Artificial Intelligence for Engineering Design, Analysis and Manufacturing (2005), *19*, 75–88. Printed in the USA.

DOI: 10.1017/S0890060405050080

Functional reasoning theories: Problems and perspectives

BEHROUZ HOMAYOUN FAR AND A. HALIM ELAMY
Department of Electrical and Computer Engineering, Faculty of Engineering, University of Calgary,
2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada

(RECEIVED May 19, 2004; ACCEPTED January 24, 2005)

Abstract

Functional reasoning (FR) enables people to derive and explain function of artifacts in a goal-oriented manner. FR has been studied and employed in various disciplines, including philosophy, biology, sociology, and engineering design, and enhanced by the techniques borrowed from computer science and artificial intelligence. The outcome of FR research has been applied to engineering design, planning, explanation, and learning. A typical FR system in engineering design usually incorporates representational mechanisms of function concept together with description mechanisms of state, structure, or behavior, and explanations and reasoning mechanisms to derive and explain functions. As for representation, philosophers have long argued whether function of an artifact is a genuine property of it. As for explanation and reasoning, they have produced theories for functional ascription by an external viewer as part of an explanation. To build an FR-based system, the theory based on which the system is built and the underlying assumptions must be explicitly identified. This point is not always clear in the engineering of FR-based systems. Understanding the underlying assumptions, logical formulation, and limitations of FR theories will help developers assessing their systems correctly. The purpose of this paper is to review various FR theories and their underlying assumptions and limitations. This later serves as a benchmark for comparing various FR techniques.

Keywords: Functional Explanation; Functional Reasoning; Logic; Teleology

1. INTRODUCTION

Function is an activity by which an artifact fulfills its purpose, according to the *Oxford English Dictionary* (Oxford University Press, 2003). Functional reasoning (FR) is a collective term for a variety of theories and techniques to explain and derive functions of artifacts. FR techniques usually offer a representation scheme for describing the artifacts and an interface method to infer and explain the artifacts' functions and how they can contribute to the functionality of a containing system.

FR has been studied within a variety of disciplines such as philosophy, biology, sociology, and engineering design; it has been enhanced by the techniques borrowed from computer science and artificial intelligence (AI); and the outcome of FR is applied to engineering design, planning, explanation, and learning, among others.

FR-based systems vary mainly depending on the area of study; that is, commonsense reasoning, planning, image understanding, fault diagnosis, and computer-aided design (CAD); ontological primitives; representation schemes of structures or functions; initial data (formal description of artifacts physical structure and/or behavior); focus of study; and particular problems. A survey of some important works with focus on certain research areas can be found in the following:

- general survey: Far (1992), Umeda and Tomiyama (1997), and Chittaro and Kumar (1998);
- FR modeling survey: Chakrabarti and Bligh (2001);
- planning and conceptualization approaches: Tezza and Trucco (1988);
- explanation-based approaches, qualitative kinematics: Faltings (1987, 1990) and Kara and Stahovich (2002);
- explanation-based approaches, diagnosis: Sembugamoorthy and Chandrasekaran (1986) and Fink and Lusth (1987);
- design verification approaches: Murakami and Nakajima (1988);

Reprint requests to: Behrouz Homayoun Far, Department of Electrical and Computer Engineering, Faculty of Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada. E-mail: far@ucalgary.ca

- design approaches: *AI Magazine* (1990) and Chandrasekaran (1990);
- Special Issues: *IEEE Expert* (e.g., Sticklen & Bond, 1991), *International Journal of Applied AI* (1994), *AIEDAM* (1996), and *IEEE Intelligent Systems* (1997); and
- workshops: Florida Artificial Intelligence Research Symposium special tracks on reasoning about function (1995–).

Typical FR-based systems with focus on engineering design define a function–behavior–structure (or state) framework and usually concentrate on three problem domains:

- functional design of mechanical systems, for example, Stanfill (1983), Gelsey (1987), Murakami and Nakajima (1988), Pu and Badler (1988), Ulrich and Seering (1988), Qian and Gero (1996), Umeda and Tomiyama (1996), and Deng (2002);
- explaining function of assembled systems, for example, Freeman and Newell (1971), DeKleer (1984), Lind (1988), Bradshaw and Young (1991), Franke (1991), and Pegah et al. (1993); and
- fault diagnosis, for example, Fink (1985), Sembugamoorthy and Chandrasekaran (1986), Fink and Lusth (1987), Murakami and Nakajima (1988), Abu-Hanna et al. (1991), Far and Nakamichi (1993), and Russumanno and Bonnell (1996).

