Artificial Intelligence for Engineering Design, Analysis and Manufacturing (2010), *24*, 87–105.

doi:10.1017/S0890060409990187

Evaluation of the functional basis using an information theoretic approach

CHIRADEEP SEN, BENJAMIN W. CALDWELL, JOSHUA D. SUMMERS, AND GREGORY M. MOCKO
Department of Mechanical Engineering, Clemson University, Clemson, South Carolina, USA

(RECEIVED October 15, 2008; ACCEPTED July 3, 2009)

Abstract

A metric for computing the information content of function models in mechanical engineering design is proposed. Function models are graph-based representations used to describe the functionality of engineered artifacts, where the nodes are function verbs and the edges are the objects of action. The functional basis, a controlled vocabulary of these verbs and nouns organized in a three level hierarchy, is intended to support consistent representation of function models. The Design Repository is a Web-based archive of function models of consumer products described with the functional basis. This paper presents the theoretical underpinnings of a metric for the information content of function models, the assumptions required to support it, the definitions of key terms associated with it, and its practical interpretation. Finally, the metric is used to study the usefulness of the functional basis through a series of experiments on function models within the Design Repository. The results of the experiment indicate that the secondary level of the functional basis is the most beneficial to designers, both in terms of information content and information density.

Keywords: Design Repository; Function Models; Functional Basis; Information Content; Information Theory

1. INTRODUCTION

Throughout the engineering design process, multiple representations are used by designers to describe different aspects of the product. For example, requirements lists detail the customer's needs, solid models represent the spatial form of the solution, and finite element models simulate the structural behavior of the product. The practical usefulness of a representation lies in its ability to facilitate the design process and help the designer to make decisions. In this context, a quantifiable metric to assess the *usefulness* of design representations could help designers in selecting the appropriate representation for describing the design product, the design process, or the design problem. One way of assessing the usefulness of a representation could be to measure the amount of information it provides to the designer about the domain it describes, because additional information can enable the designer to conduct more reasoning, thereby revealing more facts to support the design decisions. Thus, information content can be used as a first level surrogate of the practical usefulness or *informativeness* of the representation. A metric of information content could help answer questions such as "How much information is *generated* by creating a representation?", "How much information is *contained* in a representation?", or "How much information in *transmitted* when a representation is exchanged between designers?" To this end, the overarching objective behind this research is to develop a metric of *information content* of design representations. As a first step toward this objective, function models are studied in this paper. A function model is a representation that describes the *intended* functionality of a system (Pahl et al., 2007). This metric could help designers in the following:

1. comparing competing concepts for the same design problem (design product),
2. comparing competing function models for the same concept (design product),
3. comparing competing vocabularies and rule sets (modeling schema), and
4. evaluating the rate of evolution of the design (design process)

In this paper, information content of function models is measured from the syntactic point of view. To evaluate the quality of a function model and its usefulness in design, a complete evaluation of the semantic content of the model and its ability to support reasoning actions needs to be

Reprint requests to: Joshua D. Summers, Department of Mechanical Engineering, Clemson University, 250 Fluor Daniel EIB, Clemson, SC 29634-0921, USA. E-mail: jsummer@clemson.edu

captured in a metric. However, the metric presented here does not capture the semantic information.

Constructing function models using controlled vocabularies and rules has been studied over the past three decades (Collins et al., 1976; Keuneke, 1991; Kirschman & Fadel, 1998; Szykman et al., 1999; Stone & Wood, 2000; Stone et al., 2005). The functional basis is a controlled vocabulary of functional verbs and nouns organized in a three level hierarchy, which was developed in a collaborative effort between industry and academia for enabling consistent modeling of product functionality (Szykman et al., 1999; Stone & Wood, 2000; Hirtz et al., 2002). This vocabulary was used to construct function models of 129 consumer products, which are stored in a Web-based Design Repository (http://repository.designengineeringlab.org/, accessed October 13, 2008). In previous research, an information metric for the verbs in a function model was initially proposed to assess the usefulness of the functional basis (Caldwell et al., 2008). This paper presents the theoretical underpinnings of this metric, the assumptions required to support it, the definitions of key terms associated with it, and its practical interpretation. The metric is used to study the usefulness of the functional basis vocabulary through a series of experiments on function models within the Design Repository. The results indicate that the secondary level of the functional basis hierarchy is the most beneficial one for constructing function models, as it provides an optimal balance between the two quantifiers presented in this paper, information content (Section 4) and information density (Section 7.2).

2. FUNCTIONS IN ENGINEERING DESIGN

The functionality of technical systems has been studied from multiple viewpoints in engineering design research. Pahl and Beitz (Pahl et al., 2007) describe function as "the intended input/output relation of a system whose purpose is to perform a task." Ullman describes it as "the desired output from a system" (Ullman, 1992). Otto and Wood (2001) provide a systematic method of describing functionality through reverse engineering. A solution-neutral description of system functionality is generally accepted to help widen the search for design solutions (Ullman, 1992; Otto & Wood, 2001; Pahl et al., 2007). The function models stored in the Design Repository are based on the definition provided by Pahl and Beitz.

Multiple models have been proposed in previous research for the representation of system functionality. Gero modeled functionality as an interaction between three aspects of the system, *function*, *behavior*, and *structure* (Gero, 1990). Tomiyama and colleagues (Umeda & Tomiyama, 1995) introduced the concept of *state* of the system, in the *function–behavior–state* model. Chandrasekaran and colleagues (Chandrasekaran, 2005) described function as *effect*. Goel and colleagues (Bhatta et al., 1994; Goel & Bhatta, 2004) described function as the *link* between structure and behavior. More recently, Albers and colleagues (2008) modeled

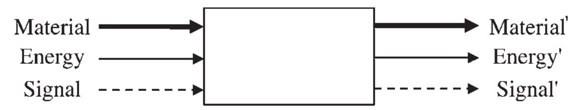

Fig. 1. A generic function block based on Pahl et al. (2007).

functions in terms of *working surface pairs* between interfacing system elements. Pahl and Beitz discussed the use of block diagrams for representing functions as a "solution-neutral relation between the inputs and the outputs" of the system (Fig. 1). The blocks describe a *reproducible transformative action*, whereas the arrows represent the *input* and *output* flows. The flows are broadly classified in previous research as that of *materials, energies, and signals* (Rodenacker, 1971; Pahl et al., 2007). The overall function of the system is described by linking together multiple functions via their flows, creating a function structure (Pahl et al., 2007). The function models stored in the Design Repository are based upon this representation.

Significant advances have been made in previous research in the directions of controlled vocabularies for function models (Collins et al., 1976; Keuneke, 1991; Kirschman & Fadel, 1998; Szykman et al., 1999; Stone & Wood, 2000; Hirtz et al., 2002), the representation of functions in computers (Szykman et al., 1999; Deng, 2002; Kitamura & Mizoguchi, 2003; Kitamura et al., 2004; Chandrasekaran, 2005; Kitamura et al., 2005), and the ways to reason upon those representations (Kurtoglu et al., 2005; Sridharan & Campbell, 2005).

2.1. The functional basis

The functional basis is a controlled vocabulary containing 54 function verbs (Table 1) and 45 flows or objects of action (Table 2) arranged in a three-level hierarchy. The left column in each table is the primary level, with the middle column being the secondary level, and the right column being the tertiary level. Here, the primary level is considered to be a *higher* level than the secondary, and the tertiary level is considered a *lower* level than the secondary.

The functional basis appears to be one of the most popular controlled vocabularies in literature (Kurfman et al., 2001; Hirtz et al., 2002; Kurtoglu et al., 2005; Vucovich et al., 2006). The wide use of the functional basis necessitates an objective external examination of its features, particularly, the usefulness of the hierarchical organization of terms to construct function models.

2.2. The function-based Design Repository

The Design Repository is a Web-based archive of function models of consumer products that was created through reverse engineering and product dissection to catalog the function of each component or subsystem using the functional

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

Table 1. *Functional basis verbs hierarchy*

Primary	Secondary	Tertiary
Branch	Separate	Divide
		Extract
		Remove
	Distribute	
Channel	Import	
	Export	
	Transfer	Transport
		Transmit
	Guide	Translate
		Rotate
		Allow DOF
Connect	Couple	Join
		Link
	Mix	
Control magnitude	Actuate	
	Regulate	Increase
		Decrease
	Change	Increment
		Decrement
		Shape
		Condition
	Stop	Prevent
		Inhibit
Convert	Convert	
Provide	Store	Contain
		Collect
	Supply	Supply
Signal	Sense	Detect
		Measure
	Indicate	Track
		Display
	Process	
Support	Stabilize	
	Secure	

Vocabularies based on Hirtz et al. (2002).

Table 2. *Functional basis nouns hierarchy*

Primary	Secondary	Tertiary
Material	Human	
	Gas	
	Liquid	
	Solid	Object
		Particulate
		Composite
	Plasma	
	Mixture	Gas–gas
		Liquid–liquid
		Solid–solid
		Solid–liquid
		Liquid–gas
		Solid–gas
		Solid–liquid–gas
		Colloidal
Signal	Status	Auditory
		Olfactory
		Tactile
		Taste
		Visual
	Control	Analog
		Discrete
Energy	Human	
	Acoustic	
	Biological	
	Chemical	
	Electrical	
	Electromagnetic	Optical
		Solar
	Hydraulic	
	Magnetic	
	Mechanical	Rotational
		Translational
	Pneumatic	
	Radioactive/nuclear	
	Thermal	

Vocabularies based on Hirtz et al. (2002).

basis. Approximately half of these products are available in graph-based function models, whereas customer requirements, function–component matrices, and component–assembly matrices are available for all. Figure 2 shows a screenshot of the Design Repository Web page, illustrating the data stored for a specific component (heating coil frame) of a specific product (Supermax hair dryer). Much of the information captured in the database is not directly related to functionality, such as material, manufacturing process, and physical parameters. The functionality of the components and subsystems are captured through the function list with input and output flows. For the heating coil frame, the function is listed as to couple one solid (heating coil) to another solid (motor housing). These supporting functions are used to represent the physical attachments between components. In addition, graphical function models, such as Figure 3 are included in the Design Repository. Unfortunately, these models are static and do not directly support computational reasoning. The intent of the Design Repository was to record enough information based on which the graphical function models could be recreated. However, no computational tool has been developed to actually accomplish that task, nor has the feasibility of accomplishing this translation been explicitly demonstrated.

