Namely, there is a mismatch between subjective B(s, ***U***d) and objective reality W(***Th***s, d) and both the subjective and the objective reality need to be adapted to each other. In this sense, a sketch s expresses the conditions of satisfaction of a desire that is fulfilled only when there is theory of a possible reality that satisfies its models. In contrast, a sketch s also expresses the conditions of satisfaction of a belief that becomes true (or false) only when there is a model of a possible world that validates the theory. To put it differently, design arises when the diagram in Figure 14 does not commute.

The following definition specifies the core hypothesis regarding the mathematical structures that characterize a design capable intentional system when faced with a design problem or design task.

THESIS 1 (design intentionality and the design task). Design intentionality arises in relation to an intentional state s with a "two-directional freedom of fit" between a subjective

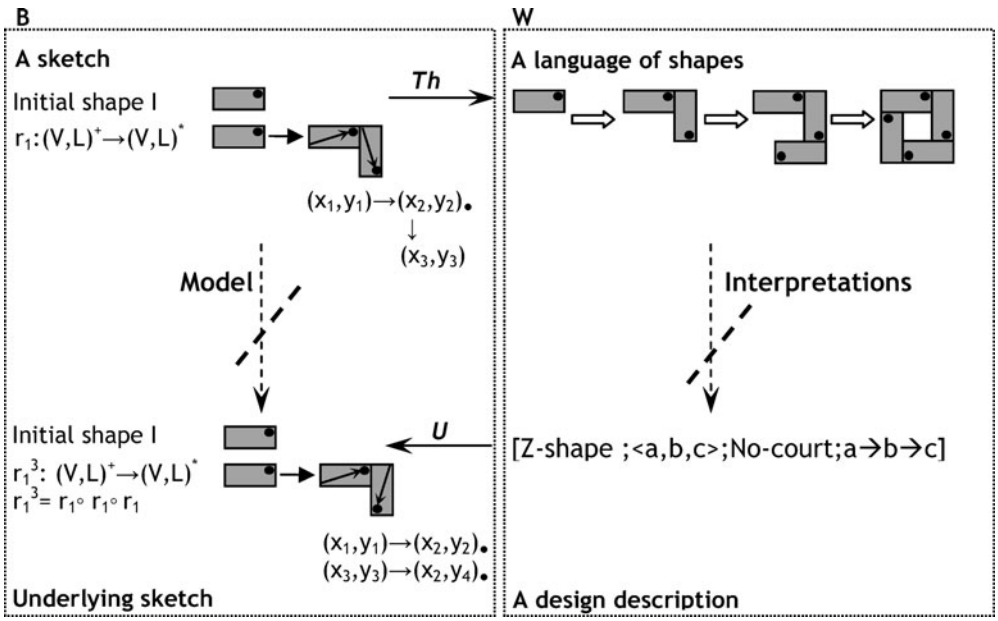


Fig. 13. The language of shapes in W does not have the expressive power to generate configurations that would satisfy the underlying sketch of Z-shape configurations.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

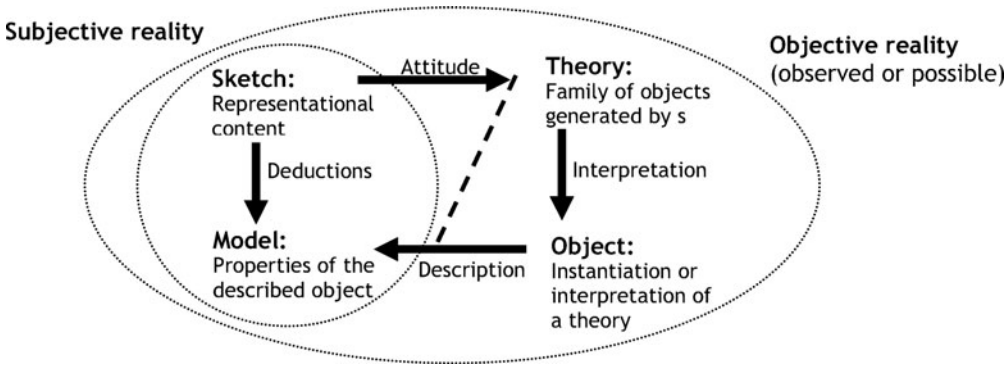
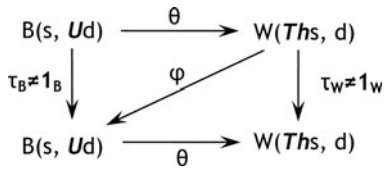


Fig. 14. A design situation arises when there is no natural bijection between subjective B(s, ***U***d) and objective W(***Th***s, d) reality.