A recent trend in FR research in engineering design is manifested by splitting and decomposing the function, purpose and goal concepts and specifying them in detail. There is a quite number of such proposals, such as Deng's *purpose function* and *action function* (Deng, 2002), Chittaro's *purposive* and *operational* functions (Chittaro & Kumar, 1998) and Chandrasekaran's *function as effect* and *function as role* (Chandrasekaran & Josephson, 2000). Similar concepts, such as *abstract* and *concrete requirements*, *analysis mechanisms*, and *design mechanisms* have emerged in other disciplines, such as software requirement engineering and software design.

Similar to the other emerging research areas, developers of FR-based systems usually start with defining their own set of underlying concepts and build their technology and prototypes on top of it. Some typical examples are Chandrasekaran's environment-centric viewpoint of *function as effect* and device-centric viewpoint of *function as role* (Chandrasekaren & Josephson, 2000), Deng's *purpose* and *action* function (Deng, 2002), Hubka's *purposive* and *working* functions (Hubka & Eder, 2001), Far's function as an interpretation of persistence or an order in the sequence of system states (Far, 1999), Qian's function–behavior–structure (Qian & Gero, 1996), and Umeda's function–behavior–state model (Umeda & Tomiyama, 1996). This has lead to multiple interpretation, overlapping and partially conflicting FR techniques.

For many centuries, philosophers have argued for and against the idea that function of an artifact is a genuine property of it. Some have described function as an ascription by an external viewer as part of an explanation, and some have advocated for function as an abstraction of structure and behavior. Basically, the assumptions and theories behind the FR techniques and systems are not usually clear. The intended purpose of this paper is to present a comparative survey of theories of FR and their assumptions and limitations. This may later serve as a benchmark for comparing various FR techniques. To the best of our knowledge, although there are several surveys of FR techniques and systems, no comprehensive survey of theories exists.

This paper is structured as follows: Section 2 accounts for the FR's basic definitions and assumptions. In Section 3, a brief history of FR is presented. Common problems that FR theories should tackle are identified in Section 4. A survey of theories follows in Section 5. In Section 6 a brief discussion on the basic assumptions of FR and its orientation is presented. Finally, a summary is given in Section 7.

2. DEFINITIONS

2.1. Function

The term "function" has a multilateral spectrum of meanings. In mathematics it refers to "an expression which contains a variable term and whose meaning or truth is determined when concrete values of the variable are specified" (Merriam–Webster, 2002). In sociology, function is defined as "an individual or an organizational unit performing a group of related acts and processes" (Merriam–Webster, 2002). In software engineering, function refers to a part of the requirements that "specify the software functionality that the developers must build into the product to enable users to accomplish their tasks" (Wiegers, 2003). In engineering design, function is defined as "an activity by which a thing fulfills its purpose" (Oxford University Press, 2003). In AI, function is usually mentioned along with the terms *behavior*, *goal*, and *purpose* with respect to system's inner and outer environments (Simon, 1969). In addition, it has strong connections with the notion of making efforts to obtain a certain result (mainly in engineering design), a certain future event (Bigelow & Pargetter, 1987), or to the notion of something *good* (e.g., survival in natural system or efficiency for designed artifacts; Sorabji, 1964).

Typical definitions of function in engineering design are: *function as intended behavior* and *function as purpose* (Chakrabarti, 1998) and typical viewpoints are *objective* and *subjective* viewpoints. The following are a few common definitions:

> The word function is regarded as a description of the action or effect required by a design problem, or that supplied by a solution. A functional representation, there-

> fore, should allow one to describe design problems and solutions in terms of their functions. (Chakrabarti & Bligh, 2001)
>
> Function is a source of knowledge that abstracts behavior. Function of a component can be defined as operational, i.e., a relation between the input and output in the component; or purposive, i.e., a relation between the goal of a human user and the behavior of the component. (Chittaro & Kumar, 1998)
>
> Function of a system is its intended purpose. The functional specification describes the system's goals at a level of abstraction that is of interest at the system level. (Keuneke, 1991)
>
> Function of a mechanical object is dependent on (and derived from modeling and simulation of) the way that motion and forces are transmitted through the contacts between parts. (Faltings, 1990)
>
> Function (of a mechanical assembly) is defined with: (a) transformation between states of physical quantities and substances; and (b) physical features that describe the relation between a physical structure and functions indirectly. The function of an assembly is derived as causalities of transformation, using physical features. (Murakami & Nakajima, 1988)
>
> Function is a relation between the goal of a human user and the behavior of a system. In an assembly, the function of a component relates the behavior of that component to the function of the assembly. (Bobrow, 1984)
>
> Function is the purpose of the system as described by the human user. Function of a system (e.g., electronic circuit) is derived from its behavior and expresses with the technical terms of the domain that it is applied to (e.g., latching, amplification, etc.). (DeKleer, 1984)

In the objective viewpoint of function, a goal describes some outcome toward which certain activities of a system or of its components are directed. It is argued that the goal and function can be used interchangeably, depending on the way of viewing the system (or a part of it) and where to put the boundary. Looking at the system externally, the effect will be regarded as a functional ascription. However, from the perspective of the system itself, it can be considered as a goal that guides the organization of resources internal to the system (Lind, 1988). Some have differentiated between the goal and function concepts, arguing that although sometimes the end product of a goal directed processes is a function, it is not necessarily so (Nagel, 1977*a*), and even the function may be different from the achievement of goals (Wright, 1973).

In the subjective viewpoint of function, the function of a system is addressed with reference to the external intention of humans. The term *intention* is usually used in the narrow sense of a kind of *plan* that includes a structural and behavioral representation of a system and its future effects. In this sense, *function* and *behavior* of a system are closely related.

In some works, functions are classified as: *conscious functions*, that is, functions with respect to a conscious designer or creator; and *natural functions*, that is, functions build naturally into a system (see Section 3.1). Some schools of thoughts have argued against such distinction. For instance, evolutionary Darwinism has denied entities having natural function.

2.2. FR

FR is a collective term for a variety of theories and techniques that enable people to explain the presence and function of artifacts in a containing system; to derive the purpose of the artifact, and to explain how the function can be achieved. FR is sometimes called *functional analysis* (FA; mainly in sociology and engineering design) or *functional explanation* (FE; mainly in biology). The ultimate goal of FR is enhancing the commonsense reasoning with the functional ability.

FR as a commonsense theory usually consists of three parts:

1. Ontology describing the domain and the entities in the domain;
2. Representation scheme for modeling the entities and their interactions;
3. Reasoning method for inferring and explaining how the entities function.

Three typical definitions of FR in engineering design are the following:

> Functional reasoning is the technology that adds functional concepts into model-based reasoning technology (MBR). MBR technology reasons out a device's behavior (what a device does) from explicitly represented models of the device. Functional reasoning technology, in contrast, deals with what the device is for. (Umeda & Tomiyama, 1997)
>
> Functional explanation often takes the form of decomposition of complex systems. This consists in describing a system in terms of what it does, and then explaining its behavior in terms of what its (functionally defined) components do. (DeJong, 2003)
>
> Functional reasoning, a sub-field of model-based reasoning, uses abstractions of a device's purpose to index behaviors that achieve that purpose. (Pegah et al., 1993)

2.3. FR-based system

An FR-based system is an implementation of one or more FR theories in a computer program. A typical FR-based system incorporates the following:

1. Representational mechanisms of function and/or goal (purpose) concepts. For example, a database of artifacts and their functions.
2. Description mechanisms of state, structure, or behavior. For example, a modeling and model-based simulation tool.
3. Inference (i.e., explanation and reasoning) mechanisms to derive and explain functions based on an FR theory.
4. Presentation mechanisms. For example, a graphical user interface.

3. HISTORY OF FR THEORIES

3.1. FR in biology and sociology

FE is extensively used in biology, sociology and engineering design. (Some works seek to unify the explanation of function in biology, sociology, and engineering design, such as Beckner, 1969.) Originally, FR theories were devoted to explain the presence of entities in a containing system (Hempel, 1959; Cummins, 1974; Nagel, 1977*a*, 1977*b*). The containing system is a living organism, an organization, or a designed artifact. In biology, FR tries to answer questions, such as "why does a giraffe have a long neck?" and tries to discover the function of an organ in an organism, such as the function of the heart in the human body. These are usually called *natural functions*.