Multiple design enabling tools and methods have been developed based on the functional basis and the Design Repository. The Concept Generator Tool uses existing function models and component failure data from the Design Repository as input to produce component arrangements for new concepts, much like an automated morphological chart (Vucovich et al., 2006). The function-failure design method predicts potential failure modes of new designs in the conceptual design phase, based on the archived failure history of components performing certain functions (Tumer & Stone, 2001; Arunajadai et al., 2002; Stone et al., 2005). A graph grammar-based tool has been developed to use the probability of occurrence of functional basis terms in a model to synthesize new functions (Sridharan & Campbell, 2005). Each of these

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

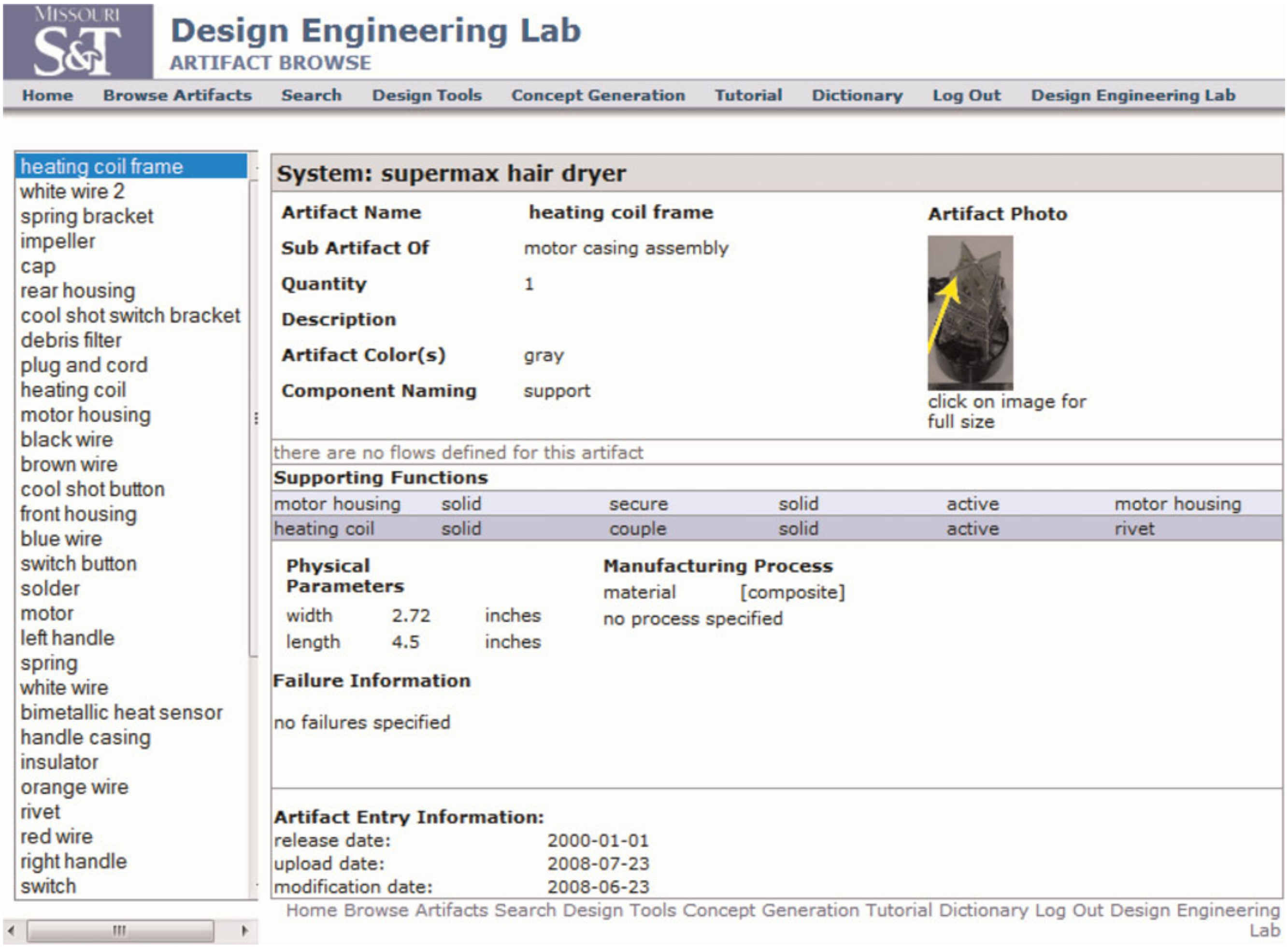

Fig. 2. The artifact browser in the Design Repository showing the heating coil frame of the Supermax hair dryer (http://repository.designengineeringlab.org/, accessed October 13, 2008). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

tools is developed on the fundamental assumption that the underlying principles of the Design Repository, and especially the functional basis, are sound. A critical area that has not been studied in previous research is the utility of the different levels of the hierarchy of the functional basis. This utility is explored here through the application of information theoretic measures on the models using different levels of the hierarchy. The next section develops this metric.

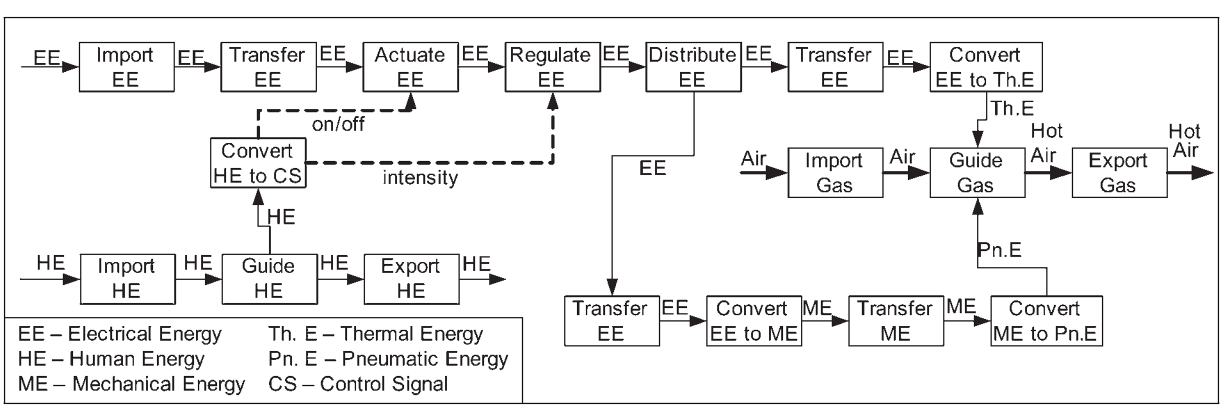

Fig. 3. Function model of Supermax hair dryer as stored in the Design Repository (http://repository.designengineeringlab.org/, accessed October 13, 2008).

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

3. INFORMATION THEORY OF FUNCTION MODELS

Information theory, originally developed in the context of communication, provides a mathematical measure for information content of a message produced by a discrete source (Hartley, 1928; Shannon, 1948). In this context, a message constitutes of a stream of events that carries some information. Conversely, an event is a unit block of information in a message. The source is discrete if the events occur as distinct units of the message with no provision for partial occurrence. The source is linear if the events are produced sequentially. The events in the message are selected from a predefined, finite list of allowed events or controlled vocabulary, where each event has a known probability of occurrence in the message. Under these premises, the information content of a single event in a message is given by Shannon (1948):

$$I_i = -K \cdot \sum_{j=1}^{n} p_j \cdot \log_b p_j, \tag{1}$$

where I_i is the information content of a single event in the message, K is a constant for scaling between different sources of information, n is the size of the finite predefined vocabulary, j is the counter of the elements in the vocabulary, p_j is the probability of occurrence of the jth element of the vocabulary, and b is the base of the logarithm, a positive integer.

The constant K scales the quantity inside the summation sign and assumes different values for different sources. Hence, K can be used to compare information content across different design representations. This comparison is reserved for future work. In this research, only a single representation, the function model, is studied. Because the metric will be used only to compare relative information content of function models and not absolute values, K is set to unity. The premises of information theory are next mapped to the features of function models to justify the use of the information metric.

3.1. Correspondents of information theory in function models

Function models, such as those stored in the Design Repository, consist of two sets of verbs and nouns, each being defined by discrete elements. Although the model contains all the elements in a graphical representation, for the sake of computing information content the elements are considered sequentially, making the model linear (see assumption 1). The elements of the model shown in Figure 3 are drawn from a specific hierarchical level of the functional basis (see Table 1 and Table 2), which are finite vocabularies of predefined sizes. The probability of occurrence of terms in the functional basis and their dependencies have been studied in previous research (Kurtoglu et al., 2005; Sridharan & Campbell, 2005); however, so far no conclusion has been generally accepted. Therefore, a uniform distribution of independent probabilities of terms in the functional basis is assumed (see assumptions 2 and 3).

3.1.1. Assumptions

1. *A function model is a linear source, that is, its elements are encountered by the observer in a sequential fashion.*
2. *The probability of occurrence of verbs and nouns in the functional basis is uniformly distributed over the respective vocabularies (Caldwell et al., 2008).*
3. *The probability of occurrence of the verbs and nouns in the functional basis is independent of each other (Caldwell et al., 2008).*

Formally, the concepts of message, event, source, vocabulary and probability distribution are mapped from information theory to corresponding concepts in function modeling (Table 3).

3.2. Information metric for functional elements: General form

Under assumptions 2 and 3, and setting $K = 1$, Eq. (2) undergoes the following change.

$$\begin{aligned} I_i &= -K \cdot \sum_{j=1}^{n} p_j \cdot \log_b p_j = -(1) \cdot \sum_{j=1}^{n} p \cdot \log_b p \\ &= -n \cdot p \cdot \log_b p = -\log_b p = \log_b(n), \end{aligned} \tag{2}$$

where I_i is the information content of a single event, j is the counter of the elements in the vocabulary, p_j is the probability of the jth element of the vocabulary, p is the uniform probability of all elements in the vocabulary under assumption 2, $n = 1/p$ is the size of the vocabulary, and b is a positive integer.

Equation (2) represents the information content per element in the function model. The base of logarithm b is essentially a scaling factor for I. As shown in Eq. (3), changing the base from b to c scales I by a constant, $\log_b(c)$.

$$\log_b(x) = \log_c(x) \times \log_b(c). \tag{3}$$

Therefore, the base can be arbitrarily chosen, as long as the choice is consistently maintained for all computations. Here, the value 2 is selected as it provides an intuitive practical interpretation of the metric, as will be discussed in Section 5.