B(_, ***U***_) and an objective reality W(***Th***_, _), whose models (i.e., *m*:s → ***U***d) in B(s, ***U***_) and interpretations (i.e., *i*:***Th***s → d) in W(***Th***_, d) are not complementary. Namely, design intentionality arises when there is a weak adjunction $< \boldsymbol{Th}, \boldsymbol{U}, \varphi, \theta, \tau_B, \tau_W >$ natural in s and d that makes the following diagram commute:



Given this intentional state, the problem of design or *design task* is to establish an adjunction B(s, ***U***d) ≅ W(***Th***s, d) by transforming both the subjective and the objective reality. ■

5.3. The capacity to design

In response to this intentional state, the design task requires the mental capacity to make the diagrams in Figures 12, 13, and 14 commute. Namely, it requires the capacity to bring the conditions of satisfaction s and the generated model *m*:s → ***U***d of a desired reality into a complementary relation (i.e., adjunction) with the generated language ***Th***s of design objects and their interpretations *i*:***Th***s → d. The main consequence of the above hypothesis is that the mental capacity to carry out design tasks (i.e., the capacity to make the diagrams in Figures 12, 13, and 14 commute) requires a type of phase transition in the organizational complexity of intentional states as described in Section 4. Let us explain further this result first in an informal and then a formal way.

Going back to the building configuration task example, the ability to design alludes to the capacity to generate an interpretation or vision of a desired building layout together with the capacity to generate a language of shape configurations that will instantiate this interpretation or vision. Let us for instance focus on the desired model and description for a "jagged layout." In this context, the ability to design requires the capacity to transform the description of a "jagged layout" into a description for a Z-shape configuration, but also the capacity to identify certain spatial relations between the wings of the building that generate a language of Z-shape configurations. These transformations create a new interpretation of what is a jagged layout configuration, but also a new language of how this vision can be instantiated. More specifically, in Figure 15 we see that the desire for a jagged-shape configuration expressed by the description d_D (bottom right of the figure) is part of an intentional state that expresses inconsistencies. The dashed lines denote that an adjunction between B(s, ***U***d) and W(***Th***s, d) cannot be formed given this description. A new interpretation of what is a jagged layout configuration d_U, but also a new sketch (rule) r_U and language of how this vision can be instantiated (see thick arrows) are needed in order to realize a transition to a universal state (an adjunction represented by solid lines).

However, what are the organizational structures and processes that drive these mental transformations during a design task? The transformation of a sketch T_s:s → s_U (or more narrowly in this example the transformation of shape rules T_R: r_I → r_U), as well as the transformation of design descriptions T_d:d → d_U, are components of a transition that leads into a new relation between desired models, language of shapes, and interpretations. The core idea is that the adjunction corresponds to an organizational state that works as an attractor for these transformations. Thus, these transformations are defined in relation to a universal state (the codomains of the arrows are part of an adjunction) although the transformations are not universal arrows. In effect, these transformations can be understood as *anticipatory representations* of a transition to complementary relations between desired models and interpretations. The notion of anticipatory representation simply refers to the fact that both transformations express mental actions in preparation of a new language of shapes and a new model of the desired building that would make the diagram commute. Moreover, both transformations express mental actions that lead to significant changes in the organizational complexity of the intentional state (from a weakly universal to a universal state).