In sociology, FR is used to explore the necessary conditions of existence of a social system and in a structural–functional approach that employs the concept of function as a link between relatively stable structural categories. FR research in sociology is governed by Malinkowski's principle: "All social phenomena have beneficial consequences (intended or unintended, recognized or unrecognized) that explains them" and a weaker version of this principle was mentioned by Merton: "Whenever social phenomena have consequences that are beneficial, unintended and unrecognized, they can also be explained by these consequences." (Elster, 1983). Some researchers have argued against the use of FR in sociology in the very same way as one explains the biological phenomenon. The reason is that in biology the optimal consequence is much stronger than the beneficial consequence in sociology.

Plato and Aristotole were among the earliest philosophers talking about functions. Plato described the function of an item conferring to some *good*. This idea still exists in some works such as Sorabji's natural functions connected with the notion of something *good* (Sorabji, 1964), or Canfield's explaining function by its *usefulness* to the containing system (Canfield, 1964). Later, philosophers from Spinoza to those of the late nineteenth century were engaged with explaining the design into nature using teleological notions of *means* and *ends* (Allan, 1952).

Among the recent works, apart from the Beckner's theory of FE using positive and negative evidences (Beckner, 1969), the rest of the FR theories are either derivations or reformulations of the works by Hempel and Nagel (Hempel, 1959; Nagel, 1977*b*). Among the followers are Lehman (1965), Ayala (1970), and Ruse (1971).

Hempel (1959) provided an analysis of functional ascription in terms of sufficient conditions. In contrast, Nagel (1977*b*) tried to specify the necessary conditions. These two attempts were somehow problematic in scientific terms. As Cummins (1974) mentioned, "Any analysis in terms of sufficient conditions may lead to a schema with true premises but invalid, and any formulation specifying necessary conditions may yield to a valid but unsound explanation."

Two other works are worthy of mention: Wright's etiological theory (1973, 1976), and Cummins' (1974) functional ascription. In Wright's etiological theory of function the unification of functional and causal explanations is the central idea (Wright, 1973, 1976). According to this theory the function of an entity is explainable in terms of its history, not its present behavior. For instance, when etiologists define the function of the kidneys in human body, they would take how it is evolved to function and ignore what it does at present. Therefore, it is possible to make distinction between success by function and success by accident. The genetic algorithm is considered as a computational model for etiological theory of function.

Cummins (1975) argued against the validity of the underlying assumptions of traditional FEs and suggested an alternative scheme: functional ascription to an item is ascribing a *capacity* to the item that can be recognized by its role in an analysis of some capacity of a containing system. These theories are reviewed in Section 5.

3.2. FR techniques in engineering design

In engineering design, there is an implicit assumption that even the most fanciful assemblies (e.g., buildings, devices, hardware, and software) have practical functions to fulfill. The structure–behavior–function framework is considered as one of the dominant methods for analysis of engineering artifacts. Cases such as function of physical components (e.g., a pedal) in a designed artifact (e.g., a bicycle) are discussed and functional decomposition is a popular method for designing real (e.g., an electric power plant) and virtual (e.g., software) artifacts.

Although software system designers developed the concept of functional decomposition and structural analysis in the 1960s, the idea of using functional concept in engineering design, the AI way, was first mentioned by Herbert Simon with respect to system's inner and outer environment (Simon, 1969). The idea was later presented by Freeman and Newell (1971). Recent advances in AI, computation theories, and distributed systems have led to new interpretation and implementation of the FR theories in programs. In typical systems, the initially given data consists of artifacts (objects, processes, or mechanisms), and a formal or semiformal description of their physical structure, behavior or functions. The outcome is describing and explaining the function of an artifact in terms of the structure or behavior of its components and their functions. These are mainly inspired by the Beckner's theory (1969; first-generation FR systems), Cummins' (1974) analytical explanation and the capacity concept, Nagel's causal/FE of goal-directed processes (1977*a*; second-generation systems) and Canfield's usability concept (1964; modern FR systems). There is also a shift of attention from justification of the theory to practical implementations. (See Section 5 for details of the FR theories.)

FR-based systems in engineering design can be classified in three groups:

- *Planning and design approaches:* a method for design verification (Murakami & Nakajima, 1988); a method for capturing qualitative design knowledge (Pu & Badler, 1988); a method for hierarchical design using functional knowledge (Acar & Ozguner, 1990); a design evaluation using functional knowledge (Bradshaw & Young, 1991); a scheme for FR in conceptual design (Chakrabarti & Bligh, 2001); and a function and behavior representation in conceptual mechanical design (Deng, 2002).
- *Conceptualization approaches:* scripts, plans, goals, and understanding (Shank & Abelson, 1977); a method for functional representation and compilation (Sembugamoorthy & Chandrasekaran, 1986); a functional representation through consolidation (Bylander, 1988); a representation and planning method (Tezza & Trucco, 1988); temporal and cohesive clustering of functional knowledge (Shekar, 1990); a functional representation of mechanical devices (Keuneke, 1991); a qualitative function formation technique (Far, 1999); a scheme for FR in conceptual design (Chakrabarti & Bligh, 2001); and a function and behavior representation in conceptual mechanical design (Deng, 2002).
- *Explanation-based approaches:* an explanation-based method for electronic circuits (DeKleer & Brown, 1984); explanation-based methods for fault diagnosis (Fink, 1985; Fink & Lusth, 1987); explanation-based methods for higher order mechanical devices (Faltings, 1987, 1990); an explanation of serial assemblies (Dormoy & Raiman, 1988); an explanation-based method for mechanical devices (Joskowicz & Addanki, 1988); an explanation-based method for fault diagnosis using hierarchical knowledge (Abu-Hanna et al., 1991); teleological descriptions for designed artifacts (Franke, 1991); FR in failure modes and effect analysis (Russumanno & Bonnell, 1996); and a qualitative function formation technique (Far, 1999).

Planning and design approaches (sometimes called CAD approaches) take advantage of the representational FR theories. In this case, a representation of a functional concept exists prior to any realization of the object having such a function, and such a representation contributes to the process of bringing that object into being (Bigelow & Pargetter, 1987). Planning problem is devising a plan that can achieve some functions while satisfying some constraints. Planning approaches have a finite set of symbols, standing for activities, and a finite set of rules showing the possible interactions between activities. The problems considered are composition, decomposition, and verification of the plan. Design problem is devising a device that can achieve some functions while satisfying some constraints. Design approaches have a finite set of symbols, standing for components, and a finite set of rules showing the possible interactions among components. The problems considered are composition, decomposition, and verification of the design. A common limitation of planning and design approaches is that they can only deal with the entities falling within their defined symbol set. Although being good and efficient, they can at most support the user through providing a more abstract (i.e., higher level) planning (or design) environment, more useful than detailed planning (or geometric design), leading to an increase of the quality and efficiency of planning (or design) task.

Conceptualization approaches suggest a hierarchical classification scheme for the functional concepts, define classes objectively, and aggregate objects into classes. The class types are defined by *functional primitives*. The necessity and sufficiency of the primitives, and whether they are appropriate for functional representation in terms of means–ends hierarchy (Rasmussen, 1985, 1990) is somehow doubtful. A main problem is that almost all of the methods try to define the primitives objectively: assign meaning to the behavior of the objects at the first place, and then recover it as a function.

In explanation-based approaches traces of qualitative reasoning in explaining function of artifacts in pioneer FR systems can be found along with the three major theories of qualitative physics, that is, the qualitative process theory (Forbus, 1984) influenced deriving function for mechanical devices (see Faltings, 1990; Kara & Stahovich, 2002), qualitative confluence theory (DeKleer & Brown, 1984) has influenced explanation of function of electronic circuits (see DeKleer, 1984) and qualitative simulation (Kuipers, 1986) has led to explaining function of designed artifacts (see Franke, 1991). There exists an analogy between the explanation-based FR systems and explanation-based learning (EBL) techniques (Ellman, 1989). The above three FR methods each resemble a kind of EBL using either chunking or generalization.

3.2.1. First-generation FR-based systems: Direct match

The first-generation systems using functional knowledge start with either a semiformal description of physical structure (design verification approaches) or a description of shape (conceptualization approaches). In addition, systems

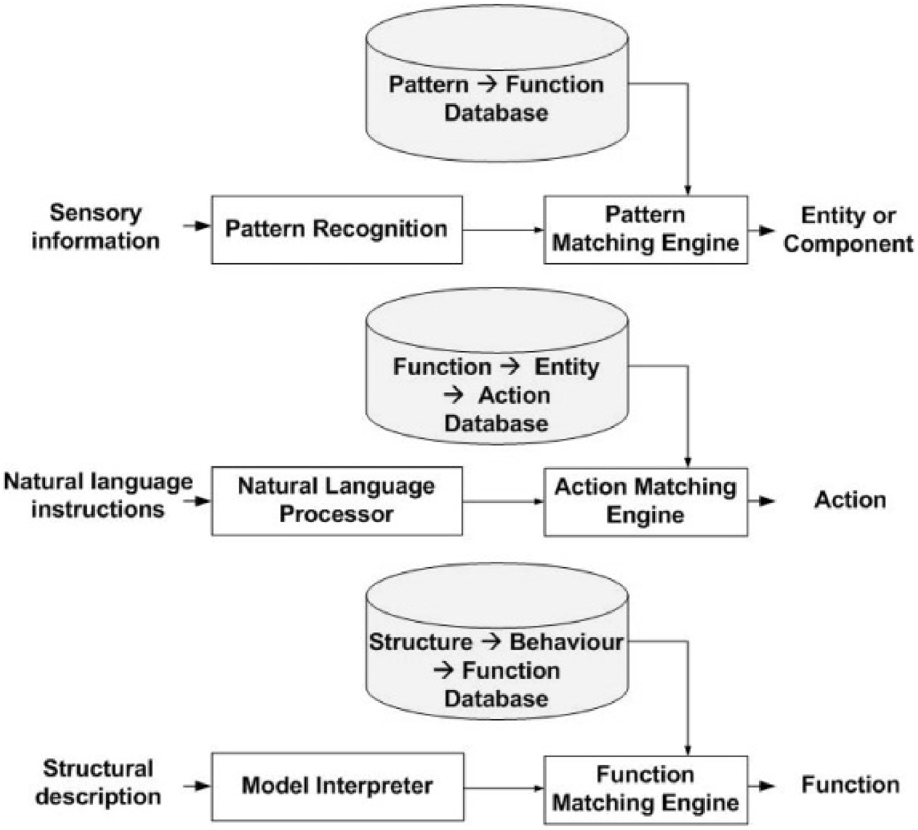

Fig. 1. First-generation functional reasoning systems.

starting with natural language instructions have been reported (Asai et al., 1990). Figure 1 shows the basic building blocks of the first-generation systems. They process input data and relate it to a functional concept that has already been recorded in the database. The functions in the database can either be rigid symbolic names for a property of a given artifact, or include some attributes filled by the data measured or interpreted from the real world. Being good as they are, none of those first-generation systems can assign several functions to an artifact or provide solution to all of the FR problems (see Section 4). The main drawback of the first-generation systems is the restricted view of the direct list matching inferences. All the artifact and functions are identified in advance, and the essence is recorded in one or more of the three-structure function databases.

3.2.2. Second-generation FR-based systems: MBR

Second-generation FR-based systems were developed based on representational FR theories, in which FE can be derived from a causal account of system's structure (or state) and behavior, and offer more flexibility through employing a kind of MBR approach. There are several methods suggested for the model-based approach to assign functions to physical structures (DeKleer, 1984; Pu & Badler, 1988; Tezza & Trucco, 1988; Faltings, 1990; Abu-Hanna et al., 1991; Franke, 1991; Qian & Gero, 1996; Umeda & Tomiyama, 1996; Deng, 2002). They all relate a semiformal description of the physical structure of a system to its function. Using qualitative simulation to derive the behavior from structure, causal, and FR to explain how such behavior is achieved and to derive functions, are typical. In such systems identification of system boundary is extremely important. The environment and interaction with the environment is expressed by the context (Tezza & Trucco, 1988), constraints (DeKleer, 1984), physical features (Murakami & Nakajima, 1988), or connection frames (Pu & Badler, 1988) or is specified by the human designer (Umeda & Tomiyama, 1996).

Recent works, such as Deng's (2002), define multiple views of function, such as *purpose function*, which is "a description of the designer's intention or the purpose of a design" and *action function*, which is "an abstraction of intended and useful behavior that an artifact exhibits." They are related to the different levels of design hierarchy and abstraction, have layered semantic representations, and are useful in developing conceptual design synthesis strategies (Deng, 2002).

3.2.3. Emerging FR-based systems: User-centered design

A relatively new paradigm in engineering design is designing systems to enhance usability (Constantine & Lockwood, 1999; Nielsen, 2000). This is based on the subjective viewpoint of function; that is, function of a system with reference to the goal (or intention) of humans. People employ designed systems for certain purposes and the trend is manifested by the used-centered design in industry. A necessary prerequisite for designing a system for proper use depends on understanding what the users intend to do.

In the past few years, analysts have tried to elicit user requirements and map them to user profiles and usage scenarios. A profile is a set of actions that a user can take and their relative frequencies. A scenario is a description of a single instance of usage of the system yet to be designed. Jacobson et al. (1992), Constantine and Lockwood (1999), Cockburn (2001), and others have formalized the usage-centered perspective into the *use-case* approach to requirements elicitation. However, use cases describe system behavior from a user's point of view, which may omit a lot of details that are invisible to the user. Designers need many other views to properly design and implement a system. Another way to organize and document user requirements is to identify the external events to which the system must respond in an event–response table (also called an event table or an event list; McMenamin & Palmer, 1984). An event–response table lists all such events and the behavior the system is expected to exhibit in reaction to each event.

In both cases verification and validation techniques are used to confirm whether the system requirements have been met, and that the designed system meets its predefined specifications. This answers the question "is the system built right?" and proves that the system can perform its intended mission and lives up to the user expectations, which answers the question "is the right system built?"

4. FR PROBLEMS

4.1. Informal description

Traditionally, biology, engineering design, and AI are considered as FR problems.

4.1.1. Biology

In biology, FR theories have to find answer to a set of problems, among which, why an organ (e.g., heart) is in an organism (e.g., human's body) in terms of its contribution to the functionality of the whole organism. In addition, it may be required to derive the natural function of an organ (e.g., heart for pumping blood vs. generating pulses or making sound, etc.). Finally, there are also some classes of problems requiring explanations with reference to functions (e.g., why animals in the Arctic have white fur).

4.1.2. Engineering design

In engineering design, first, FR has to explain why a component is used in a design artifact in terms of its contribution to the functionality of the whole system; second, it has to find answer to design for usability problem; and third, it must answer to the verification and validation problems (see Section 3.2.3).

4.1.3. AI

Explaining the functions of artifacts, generating understandable and sound explanation of functions with reference to common physical laws is considered as an area of study in AI, in general, and in model-based research, in particular. Among possible problem areas, action planning, functional design of artifacts, and fault diagnosis fall within the scope of FR techniques.

4.2. Classification of FR problems

FR problems can be classified into four groups: identification, explanation, selection, and verification problems (see Fig. 2).

4.2.1. Identification problem

Given a system, its function is explained using the knowledge of the structure and behavior of its components and

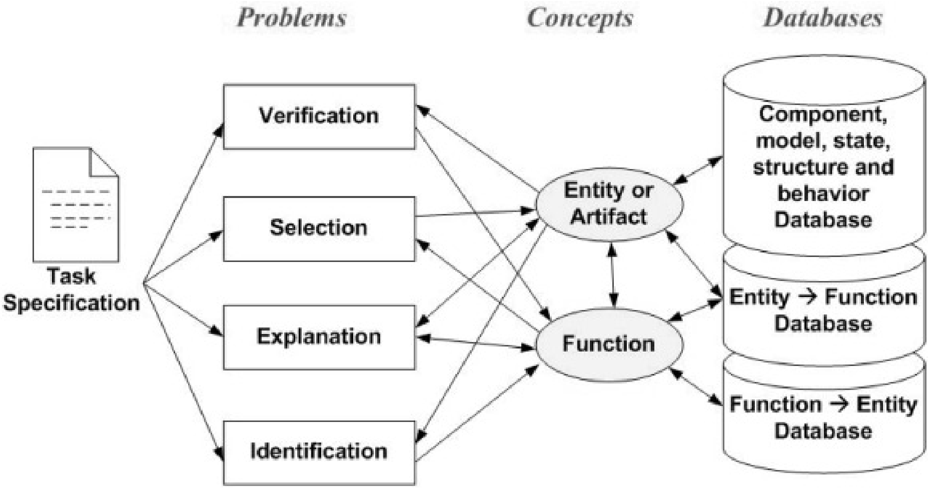

Fig. 2. Functional reasoning problems.

their organization. For example, what is the function of a pair of scissors? Typical works are Freeman and Newell (1971), DeKleer (1984), Joskowicz (1987), Dormoy and Raiman (1988), Tezza and Trucco (1988), Iwasaki (1989), Faltings (1990), and Far (1999).

4.2.2. Explanation problem

An explanation problem is explaining the presence of a component in a containing system in terms of its contribution to the overall function of the system. For example, in a software system, a user interface window is used to help the system acquire user input and display the results to the user. Typical works are Hempel (1959), Canfield (1964), Lehman (1965), Beckner (1969), Ayala (1970), Ruse (1971), Wright (1973), