Behavior of a Technical Artifact: An Ontological Perspective in Engineering

Stefano BORGO [a,1], Massimiliano CARRARA [b], Pieter E. VERMAAS [c] and Pawel GARBACZ [d]

[a] *Laboratory for Applied Ontology, ISTC-CNR, Italy*
[b] *Department of Philosophy, University of Padua, Italy*
[c] *Department of Philosophy, Delft University of Technology, The Netherlands*
[d] *Catholic University of Lublin, Poland*

Abstract. The term 'behavior' is used ubiquitously in engineering. It refers roughly to the way technical artifacts 'behave' in a given or hypothetical situation, and plays a pivotal role in specific design methodologies since it allows connecting descriptions of the physical structure of technical artifacts to descriptions of their technical functions. However, behavior does not have a precise meaning: engineers use the term loosely and when attempting to pinpoint it, end up with incompatible characterizations. Here we formalize the different notions underlying the engineering usage by providing a uniform framework in which they can be related. This framework lays also a conceptual basis for a precise characterization of the notion of technical function in engineering. Our approach develops within the DOLCE ontology and introduces behavior as a new type of individual quality that relates a technical artifact to the event to which it participates. Starting with this assumption, one can distinguish actual, possible and general behaviors of an artifact (token). We add a few more definitions to capture more specific aspects and show their role in capturing engineering usage.

Keywords. Technical artifact, Behavior, Function, Formal Ontology, Engineering.

Introduction

Despite the central importance of *behavior* and *function* to design, it is commonplace that these terms are used in the engineering literature on engineering with various and possibly conflicting meanings. This fact mirrors the differences in the use of these terms in natural language. Talking of behavior, for instance, we point prima facie to whatever people, organisms or artificial (e.g., mechanical) systems do. But at a closer look one can distinguish, at least, four types of behavior [9]: physiological reactions and responses (e.g., salivation, increase in the blood pressure), bodily motions (e.g., walking, the motion of a robot), actions involving bodily motions (e.g., going shopping, telephoning your mother), and actions not involving overt bodily motions (e.g., reasoning, calculating).

[1]Corresponding Author: Stefano Borgo, Laboratory for Applied Ontology, ISTC-CNR, Italy; E-mail: borgo@loa-cnr.it.

In engineering the term behavior is used with less diverse meanings, and there is consensus that it plays a pivotal role in a number of design methodologies by allowing engineers to connect descriptions of the physical structure of technical artifacts to descriptions of their technical functions: structural descriptions should furnish a ground for behavioral descriptions, which, in turn, provide the grounds for functional and finally purposive descriptions. But notwithstanding its central role, engineers and design methodologists often use behavior without a precise meaning. The terms behavior and function are used mostly in a loose and informal way; the attempts to pinpoint their meanings highlighted their incompatible aspects. Chandrasekaran and Josephson [3] distinguish, for instance, five different meanings for the term 'behavior' (see section 2), and Chittaro and Kumar [4] mention three main approaches towards capturing functions.

Our work stems from the framework of ontological research [14] and aims at a uniform and precise characterization of the term 'behavior' to relate the different engineering meanings and to lay a conceptual basis for the notion of technical function. Our approach develops within the framework of DOLCE [11] and consists of introducing behavior as a new type of individual quality that relates an artifact (token) to the event to which it participates. Given this approach, we show that the notions we propose are well motivated and enough to formalize and relate the different engineering meanings.

1. Behavior: Some Philosophical Insights

Informally 'behavior' means whatever (publicly observable) people, organisms or artificial systems do. Some constraints of this characterization are usually adopted in philosophy. If public observablility is central, only physiological and bodily motions that are "overt" and "external" are to count as a behavior. Thus, reasoning and calculating are ruled out. Furthermore, 'doing' is to be distinguished from 'having something done'. If you grasp my arm and pull it up, the rising of my arm is not a behavior of mine. In general, it suffices that the doing is caused by some occurrence internal to the behaving system so that the motion of a robot that goes toward a table and picks out an apple is considered part of its behavior whether or not the robot 'knows' what it is doing [6]. This latter constraint corresponds to the philosophical distinction between physical and agentive behavior. With "physical behavior" we refer to any merely physical action or passion of the body; an "agential behavior" (or "behavior proper") implies a relationship with the mind. Agential behavior entails physical behavior, but not vice versa, e.g., a reflex knee jerk (physical behavior) is not agential.

Philosophical discussion about behavior is sometime associated with "behaviorism." Three main conceptions of behavior are usually distinguished in this viewpoint. Behavior is alternatively (*a*) a physical event, (*b*) a disposition, (*c*) a process.

Physical event. In this approach, behavior is simply an observable physical event or a physical state. There is no commitment to agents nor to the existence of psychological events over and above behavioral events or states.

Dispositions. Here the focus is on agential behavior. In this view behavior is something an agent does in terms of its "dispositions". One can easily specify this view to technical artifacts: the behavior of a technical artifact is the collection of its (physical) dispositions [13]. It remains open if dispositions should be analysed in terms of events, qualitive, or non qualitative properties. A qualitative property is a property that can be perceived.