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

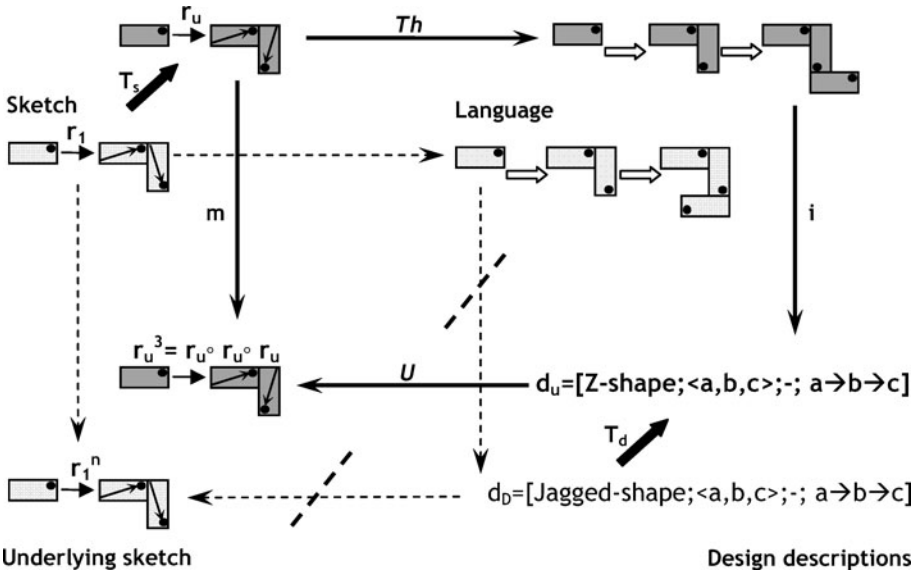


Fig. 15. An example of a phase transition. In the absence of an adjunction, a new interpretation of what is a jagged layout configuration d_U, but also a new sketch (rule) r_U and language of how this vision can be instantiated (see thick arrows) are needed in order to realize a transition to a universal state (solid arrows).

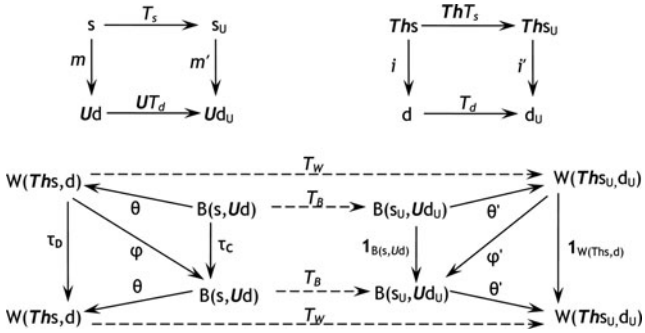
Let us explain the meaning and peculiarity of this phase transition, but also why this is a consequence of the very nature of design task. According to the proposed thesis on the nature of design tasks, design requires a transition from a situation that is characterized by a weak theory to a situation that is characterized by a universal construction. One important peculiarity of this transition is that it leads to universality. However, although this transition leads to a universal construction, there is *no universal* transition to universality. This is the very significance of the concept of phase transition. If the transformation T_s:s → s_U (from a sketch s participating in a weak adjunction to a sketch s_U participating in an well-defined adjunction) could be constructed through a universal arrow, then the theory ***Th***s would be a universal theory. However, this would contradict the main premise behind the definition of the design task that demands the existence of a weak theory. In this sense, design necessitates the capacity of generating theories and models of a sketch in preparation of their adjunction, that is, before such correspondence is constructed. It is in this sense that the mental process of design is characterized as a phase transition to universality, a phase transition that requires the capacity to hold anticipatory representations of universal constructions (for more details on the mathematical conditions that explain the construction of anticipatory representations during design tasks, see Zamenopoulos & Alexiou, 2007*b*).

We now provide a more formal presentation of these ideas. The mental activity of design is a phase transition to universality and as a result requires the existence of transformations:

$$T_B = \mathrm{B}(T_s, \boldsymbol{U}T_d):\mathrm{B}(s, \boldsymbol{U}\mathrm{d}) \rightarrow \mathrm{B}(\mathrm{s_U}, \boldsymbol{U}\mathrm{d_U}) \quad \text{and}$$

$$T_W = \mathrm{W}(\boldsymbol{Th}T_s, T_d):\mathrm{W}(\boldsymbol{Th}\mathrm{s}, \mathrm{d}) \rightarrow \mathrm{W}(\boldsymbol{Th}\mathrm{s_U}, \mathrm{d_U})$$

such that the following diagrams commute:



Given an intentional state s, the capacity to carry out design tasks alludes to the existence of transformations T_S:s → s_U in a subjective reality B(_, ***U***_) and transformations T_d:d → d_U in an objective reality W(***Th***_, _) that are *components of a phase transition to universality* $T = <T_B, T_W>$. The transformation T_s:s → s_u is a component of the transformation B(T_s, $\boldsymbol{U}T_d$) = T_B:B(s, ***U***d) → B(s_U, ***U***d_U), and similarly the transformation T_d:d → d_U is a component of the transformation W(***Th***T_s, T_d) = T_D:W(***Th***s, d) → W(***Th***s_U, d_U). It is informative to think that the transformation T_s:s → s_u has a mind to world direction of fit, in the sense that any mismatch between a subjective B(_, ***U***_) and objective reality W(***Th***_, _) is addressed by changing the properties of the sketch. Similarly, the transformation T_d:d → d_U has a world to mind direction of fit because any mismatch between the subjective and objective reality is addressed by changing the properties of the object.

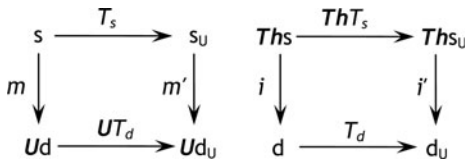
In category theoretic literature, the notion of adjunction B(s, ***U***d) ≅ W(***Th***s, d) can be alternatively presented as a relation between two functors W(***Th***s, d):$\mathrm{B^{op}}$xW → **Set** and

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

B(s, ***U***d):B^{op}xW → **Set** that are naturally isomorphic. In this context, the theory ***Th***s is said to be a *representing object* for the functor B(s, ***U***d). Similarly, the underlying sketch ***U***d of an entity d is said to be a representing object for the functor W(***Th***s, d). However, what can be said for the functors W(***Th***s, d) and B(s, ***U***d) that are not naturally isomorphic, that is when B(s, ***U***d) ≅ W(***Th***s, d)? In this case, the notion of an anticipatory representation is defined. The theory ***Th***s generated by a sketch s is called an *anticipatory representation of models* in B(s_U, ***U***d_U) if and only if the sketch s has models *m* and interpretations *i* that are components of a phase transition to universality $T = <T_B, T_W>$. The underlying sketch ***U***d is then called an *anticipatory representation of interpretations* in W(***Th***s_U, d_U). Based on this definition, the capacity to carry out design tasks is now identified as the capacity to hold anticipatory representations/interpretations of universality.

THESIS 2 (the capacity to design). The capacity to design is the capacity of an intentional state s to hold models T_s:s → s_U and generate interpretations ***Th***T_s:***Th***s → ***Th***s_U that are *components of a phase transition to universality*. The capacity to design is therefore the capacity of a sketch s to generate theories ***Th***s and sketches ***U***d of an object d that are anticipatory representations of a universal construction B(s_U, ***U***d_U) ≅ W(***Th***s_U, d_U). ■

Note that the models *m* and theories ***Th***s that are derived from an intentional state s are specified in relation to *a phase transition* to universality. Such anticipatory representations are clearly distinct from representations where models and theories are specified in relation to *a universal construction*. The adjunction B(s_U, ***U***d_U) ≅ W(***Th***s_U, d_U) characterizes an intentional state s in the sense that every model of s and interpretation of theories ***Th***s are constrained by the specified phase transition to universality. However, the reverse is not true. The models of the sketch s are not universal representations of W(***Th***s_U, d_U). In this sense the adjunction B(s_U, ***U***d_U) ≅ W(***Th***s_U, d_U) can be perceived as an "emergent" state of s.



6. SUMMARY AND CONCLUSIONS

This paper presented a mathematical theory that aimed to associate our understanding of design activity to a theory of mind and intentionality. The paper focused on the property of the mind to hold intentional states (that is, the capacity to represent or reflect existing and nonexisting realities) and on the way that these mental states are constructed during design tasks. The core objective was the development of a mathematical framework that would explicate a complexity theoretic hypothesis about the organizational structures and processes that govern the formation of intentional states during design tasks. The paper elaborated a mathematical framework regarding the organizational complexity of intentionality, which was then used in order to propose an organizational level description of design intentionality. This framework was based on the introduction and development of the category theoretic notion of weak theory.

