FROM FUNCTIONS TO MEANS
PRACTICAL CONSEQUENCES OF ARTIFACTUAL FUNCTIONS

JESSE HUGHES

Abstract. A functional ascription tells one what an artifact is "for". There is practical content to such ascriptions: they tell one that the artifact can be used to achieve certain goals. This is an essential part of the natural language discussions of function. Consequently, one expects that a functional ascription produces certain practical consequences, and in particular means-end relations. However, in existing theories of artifactual function, it is difficult to see how to infer such consequences from a functional ascription. We provide here a sketch of the relevant features of functional ascriptions that produce sufficient means-end relations. We also relate artifact failure and malfunction to the induced means-end relations.

One of the defining characteristics of artifacts is that they have functions; they are *for* something. Many, perhaps most, non-philosophical discussions of artifacts and their functions have a decidedly practical component. In general, if one knows an artifact's function, she knows that the artifact can be used to realize certain goals. Knowing the function provides one with means to related ends she may wish to pursue. We interpret such means-end claims as suggested by Georg Henrik von Wright [7]: an end is a state of affairs one wants to realize and a means is an action—something to be done—that can bring about the end. We will identify the features of functional ascriptions that allow one to infer related means-end relations. In fact, the inference appears deceptively simple, but the expectations regarding artifact behavior are difficult to capture correctly. Some artifact tokens are defective and incapable of doing what they are supposed to and our expectations must reflect this fact. We address this issue by introducing "normal tokens", a controversial notion that we justify by analogy with informal reasoning about functions.

Furthermore, we claim that the means-end relations induced by functional ascriptions can be used to define functional terms, including artifact failure and malfunction. Thus, our presentation serves two distinct purposes. First, we clarify the practical consequences of functional ascriptions and thus provide a missing component of current theories of function. Second, we use this analysis to show that failure and malfunction can be naturally defined using the means-end relations that are consequences of function ascriptions.

Philosophical theories of function often focus on different aims. Many of these theories are interested in functional explanations rather than practical knowledge. That is, they provide an analysis of artifactual function that explains the presence, production or prevalence of artifacts of a particular kind. This is largely because recent interest in artifactual functions has been a by-product of work on biological functions. Functions in biology are used to explain how natural selection has shaped

organisms or ecosystems and this explanatory role is more important than practical consequences, if any.

Indeed, it is not clear that biological functional ascriptions *do* produce means-end relations in any obvious or systematic way. That hearts are for pumping blood explains why humans have hearts, but it is difficult to see how this function suggests means-end relations. A heart is not usually called a *means* to pumping blood because, in practical reasoning, a means is something which one may *do* or *use* in order to realize an end. One doesn't usually *use* her heart to pump blood. This is what the heart normally does on its own and the heart-bearer pays it little attention (until her heart begins to behave unexpectedly).

Artifactual functions, on the other hand, are closely connected to means-end relations. Often, these functions are relevant for users precisely because they suggest ways of realizing certain goals. Existing theories differ widely in the importance they place on practical consequences, but none of them provide any clear explanation of how one infers means-end relations from functions. Distant historical theories (in Mark Perlman's terminology [6]), including Larry Wright [8] and Ruth Garrett Millikan[1] [4], define function in terms of what the artifact (or its "ancestors") once did. In similar circumstances, the artifact is supposed to perform similarly, but teasing out a clear means-end relationship from this is difficult at best. The goal-contribution account of Christopher Boorse [1] places emphasis on functional goals, but lacks a specific account of how means are derived from previous uses at particular times. More recently, Wybo Houkes and Pieter Vermaas [3] have suggested the ICE theory of artifactual functions. Here, use plans and goals are both essential to their account of function, but nonetheless they give no clear method of inferring means-end relations from functional ascriptions.

We believe that these theories of function can be strengthened by an account of how artifactual functions induce tentative means-end relations. Our account should be more or less compatible with each theory[2], at least to the extent that the theory provides each of the features of function that we discuss below. We intend to be agnostic as far as possible about which theory is the "correct" theory of function. Instead, we will identify the key components of functional ascriptions that allow one to infer means-end relations and discuss the related concepts of artifactual failure and malfunction.

We adopt Houkes and Vermaas's term "use plans", loosely, a description of what the user should do in order to realize an artifact's function. However, we use the term somewhat differently than they do, since we have separated the contexts of use from the use plan for conceptual clarity. Additionally, we are not committed to their claim that all artifactual functions include use plans: an existing retaining wall may have a function, but it is not clear that it comes with a use plan. Nonetheless, those artifacts that induce means-end relations in our sense *do* come with use plans, since

[1]Since Millikan is introducing a new technical meaning of "function" rather than analyzing an existing usage, it is not clear that our criticisms regarding practical consequences apply.

[2]Cummins's theory of functions [2] is a notable exception. He defines function in terms of system capabilities and it is not at all clear that this definition produces means-end relations in the way we describe here.

a means is something that "has to be done"[3]. Thus, we restrict our attention to artifactual functions that include use plans.

The following features of functional ascriptions account for their practical consequences. We also claim that each of these features are natural products of the design process[4], but a full argument for this claim cannot be presented here.

(1) An artifactual type T. We agree with Karen Neander [5] that functional ascriptions apply to artifactual types and to tokens only derivatively[5]. We require that such types are narrow enough to include use plans. A broad type like "bottle opener" can be realized in many different ways and with many different use plans, so is too broad for our needs here. Instead, means-end relations are introduced by narrower subtypes like "corkscrew".
(2) A functional goal φ. This is a condition or state of affairs that one aims to realize by using the artifact as specified by the use plan. Some authors identify the functional goal with the function itself, but we find that confuses issues more than necessary. For example, on Wright's account [8], the function "Z is a function of X" is described alternatively as an activity — X does Z — and a condition — Z is a consequence of X's being there. We believe that this confusing situation can be clarified by separating use from goal.
(3) A use plan α. By this, we mean a prescription for how the artifact is to be used—what manipulations, placements, maintenance, etc. that the user should do so that the artifact token is reasonably expected to realize the functional goal. Such plans may be missing from simple ascriptions (such as, "staplers are for fastening papers"), but even these ascriptions assert that there is *some* use plan for achieving the goal.
(4) A set C of normal contexts of use. This is a description of the expected context in which the artifact will be used. Such descriptions are in practice incomplete, due to the usual thorny issues of practical reasoning and relevance, but we assume that users have some idea of when and where the artifact is intended to be used.

Some of these features are hard to discern in some informal functional ascriptions. We often speak tersely about functions, giving only the type and the functional goal. But such terse ascriptions cannot generate clear means-end relations unless the missing features of use plan and context are assumed to be implicit.

We do not assume that each artifact type has only one function, proper or otherwise.

These features are fairly modest and certainly fall short of characterizing functions. Issues regarding what counts as an artifactual function or how functional ascriptions are justified are outside the scope of our work here. We aim to present an analysis that

[3]Sometimes, one speaks of an object as a means, as when we say that a bridge is a means to crossing a river. We regard such locutions as shorthand for some action involving the object: in this case, walking across the bridge is the means we have in mind.

[4]Here, we interpret "design" loosely enough that applications of accidental functions count as "re-design", as in [3].

[5]We allow that some artifactual types include only a single token—or even no existing tokens at all—to account for ascriptions involving prototypes, unique artifacts and artifacts in the early design process.

is coherent with each of the major theories of artifactual function, whether historical, intentionalist or some combination thereof.

Note as well that features (1)–(4) do not distinguish proper function from accidental. Again, we regard this flexibility as appropriate: both proper and accidental functions induce means-end relations (but we follow the generally accepted tradition that malfunction claims apply only to proper functions).

One can extract means-end relations from just these features. Given any a functional ascription with features T, φ, α and C, we tentatively claim:

(∗) In situations realizing some context in C, one expects that a token t of type T can be used as prescribed by α in order to realize φ. In other terms, in such contexts, $\alpha(t)$ is a means to φ.

Note that the means-end relation here is that of a sufficient means to an end. Functions do not provide necessary means, but instead suggest one way to realize the functional goal.

The step from features (1)–(4) to (∗) seems deceptively simple. The features were chosen just to fit into an appropriate means-end relation, but the difficulty comes in interpreting the expectations expressed in (∗). A functional ascription does not entail that every token of the appropriate type will reliably realize its end in appropriate contexts. Tokens do not always behave as one would like. Sometimes this comes as a surprise, but other times this is predictable. Some tokens are visibly defective and one should not expect such tokens to realize their functional goals. The statement (∗) should thus be taken as defeasible. We propose to interpret it as: "*normal*" tokens of type T satisfy such means-end relations, rather than *every* T-token will work to realize the end.

This introduction of normal tokens is not intended as a definition of normality, but a clarification of the expectation in (∗). Such expectations may fail, for instance, when an artifact type is poorly designed and incapable of realizing its functional goal. One needs an account of such normal tokens, then, that is independent of (∗). We will sketch an account here and justify our reliance on normality in terms of semantic similarity to natural language interpretations of functions and the means they entail.