Process. In this case, behavior is a process having, in the case of physical body, bodily movement as parts. It consists of a complex internal state causing a movement of the agent's body. If a movement M is brought by an internal cause C, the behavior is a process "C causing M" (cf. [6]). Behavior is so distinct from mere output in that it requires a form of explanation. Consider this analogy: an amplifier produces output and we can study such output independently of the machine itself and of the input. Note that the output itself does not tell us what the amplifier does. Therefore the behavior (C causing M) must be distinguished from the causes of the output. We may know what causes the output without knowing what causes behavior. We design, manufacture, and install the machine to do what we want it to do, i.e., we cause some C in the machine to bring about M in the desired circumstances.

2. Behavior in Engineering

In the domain of engineering design methodology 'behavior' (applied to technical artifacts) plays an important though subsidiary role. The term of 'function' is ultimately more important (and problematic): a description of the functions of a technical artifact captures how the artifact is related to the uses for which it is employed and designed; and the important and creative initial phase of designing in which engineers survey solutions to customer needs (without being immediately concerned with the nitty-gritty of the physics of the artifact to be designed), is usually taken as a phase in which engineers reason in purely functional terms. In attempts to fix the meaning of 'function', the concept 'behavior' of technical artifacts enters the central stage. Roughly speaking two approaches to function analysis can be distinguished in engineering [2]: in the first, one relates functions of artifacts to behaviors of artifacts, and then relates these behaviors to structural-physical descriptions of the artifacts; in the second approach, one models functions of artifacts in terms of inputs and outputs, and then relates these functions directly to structural-physical descriptions of artifacts (e.g., [18]). The first approach thus grants a pivotal conceptual role to the term 'behavior', and suggests a clear ontological ordering: technical artifacts have their physical structure; this structure, in interaction with a physical environment, gives rise to the artifacts' behaviors; and these behaviors then determine in some way the artifacts' functions. We call the first approach *FBS*, adopting this acronym from Gero [7], who proposed a Function-Behavior-Structure model of designing, and from Umeda et al. [16], who proposed Function-Behavior-State models of designing. The second approach is known as the *Functional Modelling* approach [12,15]. In this case behavior seems to be side-stepped and the relation between the physical structure of technical artifacts and their functions be established directly on the basis of ontologically more opaque phenomenological engineering hands-on experience as captured by technical handbooks.

With our project of giving an ontological account of behavior of technical artifacts, we position ourselves within the first *FBS* approach towards analysing functions. Ultimately this project should amount to an ontological account of artifact functions, and a first step in this direction is taken towards the end of paper. But, taking the ontological ordering suggested by the *FBS* approach seriously, we start by focussing on a solid account of behavior, built on the basis of existing ontological accounts of the physical realm.

Despite the importance of the notion of behavior in the *FBS* approach, consensus about what counts as 'behavior' has not been reached. Different authors give different characterizations, as we will see, and these are not consistent even in the works of a single author [5]. We will not attempt to survey all meanings proposed, but introduce a few examples and then focus on the survey in [3].

Starting with the two mentioned *FBS* models, [8] defines the behavior of a technical artifact (a designed object), or more precisely, its behavior variables, as describing "the attributes that are derived or expected to be derived from the structure (S) variables of the object, i.e., what it does", where the structural variables "describe the components of the object and their relationships, i.e., what it is." And they give as examples of behavioral variables of a window for instance 'thermal conduction' and 'light transmission' [8].

In [17] behavior is defined as "a transition of states along time" where "a state is represented by entities, their attributes, and their structure"; a given example of behavior is the 'electrical charging' of a drum in a photocopier.

Chandrasekaran and Josephson [3] explore the different meanings engineers attach to the term behavior and isolates five of them, which are characterised with the help of the primitive notion of state variable. These five meanings are:

1) behavior as the value of some state variable of the artifact or a relation between such values at a particular instant (example: the car rattled when the driver hit the curve).
2) behavior as the value of a property of the artifact or a relation between such values (example: a lintel distributes the load to two sides).
3) behavior as the value of some state variable of the artifact over an interval of time (example: the BHP[2] increased for a while, but then started decreasing).
4) behavior as the value of some output state variable of the artifact at a particular instant or over an interval (example: the amplifier is behaving well – the output voltage is constant).
5) behavior as the values of all the described state variables of the artifact at a particular instant or over an interval (no example is given)

Notice that for all five meanings, a behavior of a technical artifact is partially objective and partially subjective. It has an objective aspect because it eventually depends on the properties or features of the artifact. Still, the very same behavior has a subjective aspect: it depends on the designer(s) and, indirectly, on engineering practice for the choice of the state variables. The underlying intuition of Chandrasekaran and Josephson for this subjective twist, is that a state variable of an artifact represents some feature or aspect of this artifact that might be relevant only from a specific point of view [1]. Yet, in agreement with the observations of section 1, it is important to emphasise that the behavior of a technical artifact is different from the value of its state variable(s). The behavior is somehow characterised by this value(s) in a sense to be explicated. Thus, if the value of the input voltage of an amplifier is 10 mV at a particular instant, then this value (10 mV) is not identical with any behavior of this amplifier. Rather, part of the behavior of the amplifier consists of the actual situation that its input voltage is 10 mV.

[2] BHP stands for Brake Horse Power and it is described as the amount of real horsepower going to the pump.

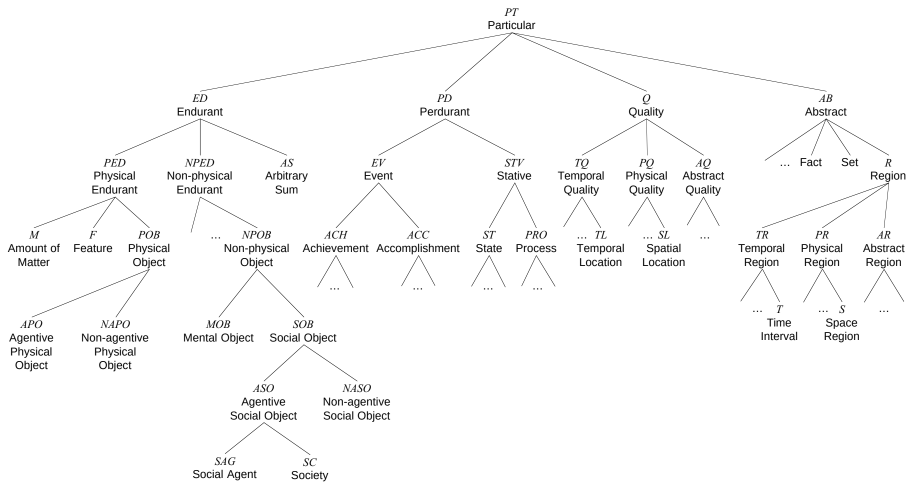

Figure 1. DOLCE basic categories (taken from [11].)

3. Scope and Adopted Framework

Our goal is to formalize a notion of *artifact behavior* that, on the one hand, captures the informal meaning this term has in engineering practice and, on the other, is ontologically motivated. In this work, we take the meanings of section 2 at face value and neither discuss their adequacy nor try to reduce or modify them. Once the characterization is posited, we will see that the relationships among these meanings becomes clear and formally expressible. The analysis is carried out within the framework of the DOLCE ontology [11] and the formal system we obtain can be considered an extension of that ontology. Our final aim is to incorporate both the notions of engineering behavior and function in DOLCE. The second notion will be only sketched in this paper by considering the notion of function as presented in [3].

The basic idea we follow to formalize behavior is to consider it as a relationship between the artifact and the event to which it participates. Since events can be actual (in the real world) or just possible, we distinguish *possible* behavior from *actual* behavior. Due to the crucial role of the notion of behavior at the design phase and at the description level, where the behavior of an artifact is (to some extent) speculative, we need to consider also events that are *believed* to be possible. To deal with these distinctions, we introduce and relate several classes of behaviors. Once these classes are formally described, one can talk of behavior according to a group of agents (typically, engineers) and of behavior descriptions. Later, we will see how this framework allows us to capture and make explicit the meanings listed in section 2.

In what remains of this section, we briefly introduce the categories of DOLCE that are relevant to our work. In parenthesis we report the formal names as used in the next sections. The full taxonomy is reported in Figure 1, while a detailed discussion of DOLCE and its categories can be found in [11]. We begin with the general notion of “endurant” (ED). This category collects those entities that are *wholly* present at any time they are present, for instance a car or an amount of water. Entities that take up time (a football game), that are qualities (the color of my car), or that are abstracts (the set of pens in my office) are not endurants. In DOLCE the term 'object' is used to capture a notion of wholeness and the them 'physical' to indicate entities that are located in space-time. In the next sections, we will concentrate on "non-agentive physical objects" (NAPO) which are physical objects to which one cannot ascribe intentions, beliefs or desires (like a tool or a notebook). Temporal entities like a thunderstorm, are called "perdurants" (PD). Informally, a perdurant is an entity that is only partially present at any time it is present. In this category we find football games as well as artifact productions. These entities have temporal parts (like the first half of the game) as well as spatial parts (the event restricted to a half of the football field during the game). Some perdurants, e.g., sitting, moving and drilling, are called "stative" (ST). They have two important properties: the sum of two stative perdurants is also stative and any temporal part of a stative perdurant is stative as well. A different type of entities find place in the "individual quality" category (Q). These can be seen as instantiations of basic aspects of endurants or perdurants (usually qualities can be perceived or measured like weight or energy). The term 'individual' is used to mark the essential relationship between the entity and its own qualities. Note that qualities are particulars, they should not be confused with properties (universals).

We will make use of several relationships among which *parthood*: "x is part of y", written P(x,y). In DOLCE the *parthood* relation applies to pairs of endurants (e.g., to state that an object is part of another) as well as to pairs of perdurants (to state that an event is part of another). For instance, if a = 'writing paper ABC' and b = 'writing the introduction of paper ABC', then P(b,a) holds. The main relation involving both endurants and perdurants is called *participation*, formally PC. This relation captures the simple fact that an endurant 'lives' in time by participating in some perdurant. For example, a person (endurant) may participate in a discussion (perdurant). A person's life is also a perdurant, in which a person participates throughout all its duration (the time spanned by the life). If endurant a participates in perdurant e at each instant of period t, we write PC(a,e,t) which reads "a participates in e during all of t" (here t may be a sub-interval of the duration of e). Note that participation and parthood are distinct relations. An endurant is never part of a perdurant, only perdurants can be parts of perdurants. Also, participation is time-indexed in order to account for the varieties of participation in time (temporary participation, constant participation, etc.) Qualities are associated to quality spaces (where comparisons like "a and b have different weight" are carried out) and the position an individual quality has in a space is called a *quale*. We write ql(r,q,t) to mean "r is the quale of the quality q during time t" and drop the temporal parameter for perdurant's qualities. (Our use of ql in section 5 is an obvious generalization of this relation). Finally, *sum* (symbol $+$) and *fusion* (symbol σ) are two operators. The sum of two perdurants x and y is written x $+$ y (e.g., closing a window is the sum of moving it and turning the handle), while the fusion is an extension of the binary sum to an arbitrary number of perdurants (e.g., my staying in Paris is the fusion of all the days I spent in Paris).

4. An Ontological Definition of Artifact Behavior

In DOLCE the most specific category for technical artifacts is NAPO. Let a be a non-agentive physical object, i.e., NAPO(a). Informally, the behavior b of a in a perdurant e is *the way* a *exists in* e. Formally, we take b to be a new kind of individual quality: it does not hold for a single endurant or perdurant, but for pairs of an endurant and a perdurant. In this way, it captures the special relationship between the artifact and the event in which it ‘behaves’. To formalize this relationship, we introduce a ternary relation $Beh(\mathtt{a},\mathtt{e},\mathtt{b})$. We talk of the behavior of $\mathtt{a}$ in a perdurant $\mathtt{e}$ if and only if $\mathtt{a}$ participates in $\mathtt{e}$ for *all the duration of* $\mathtt{e}$ and in the logic we impose that, for a pair endurant-perdurant satisfying this condition, the corresponding behavior $\mathtt{b}$ exists. A behavior $\mathtt{b}$ is uniquely identified by the pair endurant-perdurant, although a perdurant $\mathtt{e}$ may have several participants and an entity $\mathtt{a}$ may participate in several perdurants.

Given a perdurant $\mathtt{e}$, let us write $\mathtt{tm}(\mathtt{e})$ for the whole period of time during which $\mathtt{e}$ exists. Then, we formally capture the above observations with the following axioms

$$Beh(\mathtt{a},\mathtt{e},\mathtt{b}) \rightarrow \mathtt{PC}(\mathtt{a},\mathtt{e},\mathtt{tm}(\mathtt{e}))$$

$$Beh(\mathtt{a},\mathtt{e},\mathtt{b}) \wedge Beh(\mathtt{a},\mathtt{e},\mathtt{b'}) \rightarrow \mathtt{b} = \mathtt{b'}$$

$$Beh(\mathtt{a},\mathtt{e},\mathtt{b}) \wedge Beh(\mathtt{a'},\mathtt{e'},\mathtt{b}) \rightarrow \mathtt{a} = \mathtt{a'} \wedge \mathtt{e} = \mathtt{e'}$$

$$\mathtt{NAPO}(\mathtt{a}) \wedge \mathtt{PC}(\mathtt{a},\mathtt{e},\mathtt{tm}(\mathtt{e})) \rightarrow \exists \mathtt{b}\ Beh(\mathtt{a},\mathtt{e},\mathtt{b})$$

Looking at engineers activity, we need to distinguish different kinds of behavior. While an ‘actual behavior’ of the artifact is what it actually does during (a part of) its life, the more general notion of ‘possible behavior’ deals with what the artifact can possibly do. A pen may be destroyed before it happens to write, still the pen could have participated to a writing event, i.e., writing is part of its behavior although not of its ‘actual’ behavior. Furthermore, although a pen may not possibly write due to a design flaw, still engineers (not aware of the flaw) talk about its writing behavior. Now we see how to make room for these cases in the formalism.

Perdurants in DOLCE can be divided in **actual** (those that happen in the ‘actual world’, no matter when) and **possible** (those happening in a ‘possible world’, including the actual one). ‘Impossible’ perdurants, related to (designed or even constructed) artifacts that cannot perform the functionalities the designers attribute to them, are not directly isolated in DOLCE since there is no explicit constraint to contingent laws (like those of physics). Nonetheless, within the engineer domain, these perdurants are only thought but cannot possibly happen. For this reason, we assume that $\mathtt{PD}$ in the engineering domain collects all the physically possible perdurants, and introduce the more general category of **generalised** perdurants to account for all (actual, possible and impossible) perdurants with a proviso: a generalized perdurant cannot be describe by an inconsistency, that is, it must be believable by rational agents (namely, engineers). We write $\mathtt{APD}$ and $\mathtt{GPD}$ for the class of actual and generalised perdurants respectively. Then, $\mathtt{APD} \subseteq \mathtt{PD} \subseteq \mathtt{GPD}$. We can now constrain the domain of *Beh*

$$Beh(\mathtt{a},\mathtt{e},\mathtt{b}) \rightarrow \mathtt{NAPO}(\mathtt{a}) \wedge \mathtt{GPD}(\mathtt{e}) \wedge Q(\mathtt{b})$$

Some combinations of perdurants are meaningless in engineering practice. It is possible that a pen participates in a writing perdurant now (say in this world) and it is also possible that the same pen participates to a non-writing perdurant now (in another world). However, these two perdurants cannot ‘physically happen’ at the *same time* in the *same world*. If we sum these perdurants, the very same pen should be writing and not writing.[3]

[3] In principle this event could be in $\mathtt{GPD}$. However, an engineer (as such) may believe wrongly that something happen or could happen, but not that something, patently known to defy physical laws, can happen.

To account for this condition, we say that two or more perdurants are **coherent** when their sum is physically meaningful and, in this case, we write $\mathsf{a} +_c \mathsf{b}$. (A similar condition is set to restrict the fusion operator.) This condition depends on the laws of physics, which we assume are given. Note that actual perdurants are always coherent, i.e., their sum (fusion) is always defined.

We say that two perdurants are **co-temporal** if they occur at the same time period, e.g., "my attending Robert's talk today" and "Robert giving a talk today" are co-temporal.

Let $\mathcal{C}_{\mathsf{a}}$ be the (non-empty) class of all generalized perdurants e such that $\mathsf{PC}(\mathsf{a}, \mathsf{e}, \mathtt{tm}(\mathsf{e}))$ for some fixed a. Fix a perdurant e in it. We say that e is *minimal* in $\mathcal{C}_{\mathsf{a}}$ if for each perdurant e' in $\mathcal{C}_{\mathsf{a}}$ co-temporal with e, not $\mathsf{PP}(\mathsf{e}', \mathsf{e})$. We use this notion and the previous classification of perdurants to specialize the DOLCE relation $\mathsf{lf}(\mathsf{x}, \mathsf{a})$, which reads "perdurant x is the life of endurant a" (for lack of space, we do not report the formal expressions):

(a) $\mathsf{Alf}(\mathsf{x}, \mathsf{a})$ stands for "perdurant x is the actual life of endurant a" and is formally defined as the fusion of the actual perdurants which are spatially minimal in $\mathcal{C}_{\mathsf{a}}$;

(b) $\mathsf{Plf}(\mathsf{x}, \mathsf{a})$ stands for "perdurant x is a possible life of endurant a" and is formally defined as the fusion of a maximal class (wrt inclusion) of coherent possible perdurants which are minimal in $\mathcal{C}_{\mathsf{a}}$;

(c) $\mathsf{Glf}(\mathsf{x}, \mathsf{a})$ stands for "perdurant x is the general life of endurant a" and is formally defined as the fusion of the general perdurants which are spatially minimal in $\mathcal{C}_{\mathsf{a}}$.

We write $\mathsf{Alf}(\mathsf{a})$ to denote the derived term "the actual life of a". Similarly for $\mathsf{Glf}(\mathsf{a})$.

Definition 1 *(Actual, Possible, Impossible, and General Behavior)*

An actual behavior *of* a *is a* b *such that* $Beh(\mathsf{a},\mathsf{e},\mathsf{b})$ *for some* e *with* $\mathtt{APD}(\mathsf{e})$.
The maximal actual behavior *of* a *is the actual behavior* b *for which* $Beh(\mathsf{a}, \mathsf{Alf}(\mathsf{a}),\mathsf{b})$.

A possible behavior *of* a *is a* b *such that* $Beh(\mathsf{a},\mathsf{e},\mathsf{b})$ *for some* e *with* $\mathtt{PD}(\mathsf{e})$.
A maximal possible behavior *of* a *is a* b *for which* $Beh(\mathsf{a},\mathsf{e},\mathsf{b})$ *for some possible life* e *of* a, *i.e.,* $\mathsf{Plf}(\mathsf{e},\mathsf{a})$.

An impossible behavior *of* a *is a* b *such that* $Beh(\mathsf{a},\mathsf{e},\mathsf{b})$ *for some* e *with* $\mathtt{GPD}(\mathsf{e})$ *and not* $\mathtt{PD}(\mathsf{e})$.

A general behavior *of* a *is a* b *such that* $Beh(\mathsf{a},\mathsf{e},\mathsf{b})$ *for some* e *with* $\mathtt{GPD}(\mathsf{e})$.
The maximal general behavior *of* a *is the general behavior* b *for which* $Beh(\mathsf{a}, \mathsf{Glf}(\mathsf{a}),\mathsf{b})$.

Given an agent or group of agents $\mathcal{G}$, we write $\mathtt{PD}_{\mathcal{G}}(\mathsf{e})$ to state that $\mathcal{G}$ believes that e is a possible perdurant.[4] A class $\mathcal{C}$ of general behaviors of a is said to be *coherent* if b, $\mathsf{b}' \in \mathcal{C}$ with $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$, $Beh(\mathsf{a}, \mathsf{e}', \mathsf{b}')$ implies that $\mathsf{e} +_c \mathsf{e}'$ is defined. We use these two notions to define the behavior of a technical artifact *for a group* $\mathcal{G}$.

Definition 2 *($\mathcal{G}$-behavior) Let $\mathcal{G}$ be an agent or group of agents.*

A $\mathcal{G}$-behavior *of* a *is a general behavior* b *such that if* $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$ *then* $\mathtt{PD}_{\mathcal{G}}(\mathsf{e})$.

A maximal $\mathcal{G}$-behavior *of* a *is a general behavior* b *such that* $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$ *where* e *is the fusion of all the perdurants in a coherent class* $\mathcal{C}$ *of* $\mathcal{G}$*-behaviors of* a *and* $\mathcal{C}$ *is maximal with respect to inclusion.*

[4] Generally speaking, this notion should include a temporal parameter since agents may change their beliefs in time. For simplicity, we disregard this issue here.

A coherent class of $\mathcal{G}$-behaviors is a collection of $\mathcal{G}$-behaviors whose perdurants are coherent. This reflects the fact that a group of engineers may believe something which is impossible (due to construction or design flaws). Nonetheless, they do not believe that incompatible things happen.

Finally, the standard distinction between device-centric and environmental-centric views of a technical artifact [3] drives the following distinction of artifact behavior.

Definition 3 *(Internal and Environmental Actual Behavior)*

An internal actual behavior *of* a *is a behavior* b *such that if* $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$ *holds, then* $\mathsf{P}(\mathsf{e}, \mathsf{Alf}(\mathsf{a}))$.

An environmental actual behavior *of* a *is a behavior* b *such that if* $Beh(\mathsf{a},\mathsf{e},\mathsf{b})$ *then* $\mathtt{APD}(\mathsf{e})$ *and not* $\mathsf{P}(\mathsf{e}, \mathsf{Alf}(\mathsf{a}))$.

5. Artifact Behavior In DOLCE

Now we use the above ontological framework to formalize the five meanings of section 2. Here state variables are indirectly captured through the given behavior b. Also, following the discussion in [3], we implicitly take $\mathcal{G}$ to be the designer(s) of the technical artifact a, although this is not relevant.

1. The behavior b of an artifact a from an engineering perspective of an agent $\mathcal{G}$ is a $\mathcal{G}$-behavior such that if $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$, then $\mathtt{tm}(\mathsf{e})$ is an instant (a period of length zero). Example: let e = "the car rattled when the driver hit the curve", then $Beh(\text{car}, \mathsf{e}, \mathsf{b}) \wedge \mathtt{PD}_{\mathcal{G}}(\mathsf{e}) \wedge |\mathtt{tm}(\mathsf{e})| = 0$
2. The behavior b of an artifact a from an engineering perspective of an agent $\mathcal{G}$ is a $\mathcal{G}$-behavior that belongs to the class of environmental actual behaviors where $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$ implies $\mathtt{ST}(\mathsf{e})$. Example: let e = "the lintel's distributing the load to two sides", then $Beh(\text{lintel}, \mathsf{e}, \mathsf{b}) \wedge \mathtt{PD}_{\mathcal{G}}(\mathsf{e}) \wedge \mathtt{ST}(\mathsf{e}) \wedge \mathsf{P}(\mathsf{e}, \mathsf{Alf}(\mathsf{a}))$
3. The behavior b of an artifact a from an engineering perspective of an agent $\mathcal{G}$ is a $\mathcal{G}$-behavior such that if $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$, then $\mathtt{tm}(\mathsf{e})$ is a period of positive length. Example: let e = "the increasing of the BHP of the artifact for a while and its decreasing afterwards", then $Beh(\text{artifact}, \mathsf{e}, \mathsf{b}) \wedge \mathtt{PD}_{\mathcal{G}}(\mathsf{e}) \wedge |\mathtt{tm}(\mathsf{e})| > 0$
4. The behavior b of an artifact a from an engineering perspective of an agent $\mathcal{G}$ is a $\mathcal{G}$-behavior such that if $Beh(\mathsf{a}, \mathsf{e}, \mathsf{b})$ and $\mathtt{r}$ is the quale of b during e, i.e., $ql(\mathtt{r}, \mathsf{b}, \mathtt{tm}(\mathsf{e}))$, then $\mathtt{r}$ satisfies the given condition during $\mathtt{tm}(\mathsf{e})$ (i.e., it has some given value, it is constant, etc.) Example: let e = "the amplifier performing its function", then $Beh(\text{amplifier}, \mathsf{e}, \mathsf{b}) \wedge \mathtt{PD}_{\mathcal{G}}(\mathsf{e}) \wedge \forall \mathtt{t},\mathtt{t}' \; (\mathsf{P}(\mathtt{t}, \mathtt{tm}(\mathsf{e})) \wedge \mathsf{P}(\mathtt{t}', \mathtt{tm}(\mathsf{e})) \rightarrow \forall \mathtt{r},\mathtt{r}' \; (ql(\mathtt{r}, \mathsf{b}, \mathtt{t}) \wedge ql(\mathtt{r}', \mathsf{b}, \mathtt{t}') \rightarrow \mathtt{r} = \mathtt{r}')$
5. The behavior b of an artifact a from an engineering perspective of an agent $\mathcal{G}$ is a maximal $\mathcal{G}$-behavior of a (restricted to the given period of time). Example: [no example given]

Admittedly, the formalization of the examples from [3] seems not very informative of our framework and its advantages. For this reason, here we formalize another example taken from [6]. Consider a mechanical thermostat in a room and suppose that the room temperature drops to 17 C$^\circ$. The thermostat responds turning the furnace on. This event characterizes a behavior of the thermostat: a falling of the room's temperature causes a bimetal strip in the thermostat to bend. When the bimetal strip bends to some angle A (here 17 C°), it closes an electrical circuit which connects the furnace and the furnace ignites. The event sequence could be illustrated in the following way

(I) Temperature drops to 17 C°, (II) Strip bends to angle A, (III) Switch closes,
(IV) Current flows to furnace, (V) Furnace ignites.

> "The thermostat's behavior – Dretske observes – is the bringing about of furnace ignition by events occuring in the thermosthat – in this case the closure of a switch by the movement of a temperature-sensitive strip." ([6], p.86)

In our framework we represent the thermostat behavior as Beh(thermostat, e, b) where b is the general behavior of the thermostat for the event e corresponding to the sequence (I)–(V) above. If we want to model the behavior for a sub-event e', we write Beh(thermostat, e', b') where e' is, say, event (III). Instead, the behavior of the switch itself at e' is introduced by writing Beh(switch, e', b").[5]

6. From Artifact Behaviors to Artifact Functions

Besides distinguishing different meanings of behavior, Chandrasekaran and Josephson [3] define the notion of artifact function. They presuppose a theoretical perspective in which artifact functions are construed as intended behaviors and define two concepts: device-centric function and environment-centric function. We will show in this section to what extent the ontological approach outlined above is suitable for grasping these concepts. (Note that this section wants to be as close as possible to the approach of Chandrasekaran and Josephson. Some simplifications and improvements are possibles and may make this part clearer. However, since the formalization here is tentative, we prefer to stay close to the source we consider.)

Let X be a set of non-agentive physical objects, i.e., $a \in X \rightarrow NAPO(a)$. It is said in [3] that a behavioral constraint in X is any constraint on the behaviors of the elements of X. As the examples given in [3] suggest, a behavioral constraint may be absolute, i.e., unconditional (e.g., that the value of output voltage is greater than 5 volts), or conditional (e.g., that if the input voltage is above 5 volts, the output voltage is a sinusoid.)

Following the argument in [3], we say that a behavioral constraint in X, written $Constr_{Beh}(X)$, is a set of pairs $<b_0,b_1>$, where b_0 and b_1 are behaviors of $a \in X$ such that b_0 is a condition of b_1. Write the latter as $Cond(b_0,b_1)$, then:

$$<b_0,b_1> \in Constr_{Beh}(X) \text{ iff}$$
$$\exists a \in X, e_0, e_1\ (Beh(a, e_0, b_0) \wedge Beh(a, e_1, b_1) \wedge Cond(b_0, b_1))$$

where pair $<b_0,b_0>$ is used to represent the unconditional constraint b_0. We will say that a behavioral constraint $Constr_{Beh}(X)$ is satisfied iff for $<b_0,b_1> \in Constr_{Beh}(X)$

– if $b_0 = b_1$, then b_1 is an actual behavior,

– if $b_0 \neq b_1$ and b_0 is an actual behavior, then b_1 is an actual behavior.

We are now in a position to grasp the device-centric notion of function defined in [3]:

> "Let F be a set of behavioral constraints defined on, and satisfied by, an object D. If F is intended or desired by an agent A, then D has function F for A." ([3], p.172)

[5] Note that we have not characterized the relationship between behaviors b, b', and b".

In this specific sense, a function of a technical artifact a for an agent $\mathcal{G}$ is a behavioral constraint in {a} provided that this behavioral constraint is satisfied and is desired by $\mathcal{G}$.

To define the environment-centric notion of function we need to introduce the notion of mode of deployment. By Chandrasekaran and Josephson's explanation, a mode of deployment for an artifact a consist of what they call "the specifications of the ways in the causal interactions" between a and some objects from its environment. More perspicuously speaking, a mode of deployment for an artifact a consists of the structural relations between a and these objects and the actions in which a and the objects are involved. We thus represent modes of deployment by means of perdurants. Let X be a set of physical objects. We will say that a mode of deployment for an artifact a in an environment X, written MD(a,X), is such set of generalized perdurants that for any of its element e, there exists $a_1 \in X$ such that $a \neq a_1$ and both a and a_1 participate in e. A mode of deployment MD(a,X) is said to be feasible iff all of its elements are coherent.

This is how Chandrasekaran and Josephson define environment-centric functions:

> Let F be a set of behavioral constraints that an agent, say A, desires or intends to be satisfied in some W [i.e., in some world W]. Let D be an object introduced into W, in a mode of deployment M(D, W). If D causes F to be satisfied in W, we say that D has, or performs, the function F in W. ([3], p. 171)

Here is our proposal. First, although one may represent the causal relations at stake in various ways, we will employ the theory of causality proposed in [10]. For our purposes it is important to note that by the lights of this account, the relation of causality relates perdurants. Thus, instead of saying that a technical artifact in a certain mode of deployment causes a behavioral constraint to be satisfied, we will say that a mode of deployment in which an artifact is involved causes a behavioral constraint to be satisfied.

Now let $Constr_{Beh}$(X) be a behavioral constraint in X and MD(a,X) be a feasible mode of deployment for an artifact a in X. We will say that a mode of deployment MD(a,X) causes a behavioral constraint $Constr_{Beh}$(X) iff

- all perdurants from MD(a,X) are actual,
- $Constr_{Beh}$(X) is satisfied,
- given $<b, b> \in Constr_{Beh}(X)$, if *Beh*(a, e, b), then there exists such perdurant $e_1 \in$ MD(a,X) that e_1 causes e,
- given $<b_0, b> \in Constr_{Beh}(X)$, if *Beh*($a_0$,$e_0$,$b_0$), *Beh*(a, e, b), and $e_0 \in$ MD(a,X), then there exists such perdurant $e_1 \in$ MD(a,X) that e_1 causes e.

Then, a behavioral constraint $Constr_{Beh}$(X) is said to be a function of a technical artifact a in X relative to a mode of deployment MD(a,X) if MD(a,X) causes $Constr_{Beh}$(X) and $Constr_{Beh}$(X) is desired by some agent $\mathcal{G}$.

7. Results, Limits, and Future Work

We addressed the ubiquitous character of the engineering concept of the behavior of technical artifacts by providing an ontological characterization of the meanings identified in the *FBS* approach [3]. This is the first comprehensive formalization of artifact behavior that has been proposed and we showed in section 5 that it succeeds to capture and relate the given meanings. This clarifies that an ontological and uniform characterization of artifact behavior is possible.

Further extensions are needed to formalize related notions like that ‘behaviors of artifact-types’ (as used in design) and ‘behavior-types’. We will anlyse them in the future. Moreover, we want to study the constraints between the behavior of an endurant and those of its components, and the constraints between the behavior of an endurant in an event and in its sub-events. Finally, we need to better study the engineering notion of artifact function going beyond the characterization in [3].

Acknowledgements

We thank the anonymous reviewers for their comments. Stefano Borgo was partially supported by the Provincia Autonoma di Trento (PAT), Pawel Garbacz by the Marie Curie Intra-European Fellowship (EIF-006550), and Pieter Vermaas by the Netherlands Organization for Scientific Research (NWO).

References

[1] B. Chandrasekaran. Functional representation and causal processes. In *Advances in Computers*, pages 73–143. Academic Press, 1994.

[2] B. Chandrasekaran. Representing function. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 19:65–74, 2005.

[3] B. Chandrasekaran and J. R. Josephson. Function in device representation. *Engineering with Computers*, 16(3/4):162–177, 2000.

[4] Luca Chittaro and Amruth N. Kumar. Reasoning about function and its applications to engineering. *Artificial Intelligence in Engineering*, 12:331–336, 1998.

[5] K. Dorst and P. E. Vermaas. John Gero’s function-behaviour-structure model of designing: A critical analysis. *Research in Engineering Design*, 16:17–26, 2005.

[6] F. Dretske. *Explaining Behavior*. MIT Press, Cambridge, 1988.

[7] J.S. Gero. A knowledge representation schema for design. *AI Magazine*, 11:26–36, 1990.

[8] J.S. Gero and U. Kannengiesser. The situated function-behaviour-structure framework. *Design Studies*, 25:373–391, 2004.

[9] J. Kim. *Philosophy of Mind*. Westview Press, Boulder Colorado, 1998.

[10] J. Lehmann, S. Borgo, C. Masolo, and A. Gangemi. Causality and causation in dolce. In A.C. Varzi and L. Vieu, editors, *Proceedings of the Third International Conference FOIS 2004*, pages 273–284. IOS Press, 2004.

[11] Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, and Alessandro Oltramari. Wonderweb deliverabled18. In *http://www.loa-cnr.it/Papers/D18.pdf*. 2003.

[12] G. Pahl and W. Beitz. *Engineering Design*. Springer Verlag, 1998.

[13] M.A. Rosenman and J.S. Gero. Purpose and function in design. *Design Studies*, 19:161–186, 1998.

[14] Steffen Staab and Rudi Studer. *Handbook of ontologies*. International handbooks on information systems. Springer Verlag, Berlin (DE), 2004.

[15] Robert B. Stone and Kristin Wood. Development of a functional basis for design. *Journal of Mechanical Design*, 122(4):359–276, 2000.

[16] Y. Umeda, M. Ishii, M. Yoshioka, Y. Shimomura, and T. Tomiyama. Supporting conceptual design based on the function-behaviour-state modeler. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 10:275–288, 1996.

[17] Y. Umeda, M. Ishii, M. Yoshioka, Y. Shimomura, and T. Tomiyama. Development of design methodology for upgradable products based on function-behavior-state modeling. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 10:161–182, 2005.

[18] Michael van Wie, Cari R. Bryant, Matt R. Bohm, Daniel McAdams, and Robert B. Stone. A model of function-based representations. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 19:89–111, 2005.