There are two potential contributions of the proposed treatment. The first is related with the need (or ambition) of developing explanatory theories about the capacity to carry out design tasks. Such theories should be able to uniquely distinguish design (from nondesign phenomena) on the basis of certain underlying principles. Such principles may be identified at different levels of abstraction (e.g., at the level of neurological activity, information processing, knowledge level or social level). Although the present paper does not propose a "mechanism" to explain the capacity to carry out design tasks, it does propose a framework within which certain conditions (or mechanisms) may be identified. Overall, the main value of this study alludes to the rigorous formulation of the type or category of mathematical structures that underlie a design-capable intentional system (e.g., a brain). For example, the proposed mathematical framework can be used in the analysis of brain imaging data derived from functional magnetic resonance imaging methods for the purpose of identifying characteristic patterns in the structure as well as the effective and functional connectivity between brain areas during design tasks. As discussed in the introduction, although we can empirically collect data about the neurological structures and processes that partake in the completion of design tasks, we have no formal way of linking these structures and processes to the formation of mental intentional states. The potential of using the mathematical framework for exactly this purpose is explained in more detail in Zamenopoulos and Alexiou (2011).

The second contribution is related to the need of developing theories in design that are informed by and are transferable to other domains. Although this is not an uncommon objective in design research in general, the proposed framework explicitly places the problem of design within the realm of complexity research and vice versa. The constitution of design as a universal capacity of complexity is a double contribution to complexity research as well as to design research, and makes it possible to transfer results between the two fields. In this sense, it is hoped that the proposed mathematical framework offers a plausible interpretation on how characteristic structures and processes of complex (neurological, social, or ecological) systems may be linked to higher level cognitive functions such as designing (Zamenopoulos and Alexiou, 2005*b*).

REFERENCES

Akin, Ö. (1986). *Psychology of Architectural Design*. London: Pion Limited.

Alexiou, A. (2007). *Understanding multi-agent design as coordination*. PhD Thesis. University of London, University College London.

Alexiou, K., & Zamenopoulos, T. (2008). Design as a social process: a complex systems perspective. *Futures 40(6)*, 586–595.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

Alexiou, K., Zamenopoulos, T., Johnson, J.H., & Gilbert, S.J. (2009). Exploring the neurological basis of design cognition using brain imaging: some preliminary results. *Design Studies 30(6)*, 623–647.
Archer, B.L. (1965). Systematic methods for designers. In *Developments in Design Methodology* (Cross, N., Ed.), pp. 57–82. Chichester: Wiley.
Atlan, H. (1974). On a formal definition of organization. *Journal of Theoretical Biology 45(2)*, 295–304.
Atlan, H. (1998). Intentional self-organization. Emergence and reduction: towards a physical theory of intentionality. *Thesis Eleven 52(1)*, 5–34.
Badii, R., & Politi, A. (1997). *Complexity: Hierarchical Structures and Scaling in Physics*. Cambridge: Cambridge University Press.
Barr, M., & Wells, C. (1985). *Toposes, Triples and Theories*. New York: Springer.
Barr, M., & Wells, C. (1990). *Category Theory for Computing Science*. Englewood Cliffs, NJ: Prentice Hall.
Barwise, J. (1977). *A Handbook of Mathematical Logic*. Amsterdam: North-Holland.
Barwise, J., & Perry, J. (1983). *Situations and Attitudes*. Cambridge, MA: MIT Press.
Beaney, M. (Ed.). (1997). *The Frege Reader*. Oxford: Blackwell.
Bennett, C.H. (1985). Dissipation, information, computational complexity and the definition of organization. In *Emerging Syntheses in Science* (Pines, D., Ed.), pp. 215–233. Reading, MA: Addison–Wesley.
Boccara, N. (2004). *Modeling Complex Systems*. New York: Springer–Verlag.
Brazier, F.M.T., Moshkina, L.V., & Wijngaards, N.J.E. (2001). Knowledge level model of an individual designer as an agent in collaborative distributed design. *Artificial Intelligence in Engineering 15(5)*, 137–152.
Brentano, F. (1995). *Psychology From an Empirical Standpoint*. New York: Routledge.
Cariani, P. (1991). Emergence and artificial life. In *Artificial Life II* (Langton, C., Taylor, C., Farmer, J.D., & Rasmussen, S., Eds.), pp. 775–797. Redwood City, CA: Addison–Wesley.
Casti, J.L. (1992). *Reality Rules: Picturing the World in Mathematics—The Fundamentals*, Vol. 1. New York: Wiley.
Chaitin, G.J. (1966). On the length of programs for computing finite binary sequences. *Journal of the Association of Computing Machinery 13*, 547–569.
Chaitin, G.J. (1997). *The Limits of Mathematics: A Course on Information Theory and Limits of Formal Reasoning*. Singapore: Springer–Verlag.
Chalmers, D.J. (2002). Foundations. In *Philosophy of Mind. Classical and Contemporary Readings* (Chalmers, D.J., Ed.), pp. 1–9. New York: Oxford University Press.
Coyne, R.D. (1988). *Logic Models of Design*. London, Pitman.
Davidson, D. (1970). Mental events. In *Experience and Theory* (Foster, L., & Swanson, J.W., Eds.), pp. 79–101. Amherst, MA: University of Massachusetts Press.
Davis, S., & Gillon, B. (Eds.). (2004). *Semantics: a Reader*. New York: Oxford University Press.
Dennett, D.C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
Dorst, K., & Cross, N. (2001). Creativity in the design process: co-evolution of problem-solution. *Design Studies 22(5)*, 425–437.
Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.
Ellerman, D.P. (1988). Category theory and concrete universals. *Erkenntnis 28(3)*, 409–429.
Ernst, G.W., & Newell, A. (1969). *GPS: A Case Study in Generality and Problem Solving*. New York: Academic.
Fodor, J.A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.
Fodor, J.A. (1987). *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
Fodor, J.A. (1995). Fodor, Jerry A. In *A Companion to the Philosophy of Mind* (Guttenplan, S., Ed.). New York: Blackwell.
Freeman, W.J., & Skarda, C.A. (1990). Representations: who needs them. In *Brain Organization and Memory: Cells, Systems and Circuits* (Mcgaugh, J.L., Weinberger, N.M., & Lynch, G., Eds.), pp. 375–380. Oxford: Oxford University Press.
Gärdenfors, P. (2004). *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press.
Gero, J.S. (1998). Towards a model of designing which includes its situatedness. In *Universal Design Theory* (Grabowski, H., Rude, S., & Grein, G., Eds.), pp. 47–56. Aachen, Germany: Shaker Verlag.
Gilbert, S.J., Zamenopoulos, T., Alexiou, A., & Johnson, J.H. (2010). Involvement of right dorsolateral prefrontal cortex in ill-structured design cognition: an fMRI study. *Brain Research 1312(2)*, 79–88.
Goel, V. (1995). *Sketches of Thought*. Cambridge, MA: MIT Press.
Goel, V., & Grafman, J. (2000). Role of the right prefrontal cortex in ill-structured planning. *Cognitive Neuropsychology 17(5)*, 415–436.
Goel, V., & Pirolli, P. (1989). Motivating the notion of generic design within information processing theory: the design problem space. *AI Magazine 10(1)*, 19–36.
Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science 16(3)*, 395–429.
Goguen, J.A. (1973). Realization is universal. *Mathematical System Theory 6*, 359–374.
Goldblatt, R. (1984). *Topoi. The Categorical Analysis of Logic*. New York: Elsevier.
Goldblatt, R. (2006). Mathematical modal logic: a view of its evolution. In *Handbook of the History of Logic* (Gabbay, D., & Woods, J., Eds.), Vol. 6. Amsterdam: Elsevier.
Grice, H.P. (1957). Meaning. *Philosophical Review 66(3)*, 377–388.
Haken, H. (1983). *Advanced Synergetics: Instability Hierarchies of Self-Organizing Systems and Devices*. Berlin: Springer–Verlag.
Harder, P. (1996). *Functional Semantics: A Theory of Meaning, Structure and Tense in English*. New York: Mouton de Gruyter.
Harman, G. (1989). Some philosophical issues in cognitive science: qualia, intentionality, and the mind–body problem. In *Foundations of Cognitive Science* (Posner, M.I., Ed.), pp. 831–848. Cambridge, MA: MIT Press.
Harnad, S. (1987). Category induction and representation. In *Categorical Perception: The Groundwork of Cognition* (Harnad, S., Ed.). New York: Cambridge University Press.
Harnad, S. (1990). The symbol grounding problem. *Physica D 42*, 335–346.
Hintikka, K.J.J. (1961). Modality and quantification. *Theoria 27*, 119–128.
Ho, Y.-S. (1982). The planning process: a formal model. *Environment and Planning B 9(4)*, 377–386.
Horgan, T. (1993). From supervenience to superdupervenience. Meeting the demands of a material world. *Mind 102(408)*, 555–586.
Jackendoff, R. (1976). *Semantics and Cognition*. Cambridge, MA: MIT Press.
Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
Kainen, P.C. (1971). Weak adjoint functors. *Mathematische Zeitschrift 122(1)*, 1–9.
Kim, J. (1992). Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research 52(1)*, 1–26.
Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior 4(3/4)*, 415–452.
Kolmogorov, A. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission 1(1)*, 1–17.
Kripke, S.A. (1963). Semantical considerations on modal logic. *Acta Philosophica Fennica 16*, 83–94.
Lambek, J., & Scott, P.J. (1986). *Introduction to Higher Order Categorical Logic*. Cambridge: Cambridge University Press.
Langton, C. (1990). Computation at the edge of chaos: phase transitions and emergent computation. *Physica D 42(1–3)*, 12–37.
Lawrere, F.W., & Schanuel, S.H. (1997). *Conceptual Mathematics: a First Introduction to Categories*. Cambridge: Cambridge University Press.
Mac Lane, S. (1998). *Categories for the Working Mathematician*. New York: Springer–Verlag.
Maher, M.L., Poon, J., & Boulanger, S. (1996). Formalizing design exploration as co-evolution. *Advances in Formal Design Methods for CAD. Proc. IFIP WG5.2 Workshop on Formal Design Methods for Computer-Aided Design* (Gero, J.S., & Sudweeks, F., Eds.). London: Chapman & Hall.
Maranda, J.M. (1964). Injective structures. *Transactions of the American Mathematical Society 110*, 98–135.
March, M.L. (Ed.). (1976). The logic of design and the question of value. In *The Architecture of Form*. Cambridge: Cambridge University Press.
McLaughlin, B.P. (1992). The rise and fall of British emergentism. In *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism* (Beckermann, A., Flohr, H., & Kim, J., Eds.), pp. 49–93. Berlin: Walter de Gruyter.
McLaughlin, B.P. (1997). Emergence and supervenience. *Intellectica 25(1)*, 25–43.
Millikan, R.G. (2004). *Varieties of Meaning. The 2002 Jean Nicod Lectures*. Cambridge, MA: MIT Press.
Mitchell, W.J. (1990). *The Logic of Architecture: Design, Computation and Cognition*. Cambridge, MA: MIT Press.
Newell, A., & Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.
Nicolis, G., & Prigogine, I. (1967). On symmetry breaking instabilities in dissipative systems. *Journal of Chemical Physics 46(9)*, 3542–3550.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press

Place, U.T. (1956). Is consciousness a brain process? *British Journal of Psychology 47*, 44–50.
Portugali, J., & Casakin, H. (2002). SIRN (Synergetic Inter-Representation Network): an approach to design. In *Common Ground: Design Research Society International Conference* (Durling, D., & Shackleton, J., Eds.). Staffordshire: Staffordshire University Press.
Putnam, H. (1973). Psychological predicates. In *Art, Mind and Religion* (Capitan, W.H., & Merrill, D.D., Eds.), pp. 37–48. Pittsburgh, PA: University of Pittsburgh Press.
Pylyshyn, Z.W. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press.
Roozenburg, N.F.M. (1993). On the pattern of reasoning in innovative design. *Design Studies 14(1)*, 4–18.
Rosen, R. (1991). *Life Itself. A Comprenhensive Inquiry into the Nature, Origin, and Fabrication of Life*. New York: Columbia University Press.
Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson and Co.
Schön, D.A. (1983). *The Reflective Practitioner. How Professionals Think in Action*. New York: Basic Books.
Schuster, H.G. (2001). *Complex Adaptive Systems: An Introduction*. Saarbrücken: Scator Verlag.
Searle, J.R. (1983). *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
Seely, R.A.G. (1979). Weak adjointness in proof theory. *Proc. Research Symp. on Applications of Sheaf Theory to Logic, Algebra, and Analysis*. Berlin: Springer/Heidelberg.
Simon, H.A. (1973). The structure of ill-structured problems. *Artificial Intelligence 4(4)*, 181–201.
Smart, J.J.C. (1959). Sensations and brain processes. *Philosophical Review 68*, 141–156.
Smithers, T. (1996). On knowledge level theories of design process. In *Artificial Intelligence in Design* (Gero, J.S., & Sudweeks, F., Eds.). Dordrecht: Kluwer.
Smithers, T. (1998). Towards a knowledge level theory of design process. In *Artificial Intelligence in Design '98* (Gero, J.S., & Sudweeks, F., Eds.), pp. 3–22. Dordrecht: Kluwer.
Smithers, T. (2002). Synthesis in design. In *Artificial Intelligence in Design* (Gero, J.S., Ed.), pp. 3–24. Cambridge: Kluwer.
Snodgrass, A., & Coyne, R. (2006). *Interpretation in Architecture: Design as a Way of Thinking*. London: Routledge.
Steadman, J.P. (1979). *The Evolution of Designs: Biological Analogy in Architecture and the Applied Arts*. Cambridge: Cambridge University Press.
Stiny, G. (1981). A note on the description of designs. *Environment and Planning B 8(3)*, 257–267.
Stiny, G. (2006). *Shape: talking about seeing and doing*. Cambridge, MA: MIT Press.
Stiny, G. & March, L. (1981). Design machines. *Environment and Planning B 8(3)*, 245–255.
Takeda, H., Veerkamp, P., Tomiyama, T., & Yoshikawa, H. (1990). Modeling design process. *AI Magazine 11(4)*, 37–48.
Tarski, A. (1956). *Logic, Semantics, Metamathematics*. New York: Oxford at the Charendon Press.
Van Gelder, T., & Port, R.F. (1995). It's about time: an overview of the dynamical approach to cognition. In *Mind as Motion: Explorations in the Dynamics of Cognition* (Port, R.F., & Van Gelder, T., Eds.), pp. 1–43. Cambridge, MA: MIT Press.
Vartanian, O., & Goel, V. (2005). Neural correlates of creative cognition. In *Evolutionary and Neurocognitive Approaches to the Arts* (Martindale, C., Locher, P., & Petrov, V.M., Eds.), pp. 195–207. Amityville, NY: Baywood Publishing.
Von Foerster, H. (1984). *Observing Systems*. Salinas, CA: Intersystems Publications.
Wells, C. (1994). *Sketches: Outline with References*. Accessed at http://www.case.edu/artsci/math/wells/pub/pdf/sketch.pdf
Wolfram, S. (1994). *Cellular Automata and Complexity*. Reading, MA: Addison–Wesley.
Yoshikawa, H. (1981). General design theory and CAD system. In *Man–Machine Communication in CAD/CAM* (Sata, T., & Warman, E., Eds.), pp. 35–58. Amsterdam: North-Holland.
Zamenopoulos, T. (2008). *Design out of complexity: a mathematical theory of design as a universal property of organization*. PhD Thesis. University College London, University of London.
Zamenopoulos, T., & Alexiou, K. (2005*a*). Linking design and complexity: a review. *Proc. ECCS 2005 Satellite Workshop Embracing Complexity in Design*, pp. 91–102. Paris, November 17, 2005, Open University.
Zamenopoulos, T., & Alexiou, K. (2005*b*). The problem of design in complexity research. *Proc. European Conf. Complex Systems (ECCS'05)*, pp. 137–138, Paris.
Zamenopoulos, T., & Alexiou, K. (2007*a*). Rethinking the cybernetic basis of design: the concepts of control and organization. *Kybernetes 36(9/10)*, 1570–1589.
Zamenopoulos, T., & Alexiou, K. (2007*b*). Towards an anticipatory view of design. *Design Studies 28(4)*, 411–436.
Zamenopoulos, T., & Alexiou, K. (2011). A complexity theoretic view of cities as artefacts of design intentionality. In *Complexity Theories of Cities Have Come of Age* (Portugali, J., & Meyer, H., Eds.). New York: Springer.

Theodore Zamenopoulos is a Lecturer in design at Open University. He attained a PhD in complexity science and design from University College London. Dr. Zamenopoulos has trained and worked as an architect. His research includes empirical studies of the neurological correlates of design cognition, as well as mathematical and computational models of mental and social processes involved in design tasks.

https://doi.org/10.1017/S0890060411000023 Published online by Cambridge University Press