One may be tempted to avoid normal tokens by strengthening the contexts C, describing the assumptions about the physical condition of T-tokens that count for normal usage. We avoid this alternative for two primary reasons. First, with complicated artifacts, the user is unlikely to know what structural features are relevant for use or whether a particular token possesses such features. An engineer may know how the wires inside my television should be connected, but I surely do not know this and I am unable to easily confirm that they are properly connected in any case. So this precondition for using my television cannot be part of my understanding of its function. The second reason for keeping a notion of normal token separate from conditions of use is defended below, when we give a preliminary analysis of malfunction.

Thus, we suppose that every type comes with an implicit notion of normal tokens. Such normal tokens appear to be fictional: even if every token of type T happens to be broken, it may be that tokens of type T are *for* realizing functional goal φ via user

plan α. One must of course be careful in generating expectations regarding normal tokens. Do "normal" perpetual motion machines produce energy from nothing?

An average user forms his expectations regarding an artifact's performance by previous experience with similar artifacts and whatever inferences he draws regarding the manufacturer's intentions. An engineer with a deeper structural understanding of an artifact knows better how the artifact is supposed to work and uses this knowledge to form an intuition about normally functioning tokens. In each case, it is natural to express one's expectations about artifact behavior in terms of normal tokens. This terminology gives a concise expression of the practical consequences of functional ascriptions. That *normal* T-tokens can be used to realize our end gives one reason to believe that the T-token at hand may be used for this purpose, unless it is dissimilar to normal T-tokens in some relevant way. This reasoning seems significantly analogous to natural practical reasoning about artifacts and so we depend on normality in our development here.

We can use the means-end relation ($*$) to define functional terms, including artifact failure and malfunction. An artifact t of type T *fails to fulfill its function* in a particular application if the application occurs in a C-context and the artifact is used according to the use plan α, but the functional goal φ is not realized. Otherwise, it *fulfills* its function. Failure is not the same as malfunction. An anti-aircraft missile that fails to strike its target is not necessarily malfunctioning. Sometimes, missiles miss.

A token t is *malfunctioning* with respect to a proper function if it would fail to reliably or effectively realize φ in some normal contexts C when used properly according to α. Malfunction is a comparative claim, involving reliability (probability of realizing φ) and effectiveness (degree to which φ is realized, in the case of vaguely specified goals). Here again, we lean on our expectations regarding normal tokens to justify claims of malfunction. A lighter that fails to light as often as one would expect from a normally functioning lighter is malfunctioning. A hair-dryer that fails to dry hair as quickly (i.e., effectively) as a normally functioning hair-dryer is also malfunctioning. Clearly, malfunction is also a matter of degree: the slow-drying hair-dryer may still be usable, but it is not meeting our expectations. As well, the normal tokens must be drawn from a *suitably narrow type* that includes our token. A 1950's era television is malfunctioning only if it performs poorly compared to similar (normal) televisions. One would not compare a 1950's television to modern televisions for this purpose.

We have attempted to make explicit the way in which artifactual functions produce defeasible sufficient means-end relations. In doing so, we have tried to stay as close to natural language meanings of the relevant terms as possible. Because of this, we have relied heavily on a difficult notion: normal tokens of an artifactual type. Clearly, there is more to be said about this concept, but we believe that something like intuitions regarding normal tokens is at the heart of user expectations regarding artifacts. Furthermore, expectations about normal tokens allow one to distinguish failure from malfunction. Further reflection is required to distinguish malfunction from design failures and to apply this sketch of practical consequences to existing theories of artifactual function, but we believe that this is a promising start.

References

[1] Christopher Boorse. A rebuttal on functions. In André Ariew, Robert Cummins, and Mark Perlman, editors, *Functions: New Essays in the Philosophy of Psychology and Biology*, pages 63–112. Oxford University Press, 2002.
[2] Robert Cummins. Functional analysis. *Journal of Philosophy*, 72(20):741–765, 1975.
[3] W. Houkes and P.E. Vermaas. Actions versus functions: A plea for an alternative metaphysics of artifacts. *Monist*, 87:52–71, 2004.
[4] Ruth Garrett Millikan. *Language, Thought and Other Biological Categories*. The MIT Press, 1984.
[5] Karen Neander. The teleological notion of function. *Australasian Journal of Philosophy*, 74:261–301, 1991.
[6] Mark Perlman. The modern philosophical resurrection of teleology. *The Monist*, 87(1):3–51, January 2004.
[7] Georg Henrik von Wright. Practical inference. *The Philosophical Review*, 72(2):159–179, Apr. 1963.
[8] L. Wright. Functions. *Philosophical Review*, 82:139–168, 1973.

Eindhoven University of Technology, The Netherlands, J.Hughes@tm.tue.nl