Table 3. *Correspondents of information theory in function models*

Concepts in Information Theory	Correspondents in Function Modeling
Message	Set of verbs and set of nouns in a function model
Event	Individual verbs and nouns in a function model
Discrete, linear source	The function model
Finite predefined vocabulary	The list of verbs and nouns in a specific level of the functional basis
Probability distribution of events over the vocabulary	Assumed uniform over the functional basis

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

The choice of the base determines the unit of information, which, for $b = 2$, is bits (Shannon, 1948). The unit information content per element of the function model is thus simplified to

$$I_i = \log_2(x)\ \text{bits/element}, \quad (4)$$

where I_i is the unit information, that is, information per element of the model, and x is the number of terms in the vocabulary from which the element is drawn.

For y distinct elements in the function model, the total information content is given by

$$I = \sum_{i=1}^{y} I_i = \sum_{i=1}^{y} \log_2(x) = y \cdot \log_2(x)\ \text{bits}, \quad (5)$$

where I is the information content of all elements in the message, I_i is the information content of the i th element of the message, x is the number of terms in the vocabulary from which the elements are drawn, and y is the number of elements in the functional model.

Ultimately, Eqs. (4) and (5) are defined as the general metrics of information content of function models in this research. This measure of information content has previously been used to measure size complexity of engineering models (Summers & Shah, 2003; Summers & Ameri, 2008).

4. INFORMATION CONTENT OF FUNCTION MODELS

Each element of the function model contributes to the informativeness of the model, because by the removal of any element, or a set, the model becomes less informative to the designer than the initial model. Thus, the information content of the whole function model is expected to be of the form $I_{\text{FM}} = f(V, N)$, where V and N are the respective sets of verbs and nouns in the model. The function f describes how the information from the elements contributes to the total information content. In this research, two possible definitions of f are identified, namely, *element-wise* and *combined* information content. Both of these definitions are discussed in this section.

It is noteworthy that the topology of a function model, the connectedness of the verbs with the nouns, also contributes significantly to its informativeness, as the model is more informative to the designer when the verbs and nouns are arranged in the topological arrangement rather than in a flat list. However, the aim of this research is to evaluate the functional basis, rather than to compute the total information content of function models. As the functional basis itself provides only verbs and nouns, but no guideline to construct the topology of the model, topological information content cannot be utilized to evaluate this vocabulary, even if it was computed. For this reason, this component of information is not included in this research, but saved for future work, where the complete information content of function models are computed.

4.1. Element-wise information content of function models

The element-wise information content of a function model is the algebraic sum of information contributed by the individual elements. If the number of verbs and nouns in the respective vocabularies are given by x_V and x_N, and the number of verb instances and noun instances in a specific function model is given by y_V and y_N, the following metrics are obtained from Eq. (5).

4.1.1. *Definitions*

Information content of verbs in a function model:

$$I_V = y_V \cdot \log_2(x_V)\ \text{bits} \quad (6)$$

Information content of nouns in a function model:

$$I_N = y_N \cdot \log_2(x_N)\ \text{bits} \quad (7)$$

Information content of the whole model:

$$I_{FM} = I_V + I_N = y_V \cdot \log_2(x_V) + y_N \cdot \log_2(x_N)\ \text{bits} \quad (8)$$

The unit information per verb, I_{iV}, and per noun, I_{iN}, can be obtained by putting $y_V = 1$ and $y_N = 1$ in Eqs. (4) and (5), respectively.

4.2. Combined information content of function models

The combined information content of a function model is based on a combined vocabulary obtained by concatenating the individual vocabularies of verbs and nouns. If the number of verbs and nouns in the respective vocabularies are given by x_V and x_N, and the number of verb instances and noun instances in the function model is given by y_V and y_N, respectively, the following metrics are obtained from Eqs. (9) and (10).

4.2.1. *Definitions*

Combined information content per element:

$$I_{i(N+V)} = \log_2(x_V + x_N)\ \text{bits/element} \quad (9)$$

Combined information content of the whole model:

$$I_{N+V} = (y_V + y_N) \cdot \log_2(x_V + x_N)\ \text{bits} \quad (10)$$

The element-wise and combined information metrics are used in Section 8.1.4 to measure the information content of function models.

5. PRACTICAL INTERPRETATION OF INFORMATION CONTENT

To illustrate the practical meaning of the information metric as it relates to function modeling, a hypothetical scenario is shown in Figure 4. A discrete linear source S is producing a message by drawing events from a finite, predefined vocabulary Σ, contain-

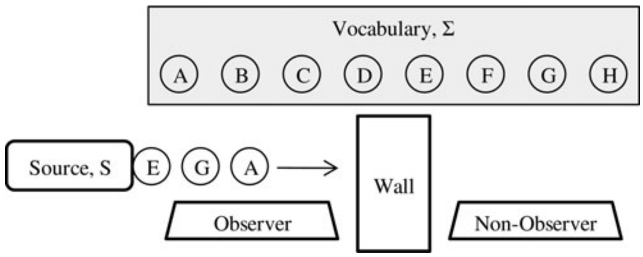

Fig. 4. Hypothetical scenario for the interpretation of information.

ing elements A through H, with equal probability. An observer is witnessing the events as they occur, and transmitting that information to a nonobserver who is separated from the source. The vocabulary is known to both the observer and the nonobserver. According to the correspondents explained in Section 3.1, this setup simulates the transmission of a function model from one designer (observer) to another (nonobserver), element by element. With each element transmitted, the nonobserver comes to know more of the function model (source S). Thus, an important question arises, "What is the value of the information transmitted by the observer per element?"

To answer this question, let the setup change so that the nonobserver is required to determine the events by asking binary questions to the observer, similar to the game "Twenty Questions." Binary questions are answered either yes or no. Given the vocabulary and its probability distribution, the nonobserver can determine an event with minimum number of questions by adopting a binary search through the vocabulary. This search is executed by recursively dividing the vocabulary in halves and asking if the occurred event is in the left half. Starting with a vocabulary of size x, the size of the search space reduces with each question according to the geometric series $\{x, x/2, x/4, \ldots, 4, 2, 1\}$, until the correct event is found. The number of questions required to determine the event is one less than the number of terms in the series, given by:

$$N_{\min} = \log_2(x), \tag{11}$$

where $N_{\min}$ is the minimum number of binary questions required to determine the event and x is the number of terms in the vocabulary Σ.

Thus, it can be argued that in the initial communication setup, the nonobserver was receiving a value of $\log_2(x)$ with each event because the information received from the observer was equivalent to receiving answers to $\log_2(x)$ questions. The form of the expression in Eq. (11) is identical with the general form of the information metric in Eq. (4). Therefore, the information content of each element practically represents the minimum number of binary questions that must be asked in order to identify an element from the vocabulary.

The nonobserver can rebuild the entire function model, element by element, by asking $\log_2(x)$ questions for each element. At this point, because of the equality between the original and the rebuilt model, it can be argued that any reasoning supported by the original model is equally supported by the rebuilt model. Thus, the entire information stored in the original model has been transmitted, though indirectly, from the observer to the nonobserver, in the form of answers to a finite number of questions, $\log_2(x)$.

6. VALIDATION OF THE METRIC AGAINST REQUIRED CRITERIA

Four requirements for information metrics have been discussed in literature (Shannon, 1948; Cover & Thomas, 2006). The metric presented in Section 4 is validated against these requirements to ensure that by adopting assumptions 1, 2, and 3, the fundamental properties of the metric are not violated.

6.1. Requirement 1

Information is always a nonnegative quantity (Cover & Thomas, 2006). In a function model there is always at least one verb and at least one noun ($y_V \geq 1$, $y_N \geq 1$), and the vocabularies contain at least one verb and one noun each ($x_V \geq 1, x_N \geq 1$). Hence, the minimum value of the expressions in Eq. (5) is $I_{\min} \geq (1) \times \log_2(1)$, that is, $I_{\min} \geq 0$. Thus, the metric satisfies this requirement.

6.2. Requirement 2

If an event has probability of 1, no information is obtained from its occurrence (Cover & Thomas, 2006). In function models, the condition implies that there is only one verb or noun repeatedly used in the function model. In that case, the term becomes fully predictable where no additional information is gained by knowing about its occurrence. Mathematically, by setting $x_V = 1$ and $x_N = 1$ in Eqs. (6) and (7), I_V and I_N vanish. Thus, the metric satisfies this requirement.

6.3. Requirement 3

If two independent events occur, whose joint probability is the product of their independent probabilities, the total information obtained is the sum of their individual information (Cover & Thomas, 2006). If i and j are two elements of a vocabulary, with independent probabilities $p_{i'}$ and p_j, the probability of their joint occurrence is given by $p_i \times p_j$. Under assumptions 2 and 3, $p_i = p_j = 1/x$, where x is the size of the vocabulary. Hence, the probability of the joint occurrence of i and j is $(1/x) \times (1/x) = 1/x^2$, which is equivalent to the independent uniform probability of a single element in a vocabulary of size x^2. Thus, if the individual information content of events i and j are I_i and I_j, the information produced by their joint occurrence is obtained from Eq. (4) as

$$I_{i+j} = \log_2(x^2) = 2 \cdot \log_2(x) = \log_2(x) + \log_2(x) = I_i + I_j, \tag{12}$$

where i, j are two distinct elements of the vocabulary of size x, I_i is the individual information content of element i in the vocabulary of size x, I_j is the individual information content of element j in the vocabulary of size x, and I_{i+j} is the individual information content of an element in the vocabulary of size x^2. The metric, therefore, satisfies this requirement.

6.4. Requirement 4

Information is a monotonic continuous function of the probabilities, that is, a slight increase in the probabilities should always result into a slight increase in information (Shannon, 1948; Cover & Thomas, 2006). Figure 5 shows the plot of unit information against the size of the vocabulary, which satisfies the criterion because of the monotonic behavior of logarithms. As discussed in requirement 1, the practically usable portion of the curve is in the range $x \geq 1$, because a null vocabulary is unusable for creating messages.

7. DISCUSSION ON THE INFORMATION METRIC

In this section, important properties of the metric are reviewed and their implications to function models are discussed.

7.1. Response to variables

As seen in Eq. (5), the information content I of a function model increases linearly with the size of the model y and logarithmically with the size of the vocabulary x. Thus, the metric is more sensitive to the change of model size than to the change of the vocabulary size. This implies that a means to arrive at larger models, such as decomposition, can help increase the informativeness of a model more than using a larger vocabulary to construct the model. Intuitively, in a large vocabulary, the distinction between the terms becomes gradually obscured. Hence, the model's informativeness to the designer does not increase significantly.

7.2. Information density of a vocabulary

Because information is a monotonically increasing function of the vocabulary size, the information obtained from a larger vocabulary is always larger, but the increase in information gradually diminishes with increasing size of the vocabulary. As observed in Figure 5, the increase of information because of unit increase of the vocabulary size from 2 to 3, indicated by $\Delta I_{2,3}$, is larger than the increase in information because of the same increase in the vocabulary size from 5 to 6, indicated as $\Delta I_{5,6}$. This observation enables the formulation of a new quantity to assess the usefulness of the vocabulary itself. This quantity, termed information density, is defined below:

> *Information density of a vocabulary is the amount of information produced by a single event, measured per unit size of the vocabulary.*

The information density of a vocabulary of size x implies the usefulness of the vocabulary in terms of the *benefit* (information produced) over *cost* (size of the vocabulary), and is obtained by dividing both sides of Eq. by the size of the vocabulary, x.

$$I' = \frac{I_i}{x} = \frac{\log_2(x)}{x}, \tag{13}$$

where I' is the information density of the vocabulary, I_i is the unit information per element of the vocabulary, and x is the size of the vocabulary.

7.3. Quantity versus quality

The metric provides a measure for only the quantity, not the quality, of information stored in a function model. The numeric value of information can be increased merely by increasing the number of terms in the function model or the vocabulary, even if the model does not describe the system correctly or consistently. The issue of measuring quality of a function model is saved for future work.

8. APPLICATION OF THE INFORMATION METRIC

The information content of function models is measured in a series of experiments on three products within the Design Repository. The products are the Supermax hair dryer, the Delta jigsaw, and the Brother sewing machine. These products are chosen for the experiments as they are representative of the many products in the Design Repository, they demonstrate the use of many of the functional basis' commonly used functions, and one of them, the hair dryer, has been used as example in previous research related to function modeling (Mocko et al., 2007; Caldwell et al., 2008). In addition, these products demonstrate a variety of size and verbs/nouns ratio. Although the function model for Supermax hair dryer has 18 verbs and 24 nouns (a ratio of 0.75), the Brother sewing machine has 44 verbs and 64 nouns (a ratio of 0.69), and the Delta jigsaw has 17 verbs and 42 nouns (a ratio of 0.40). The results of these experiments are used to evaluate the usefulness of the functional basis vocabulary.

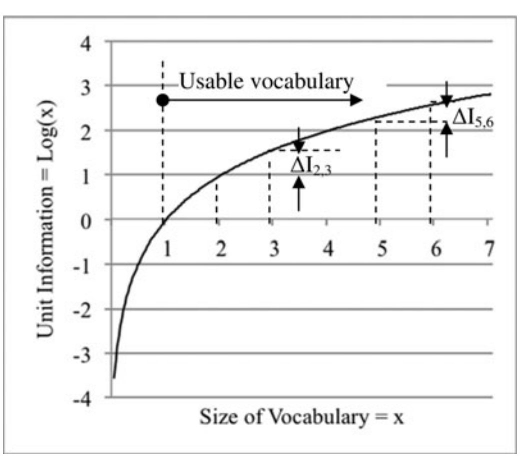

Fig. 5. Unit information versus the size of vocabulary: logarithmic plot with base 2.

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

8.1. Experimental protocol

Four experimental steps are defined in this protocol and illustrated through the Supermax hair dryer example. These steps are model cleanup, translating the models across functional basis levels, defining the vocabularies, and computing the information content.

8.1.1. Model cleanup

The three function models are selected from the Design Repository, and corrected for inconsistencies. For example, the hair dryer function model shown in Figure 3 contains some non-functional basis terms, such as “hot air” and “intensity.” These terms are replaced with terms from the same level of the functional basis as used throughout the model, such as “gas” and “control signal.” The adjective “hot” is dropped, because the functional basis does not provide any vocabulary of adjectives.

Next, the model is cleaned up from redundancies. For example, in the block containing the text “convert EE to ThE,” all words other than the verb “convert” are deleted, because the arrows associated with the block are sufficient to indicate the nouns. Figure 6 shows the cleaned up model of the Supermax hair dryer, with the six corrections for inconsistencies highlighted with circles. The function models of Delta jigsaw and Brother sewing machine, as obtained from the Design Repository, are shown in Figure A.1. in Appendix A and Figure B.1. in Appendix B.

8.1.2. Translating function models across functional basis levels

After a function model is cleaned, it is *translated*, that is, it is redefined with verbs and nouns from other levels of the functional basis, without any change to its topology. Because there are three hierarchical levels for both verbs and nouns, a model can be translated to 16 different *designations*, as shown in Table 4.

A model described with the *m*th level of verbs hierarchy and the *n*th level of nouns hierarchy of the functional basis is *designated* as M(*m*, *n*). For example, M(2, 3) designates

Table 4. *Designation protocol of function models*

	Verb Levels			
Noun Levels	0	1	2	3
3	M(0, 3)			M(3, 3)
2	M(0, 2)		M(2, 2)	
1	M(0, 1)	M(1, 1)		
0	M(0, 0)	M(1, 0)	M(2, 0)	M(3, 0)

a model with secondary level verbs and tertiary level nouns. The bottom row M(*m*, 0), and the left column M(0, *n*) designate models with only one type of terms. For example, M(3, 0) designates a model with tertiary level verbs in the blocks, but no nouns on the arrows. These models allow the measurement of element-wise information content. The diagonal designates models with the same levels of verbs and nouns. These models show trends of the combined information content. Here, M(0, 0) represents the empty topology of the model and contains zero information. The empty cells designate models described with mixed levels of verbs and nouns; these models are not used in the experiments, because conventionally, the function models within the Design Repository are defined with the same hierarchical level of verbs and nouns.

When a function model is translated from a lower to a higher level (upward translation), the taxonomical parent of each lower level element is chosen as the new element. When a model is translated from a higher to a lower level (downward translation), each new element is chosen from the taxonomical children of the higher level element using some knowledge about the product. For example, in Figure 3, the secondary function “guide” is translated to “channel” in upward translation, whereas in downward translation “allow DoF” is selected because the definition of “allow DoF” in the functional basis best describes the actual function in the product. Thus, upward translations are more objective than downward translations; however, owing to the uniform

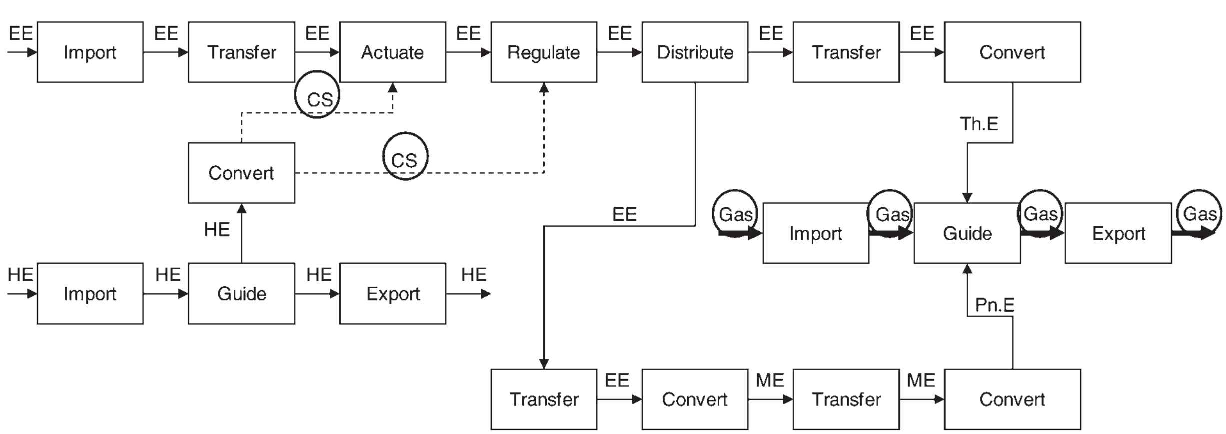

Fig. 6. Hair dryer function model with secondary verbs and secondary nouns, after cleanup.

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

probability distribution over the levels in the functional basis, the specific selection does not impact the numeric score of information content. To ensure that each higher level term is represented in the lower levels, secondary terms that have not been categorized in the tertiary level are propagated, as is, to the tertiary level. For example, in Figure 6, the secondary verbs "distribute," "import," and "export" are all propagated to the tertiary level at the time of translation. Figure 7 and Figure 8 shows the hair dryer function models of designations M(1, 1) and M(2, 2), respectively. These models are obtained by translation from Figure 6, which is of designation M(2, 2).

8.1.3. *Defining three types of vocabularies for computing information content*

Because of the hierarchical arrangement of terms in the functional basis, a downward translation enables at least three definitions of the lower level vocabulary to be used for computations, as defined below:

Definitions.

1. *The fixed vocabulary of a given level is the collection of all terms in the level.*
2. *The used vocabulary of a given level and a given function model described on that level is the set of terms that appear in the model.*
3. *The reduced vocabulary for a given function model that is obtained by translation from a higher to a lower level is the set of all lower level terms that can be obtained as taxonomical children of the higher level terms used by the higher level function model.*

In the hair dryer function model, the fixed vocabulary of verbs for all models of designations M(1, *n*), M(2, *n*), and M(3, *n*) are given by all the verbs in the corresponding columns of Table 1, that is, 8, 21, and 35, respectively. The used vocabulary of verbs for the models of designation M(1, *n*), M(2, *n*), and M(3, *n*) are the number of verbs appearing in Figures 7, 6, and 8, respectively, that is, four, eight, and eight, respectively. The reduced vocabulary of verbs for all models of designation M(1, *n*) is accepted to be identical with the fixed verb vocabulary of the same models, because primary models cannot be obtained in downward translation. Because the used vocabulary of verbs for M(1, *n*) consists of "branch," "channel," "control," and "convert," the reduced vocabulary for M(2, *n*) is taxonomically obtained as the following list:

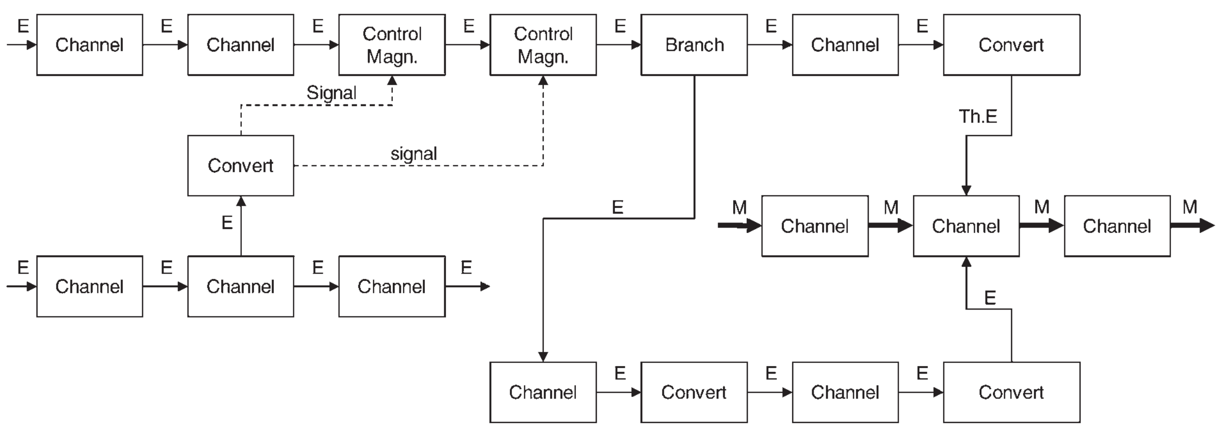

Fig. 7. Hair dryer function model with primary verbs and primary nouns, M(1, 1).

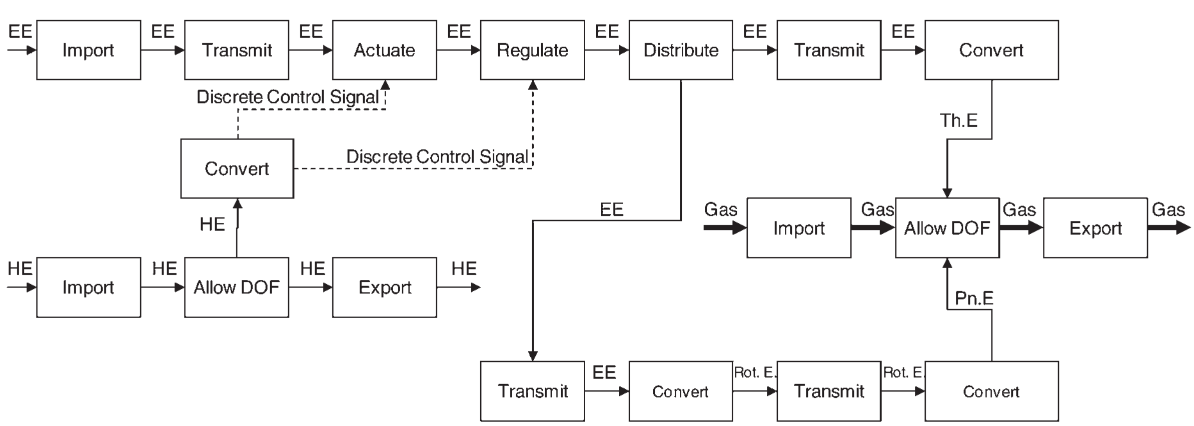

Fig. 8. Hair dryer function model with tertiary verbs and tertiary nouns, M(3, 3).

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

"separate" and "distribute" (obtained from "branch"), "import," "export," "transfer," and "guide" (obtained from "channel"), "actuate" and "regulate" (obtained from "control magnitude"), and "convert" (obtained from "convert"), a list of 11 verbs. Similarly, for all models of designation M(3, n), the reduced verbs vocabulary is of size 12.

In a similar way, the noun vocabularies of these three types are determined for each row in Table 4. The combined vocabularies are obtained by adding up the sizes of the corresponding verb and noun vocabularies. Table 5 shows a summary of the verb [M(m, 0)], noun [M(0, n)], and combined [M(n, n)] vocabularies of the fixed (F), used (U), and reduced (R) types, for all 10 designations. In each cell under columns U and R, the values separated by commas represent used and reduced vocabulary sizes for the Supermax hair dryer, the Delta jigsaw, and the Brother sewing machine, respectively. The fixed vocabulary size is essentially a property of the vocabularies, not the models, and hence remains equal for all products in each level.

8.1.4. Computing information content

To compute information content, first the sizes of the respective models are determined. Because there are 18 verb instances and 24 noun instances in the hair dryer function model (Fig. 6), the size of all function models in the bottom row of Table 5, y_V, is 18, and the size of all models in the left column of Table 5, y_N, is 24. The size of all models on the diagonal, $y_V + y_N$, is 18 + 24 = 42. The empty model, M(0, 0) is an exception, with size zero. The information content of the whole model is then computed by applying the appropriate equation, Eq. (8) or Eq. (10). In each case, the result of the logarithm is rounded up to the next higher integer so that, in accordance with the practical interpretation of the metric, a whole binary question is counted for the fractional part of the logarithm. Notably, the rounding up is done *before* multiplying by y, as opposed to rounding the total information content obtained after multiplying by y, because according to the practical interpretation, each element of the model needs a finite number of questions to be fully known by the nonobserver (see Section 5). For example, information for M(2, 0) and M(0, 2) about the fixed vocabulary are computed using Eq. (8) as $18 \times \log_2(8) + 0 = 18 \times 3 = 54$ bits, and $0 + 24 \times \log_2(20) = 24 \times 4.3 = 24 \times 5 = 120$ bits, respectively. Similarly, the combined information for M(2, 2) about the fixed vocabulary is computed using Eq. (10) as $42 \times \log_2(41) = 42 \times 5.4 = 42 \times 6 = 252$ bits. The results of the computations for the Supermax hair dryer are shown in Table 6. The results for the Delta jigsaw and the Brother sewing machine are provided in Table A.1 in Appendix A and Table B.1 in Appendix B.

8.2. Results

The results tabulated in Table 6 for the Supermax hair dryer are organized in bar charts for comparison in this section. Figure 9 shows the nine data points from the bottom row of Table 6, which are the element-wise information contents of the verbs. The three clusters of bars represent the primary, secondary, and tertiary levels of the verbs hierarchy, corresponding to models of designation M(1, 0), M(2, 0), and M(3, 0). Within each cluster, the individual bars represent information content using the fixed, used, and reduced vocabularies.

Similarly, Figure 10 shows the nine data points from the left column of Table 6, which are the element-wise information

Table 5. *Summary of vocabulary sizes for the Supermax hair dryer, Delta jigsaw, and Brother sewing machine function models*

	Verb Levels											
	0			1			2			3		
Noun Levels	F	U	R	F	U	R	F	U	R	F	U	R
3	36	7, 10, 9	9, 12, 16							71	15, 23, 24	21, 36, 34
2	20	7, 3, 7	20, 20, 20				41	15, 21, 17	31, 38, 33			
1	3	3, 3, 3	3, 3, 3	11	7, 10, 8	11, 11, 11						
0	0	0	0	8	4, 7, 5	8, 8, 8	21	8, 13, 10	11, 18, 13	35	8, 13, 15	12, 24, 18

Table 6. *Results of information content of Supermax hair dryer*

	Verb Levels											
	0			1			2			3		
Noun Levels	F	U	R	F	U	R	F	U	R	F	U	R
3	144	72	96							294	168	210
2	120	72	120				252	168	210			
1	48	48	48	168	126	168						
0				54	36	54	90	54	72	108	54	72

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

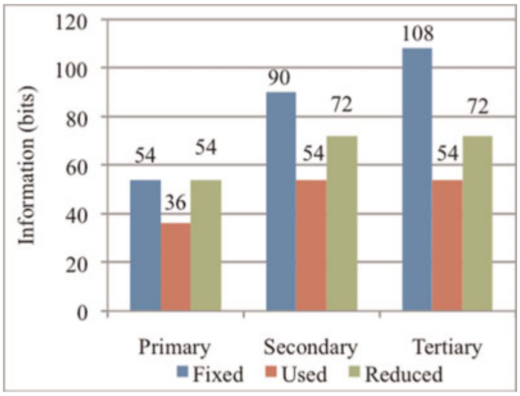

Fig. 9. Supermax hair dryer information content, verbs only: M(1, 0), M(2, 0), M(3, 0). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

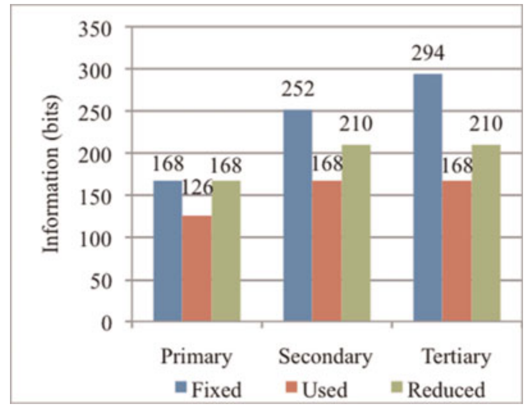

Fig. 11. Supermax hair dryer information content, combined: M(1, 1), M(2, 2), M(3, 3). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

contents of nouns in function models of designation M(0, 1), M(0, 2), and M(0, 3), and Figure 11 shows the nine data points from the diagonal of Table 6, which are the combined information contents of verbs and nouns in function models of designations M(1, 1), M(2, 2), and M(3, 3). The results for the Delta jigsaw and Brother sewing machine function models are shown in Figures A.2 to A.4 in Appendix A and Figures B.2 to B.4 in Appendix B. These figures are also organized as explained above.

9. OBSERVATIONS AND ANALYSIS

Table 7 summarizes the trends of information content based on the experimental results. There are 27 trends recorded, resulting from the combination of three products, three vocabulary types (fixed, used, reduced), and three metrics (verb, noun, combined). Here, $\Delta I_{\mathrm{I,II}}$ represents the change in information content from the primary to the secondary level and $\Delta I_{\mathrm{II,III}}$ indicates the change in information content from the secondary to the tertiary level. The symbols +, 0, and – in a cell under $\Delta I_{\mathrm{I,II}}$, for example, indicate that the information content based on the secondary level is greater than, equal to, or lower than the information content based on the primary level.

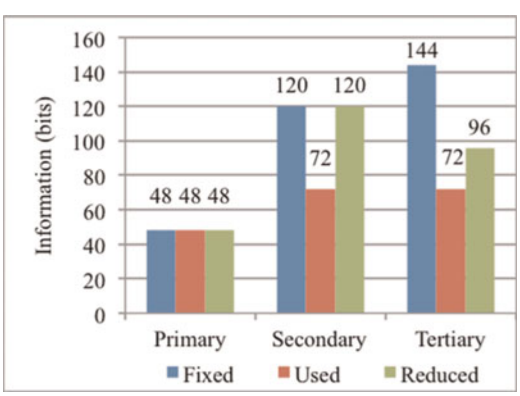

Fig. 10. Supermax hair dryer information content, nouns only: M(0, 1), M(0, 2), M(0, 3). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

Eight interpretations of the results are presented here. They address the variation in information content across the hierarchical levels of the functional basis, the comparative increase of information across those levels, and the trends in information density.

1. *Information content of function models based on the fixed vocabulary monotonically increases from the primary to the secondary to the tertiary level of the functional basis* (top three rows of data in Table 7). This trend is consistent for the verb, noun, and combined metrics, for all three products examined. This trend is expected, as the vocabularies increase in size with the levels (see Table 5).
2. *Information content of function models based on the used vocabulary increases from the primary to the secondary level, but usually remains the same between the secondary and tertiary levels* (middle three rows of data in Table 7). This trend is consistent in all but two out of nine cases. The two exceptions occurred in the noun metrics in the Delta jigsaw and Brother sewing machine models, where the information content increased from the secondary to the tertiary level. However, this increase is marginal: 168 – 126 = 42 bits in the Delta jigsaw, and 256 – 192 = 64 bits in the Brother sewing machine. As a result, the overall information, shown by the combined information content, remains the same between the secondary and the tertiary levels for both products. This observation indicates that even though the vocabulary size increases between the levels, the usage of terms in function models does not increase proportionately, which means that the tertiary level contains redundant functional terms, both verbs and nouns.
3. *Information content of function models based on the reduced vocabulary increases from the primary to the secondary level, but usually remains the same from the secondary to the tertiary level, in case of the verbs and the combined metrics* (first and third row of the last three rows of data in Table 7). This observation is consistent through all but one out of six cases: the verb metric of the sewing machine. This trend is identical

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

Table 7. *Trend of information content across functional basis levels*

		Supermax Hair Dryer		Delta Jigsaw		Brother Sewing Machine	
Voc. Type	Metric Type	$\Delta I_{\text{I,II}}$	$\Delta I_{\text{II,III}}$	$\Delta I_{\text{I,II}}$	$\Delta I_{\text{II,III}}$	$\Delta I_{\text{I,II}}$	$\Delta I_{\text{II,III}}$
Fixed	Verb	+	+	+	+	+	+
	Noun	+	+	+	+	+	+
	Combined	+	+	+	+	+	+
Used	Verb	+	0	+	0	+	0
	Noun	+	0	+	+	+	+
	Combined	+	0	+	0	+	0
Reduced	Verb	+	0	+	0	+	+
	Noun	+	−	+	−	+	−
	Combined	+	0	+	0	+	0

with observation 2, and it reinforces the analysis that the tertiary level contains many redundant terms, which add little information content.

4. *Information content of function models based on the reduced vocabulary using the noun metric increases from the primary to the secondary level, but decreases from the secondary to the tertiary level* (middle row of the last three rows of data in Table 7). As discussed in Section 8.1.3, the reduced vocabulary is obtained in two steps. First, the used vocabulary of the higher level is determined. Next, upon downward translation, this used vocabulary expands into its taxonomical children of the lower level. The vocabulary first reduces then expands in this process. Although the reduction depends entirely on the function model, the expansion is entirely dependent on the hierarchical structure of the vocabulary. This observation, then, is a consequence of the fact that the hierarchical expansion of nouns from the primary to the secondary level is much higher than the expansion from the secondary to the tertiary level, which means that the functional basis noun hierarchy is an unbalanced taxonomy.
5. *All 27 trends consistently show a significant increase of information content from the primary to the secondary level* (three columns under heading $\Delta I_{\text{I,II}}$ in Table 7). This observation indicates that the secondary level is more informative to the designer than the primary level. However, because of the mixed trends recorded under heading $\Delta I_{\text{II,III}}$, particularly in case of the used and reduced vocabularies, the tertiary level is not necessarily more informative to the designer than the secondary level.

Table 8 shows some more trends in information content in form of a truth table. Each instance of $I_{m,n}$ represents the information content of a function model of designation M(m, n), and $I'_{m,n}$ indicates the information gradient of the vocabulary measured on model M(m, n). Each row in the statement column contains a statement that predicts a relation between two quantities related to information content or information density. Each statement is being evaluated from the experimental results. The status of the evaluation is indicated using 1 for true and 0 for false in the three columns on the right. The fixed, used, and reduced columns indicate the types of vocabulary used for computing information content. The three symbols inside each cell, separated by commas, indicate the status of the evaluation for the Supermax hair dryer, Delta jigsaw, and Brother sewing machine function models. The trends that did not match the prediction are shaded.

6. *The proportional increase in information content from the primary to the secondary level is greater than the proportional increase from the secondary to the tertiary level* (trends 1–3 in Table 8). This observation is consistent for all three products, for all three vocabulary types, and for all three metrics. Thus, even though information contents based on the fixed vocabularies increase from the primary to the secondary to the tertiary level in all three products (observation 1), the proportional increase gradually diminishes for all types of vocabularies in all products, the largest jump being in the downward translation from the primary to the secondary level of both verbs and nouns. This observation supports from a different viewpoint the analysis of observation 5 that the secondary level is the most useful level in the functional basis.

Table 8. *Truth table of trends in information content*

Trend No.	Statement	Fixed	Used	Reduced
1	$\frac{I_{2,0}}{I_{1,0}} > \frac{I_{3,0}}{I_{2,0}}$	1, 1, 1	1, 1, 1	1, 1, 1
2	$\frac{I_{0,2}}{I_{0,1}} > \frac{I_{0,3}}{I_{0,2}}$	1, 1, 1	1, 1, 1	1, 1, 1
3	$\frac{I_{2,2}}{I_{1,1}} > \frac{I_{3,3}}{I_{2,2}}$	1, 1, 1	1, 1, 1	1, 1, 1
4	$I'_{1,0} > I'_{2,0} > I'_{3,0}$	1, 1, 1	1, 1, 1	1, 1, 1
5	$I'_{0,1} > I'_{0,2} > I'_{0,3}$	1, 1, 1		
6	$I'_{1,1} > I'_{2,2} > I'_{3,3}$	1, 1, 1	1, 1, 1	
7	$I_{1,1} > I_{0,1} + I_{1,0}$	1, 1, 1	1, 1, 1	1, 1, 1
8	$I_{2,2} > I_{0,2} + I_{2,0}$	1, 1, 1	1, 1, 1	1, 1, 1
9	$I_{3,3} > I_{0,3} + I_{3,0}$	1, 1, 1	1, 1, 1	1, 1, 1

7. *The information density based on the fixed vocabularies reduces from the primary to the secondary to the tertiary level* (trends 4–6 in Table 8). For example, in the case of the Supermax hair dryer, the density of the fixed verbs vocabulary for the primary, secondary, and tertiary levels is 0.364, 0.146, and 0.099 bits per verb. This trend indicates that the *usefulness* of a given level, in terms of *benefit* (information produced) over *cost* (size of the level), reduces with lower levels of the hierarchy. The tertiary level has the worst density.
8. *The combined information content of function models is greater than the sum of the element-wise information contents* (trends 7–9 in Table 8). This means that a combined model, described with verbs and nouns of the same hierarchical level, is more informative than the collection of two partial models, described with only verbs and only nouns of the same level. This observation is intuitively explainable, because, given the two partial models, some interpretation or value-added activity is required to join them into the combined model.

10. CONCLUSIONS

The secondary level of the functional basis vocabulary is clearly the most useful and informative level of the three, for both verbs and nouns, as collectively indicated by observations 2 through 6. The primary level is lower in information content, and therefore concluded to be less useful than the secondary, owing to too few terms to provide the necessary resolution for adequate function description. The tertiary level is problematic as it has too many redundant terms, which provide only a marginal proportional benefit over the secondary level, but at the cost of a poor information density. Thus, the tertiary level seems to provide little information benefit. In fact, in some cases, the information content actually reduces upon a downward translation from the secondary to the tertiary level, making the tertiary level more discouraging to the designer. Overall, the secondary level appears to be the most preferred of the three levels, providing a good balance between information content and information density. In previous research, an empirical study revealed that about 92% of the functional basis terms in function models within the Design Repository belong to the secondary level (Caldwell et al., 2008). This empirical observation reinforces the above conclusions, provided that the function models used in that study was constructed correctly using the functional basis.

The information metric acts as a measure of the usefulness of function models and the vocabulary, and behaves in agreement with practical expectations. It produces larger values for larger vocabularies, has a reasonable practical interpretation, satisfies the required criteria set by Information Theory research, and predicts a trend in information content that is practically reasonable (observation 8). This metric needs to be externally validated in order to test the metric's ability to reflect the judgment of designers about the practical value of function models and vocabularies. Further, it is hypothesized that the informativeness of function models is largely contributed by the topological information, discussed in Section 4. Hence, an extended metric that includes the topological information may be required. Another shortcoming of the metric is that it represents the quantity, not the quality, of information in a function model. To measure the quality of information, first a general representation schema and modeling rules for constructing function models is required.

REFERENCES

Albers, A., Matthiesen, S., Thau, S., & Alink, T. (2008). Support of design engineering activity through C&CM—temporal decomposition of design problems. *TMCE 2008 Symp*., Izmir, Turkey.

Arunajadai, S.G., Stone, R.B., & Tumer, I.Y. (2002). A framework for creating a function-based design tool for failure mode identification. *ASME 2002 Design Engineering Technical Conf. Computers and Information in Engineering Conf.*, Montreal, Canada.

Bhatta, S., Goel, A., & Prabhakar, S. (1994). Innovation in analogical design: a model-based approach. *3rd Int. Conf. Artificial Intelligence in Design (AID-94)*, Lausanne, Switzerland.

Caldwell, B.W., Sen, C., Mocko, G.M., Summers, J.D., & Fadel, G.M. (2008). Empirical examination of the functional basis and design repository. *3rd Int. Conf. Design Computing and Cognition*, Atlanta, GA.

Chandrasekaran, B. (2005). Representing function: relating functional representation and functional modeling research streams. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 19(1)*, 65–74.

Collins, J.A., Hagan, B.T., & Bratt, H.M. (1976). Failure-experience matrix—a useful design tool. *Journal of Engineering for Industry B 98(3)*, 1074–1079.

Cover, T.M., & Thomas, J.A. (2006). *Elements of Information Theory* (2nd ed.). Hoboken, NJ: Wiley–Interscience.

Deng, Y.M. (2002). Function and behavior representation in conceptual mechanical design. *Artificial Intelligence in Engineering Design, Analysis and Manufacturing 16(5)*, 343–362.

Gero, J.S. (1990). Design prototypes: a knowledge representation schema for design. *AI Magazine 11(4)*, 26–36.

Goel, A.K., & Bhatta, S.R. (2004). Use of design patterns in analogy-based design. *Advanced Engineering Informatics 18*, 85–94.

Hartley, R.V.L. (1928). Transmission of Information. *Bell System Technical Journal 7(1928)*, 535–563.

Hirtz, J., Stone, R.B., McAdams, D.A., Szykman, S., & Wood, K.L. (2002). A functional basis for engineering design: reconciling and evolving previous efforts. *Research in Engineering Design 13(2)*, 65–82.

Keuneke, A.M. (1991). Device representation—the significance of functional knowledge. *IEEE Expert 6(2)*, 22–25.

Kirschman, C.F., & Fadel, G.M. (1998). Classifying functions for mechanical design. *Journal of Mechanical Design 120(3)*, 475–482.

Kitamura, Y., Kashiwaseb, M., Fuseb, M., & Mizoguchia, R. (2004). Deployment of an ontological framework of functional design knowledge. *Advanced Engineering Informatics 18(2)*, 115–127.

Kitamura, Y., Koji, Y., & Mizoguchi, R. (2005). An ontological model of device function and its deployment for engineering knowledge sharing. *First Workshop FOMI 2005—Formal Ontologies Meet Industry*, Castelnuovo del Garda (VR), Italy.

Kitamura, Y., & Mizoguchi, R. (2003). Ontology-based description of functional design knowledge and its use in a functional way server. *Expert Systems With Applications 24*, 153–166.

Kurfman, M.A., Stone, R.B., Rajan, J.R., & Wood, K.L. (2001). Functional modeling experimental studies. *ASME Design Engineering Technical Conf.*, Pittsburgh, PA.

Kurtoglu, T., Campbell, M.I., Gonzales, J., Bryant, C.R., & Stone, R.B. (2005). Capturing empirically derived design knowledge for creating conceptual design configurations. *2005 ASME Int. Design Engineering and Technical Conf.*, Long Beach, CA.

Mocko, G.M., Summers, J.D., Fadel, G.M., Teegavarapu, S., Maier, J.R.A., & Ezhilan, T. (2007). A modelling scheme for capturing and analyzing multi-domain design information: a hair dryer design example. *16th Int. Conf. Engineering Design*, Paris.

Otto, K.N., & Wood, K.L. (2001). *Product Design Techniques in Reverse Engineering and New Product Development*. Upper Saddle River, NJ: Prentice Hall.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K.H. (2007). *Engineering Design: A Systematic Approach*. London: Springer–Verlag London Limited.

Rodenacker, W. (1971). *Methodisches Konstruieren*. Berlin: Springer–Verlag.

Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical Journal 27*, 379–423, 623–656.

Sridharan, P., & Campbell, M.I. (2005). A study on the grammatical construction of function structures. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 19*, 139–160.

Stone, R.B., Tumer, I.Y., & Stock, M.E. (2005). Linking product functionality to historic failures to improve failure analysis in design. *Research in Engineering Design 16(2)*, 96–108.

Stone, R.B., & Wood, K.L. (2000). Development of a functional basis for design. *Journal of Mechanical Design 122(4)*, 359–370.

Summers, J.D., & Ameri, F. (2008). An algorithm for assessing design complexity through a connectivity view. *Proc. TMCE 2008*, Izmir, Turkey.

Summers, J.D., & Shah, J.J. (2003). Developing measured of complexity for engineering design. *ASME 2003 Design Engineering Technical Conf. Computers and Information in Engineering Conf.*, Chicago.

Szykman, S., Racz, J.W., & Sriram, R.D. (1999). The representation of function in computer-based design. *1999 ASME Design Engineering Technical Conf.*, Las Vegas, NV.

Tumer, I.Y., & Stone, R.B. (2001). Analytical methods to evaluate failure potential during high-risk component development. *2001 ASME Design Engineering Technical Conf.*, Pittsburgh, PA.

Ullman, D.G. (1992). *The Mechanical Design Process*. New York: McGraw–Hill.

Umeda, Y., & Tomiyama, T. (1995). FBS modeling: modeling scheme of function for conceptual design. *9th Int. Workshop on Qualitative Reasoning*.

Vucovich, J., Bhardwaj, N., Ho, H.-H., Ramakrishna, M., Thakur, M., & Stone, R. (2006). Concept generation algorithms for repository-based early design. *ASME 2006 Int. Design Engineering Technical Conf. Computers and Information in Engineering Conf.*, Philadelphia, PA.

Chiradeep Sen is a PhD student at Clemson University. He obtained his MS in mechanical engineering from Clemson University in 2009, working on information content and formal representation of function structures. He received his BS in mechanical engineering in 1995 from Jadavpur University, India. He worked as a product designer and tool designer for Philips India Limited and as a design automation consultant for Dresser–Rand Corporation, Olean, NY. Chiradeep's areas of interest include the representation of product functionality, information content and expressiveness of design representations, complexity and uncertainty in design, design automation and knowledge-based engineering, and information modeling.

Benjamin W. Caldwell is a PhD student at Clemson University. He obtained his MS in mechanical engineering from Clemson University in 2009, working on functional similarity and formal representation of functional models in engineering design. He obtained his BS in mechanical engineering from Clemson University in 2007. His research interests include mechanical functions and forms, functional similarity, user interaction with the product, and affordances.

Joshua D. Summers is an Associate Professor in mechanical engineering at Clemson University, where he leads the Automation in Design Group. Dr. Summers earned his PhD in mechanical engineering from Arizona State University in 2004 while researching design automation under the direction of Dr. Jami Shah in the Design Automation Lab. He received his BSME and MSME from the University of Missouri–Columbia. He worked at the Naval Research Laboratory and served on the Foreign Relations/Armed Services staff of Senator John D. Ashcroft. Dr. Summers' areas of interest include collaborative design, knowledge management, and design automation with the objective of improving design through collaboration and computation.

Gregory M. Mocko is an Assistant Professor of mechanical engineering at Clemson University (since 2006). He obtained his PhD degree from Georgia Institute of Technology in 2006 in mechanical engineering, researching on engineering information management for multiobjective decision modeling with emphasis on the development of formal languages to facilitate information exchange. He received his MS degree in mechanical engineering from Oregon State University (2001) and his BS degree in mechanical engineering and material science from the University of Connecticut (1999). Dr. Mocko worked at the National Institute of Standards and Technology (NIST) where he performed research on the integration of design and analysis.

APPENDIX A: INFORMATION CONTENT OF DELTA JIGSAW FUNCTION MODEL

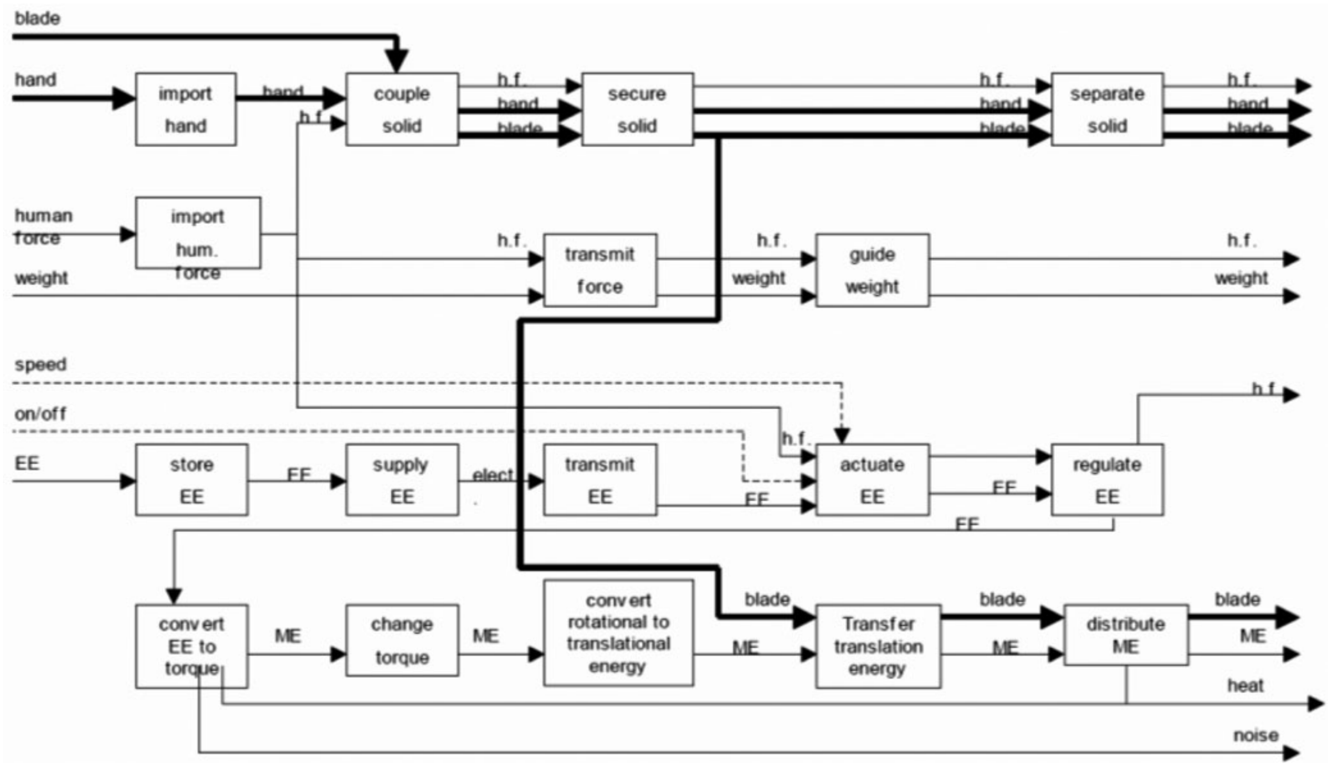

Fig. A.1. Function model of Delta jigsaw (http://repository.designengineeringlab.org/, accessed October 13, 2008).

Table A.1. *Results of information content of Delta jigsaw function model*

	Verb Levels											
	0			1			2			3		
Noun Levels	F	U	R	F	U	R	F	U	R	F	U	R
3	252	168	168							413	295	354
2	210	126	210				354	295	354			
1	84	84	84	236	236	236						
0				51	51	51	85	68	85	102	68	85

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

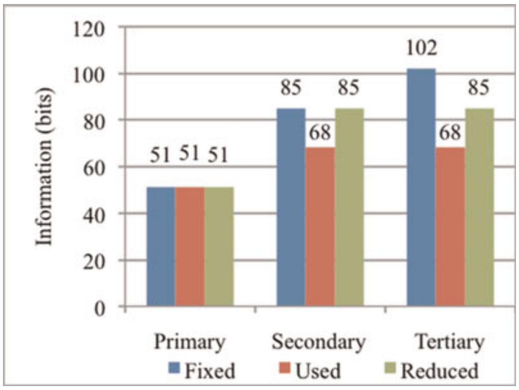

Fig. A.2. Delta jigsaw information content, verbs only: M(1, 0), M(2, 0), M(3, 0). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

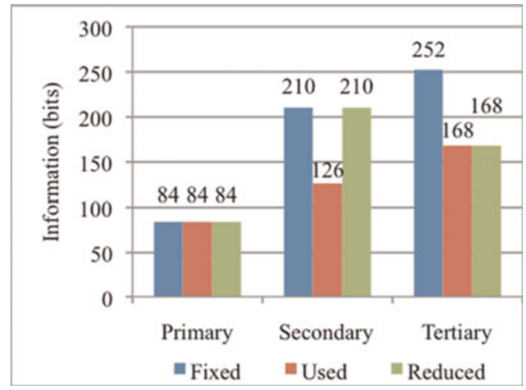

Fig. A.3. Delta jigsaw information content, nouns only: M(0, 1), M(0, 2), M(0, 3). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

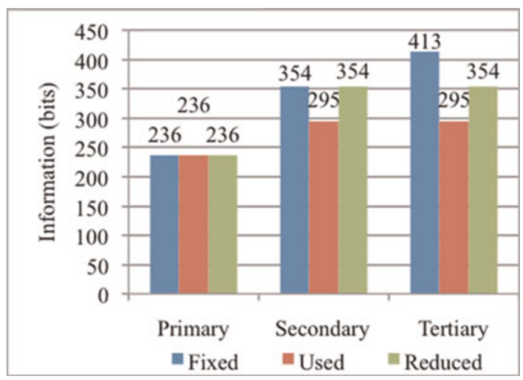

Fig. A.4. Delta jigsaw information content, combined: M(1, 1), M(2, 2), M(3, 3). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

APPENDIX B: INFORMATION CONTENT OF BROTHER SEWING MACHINE FUNCTION MODEL

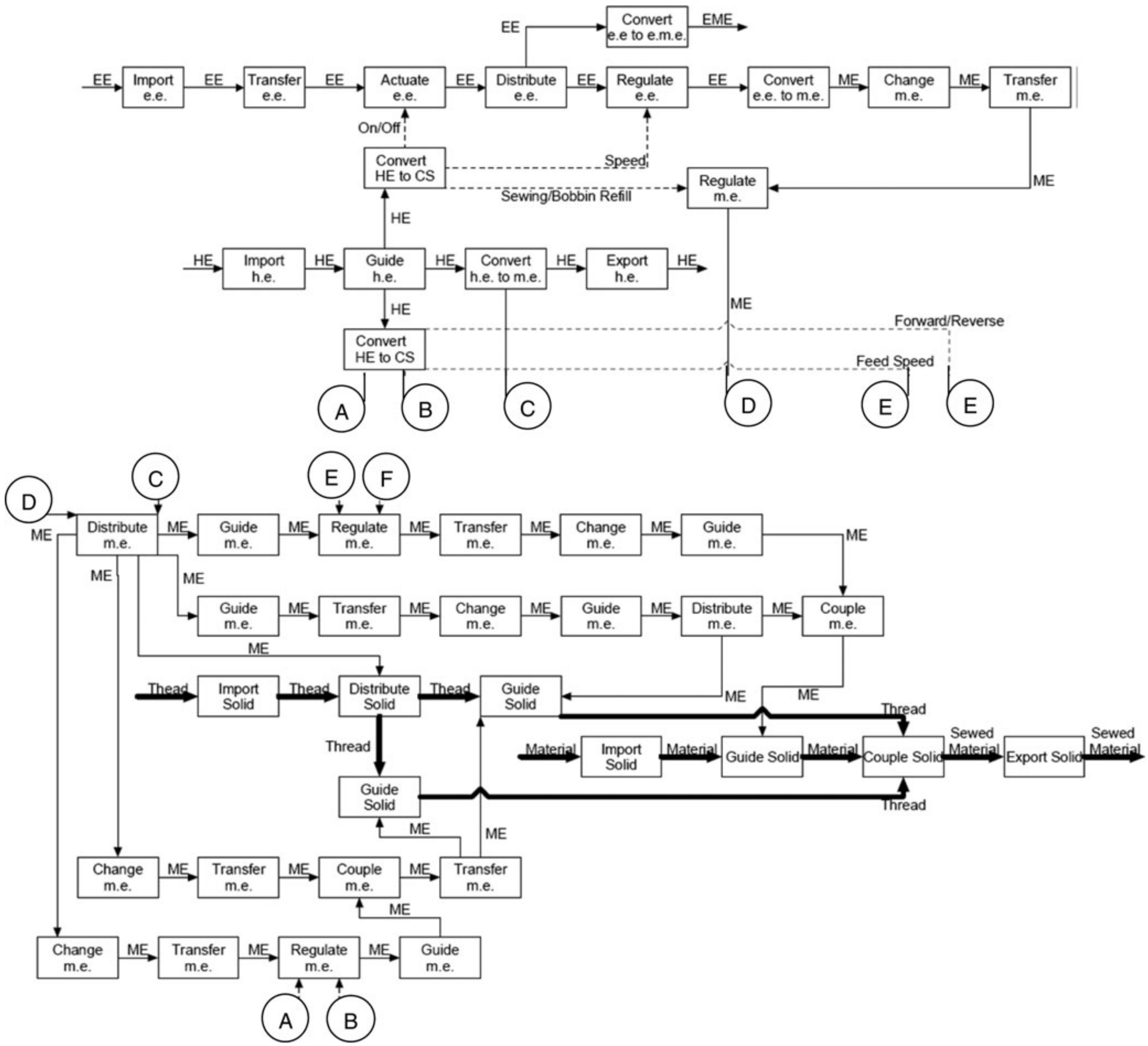

Fig. B.1. Function model of Brother sewing machine (http://repository.designengineeringlab.org/, accessed October 13, 2008).

Table B.1. *Result of information content of Brother sewing machine function model*

	Verb Levels											
	0			1			2			3		
Noun Levels	F	U	R	F	U	R	F	U	R	F	U	R
3	384	256	256							756	540	648
2	320	192	320				648	540	648			
1	128	128	128	432	324	432						
0				132	132	132	220	176	176	264	176	220

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press

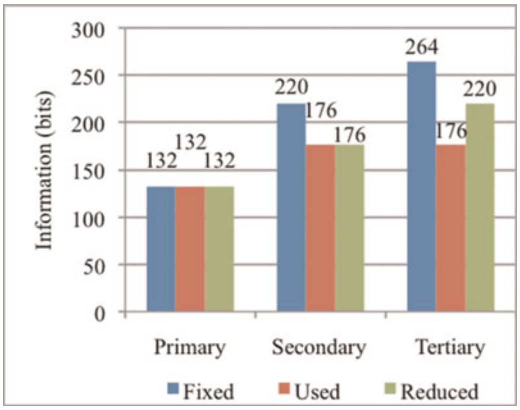

Fig. B.2. Brother sewing machine information content, verbs only: M(1, 0), M(2, 0), M(3, 0). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

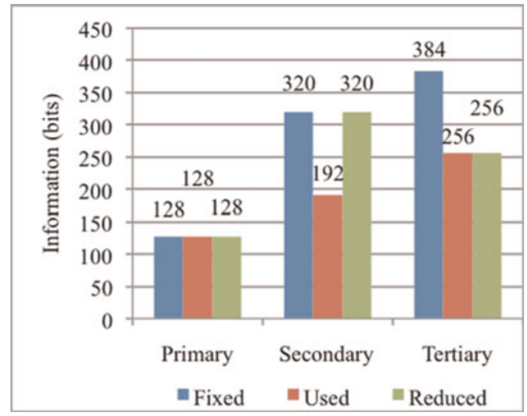

Fig. B.3. Brother sewing machine information content, nouns only: M(0, 1), M(0, 2), M(0, 3). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

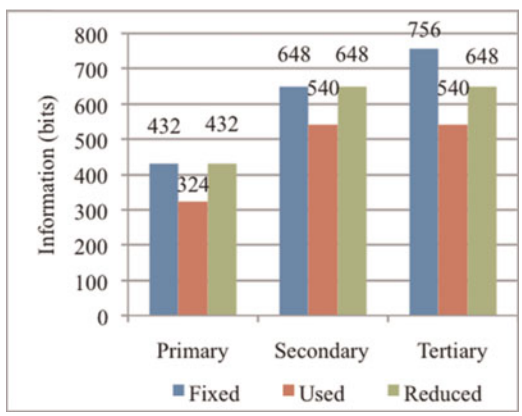

Fig. B.4. Brother sewing machine information content, combined: M(1, 1), M(2, 2), M(3, 3). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

https://doi.org/10.1017/S0890060409990187 Published online by Cambridge University Press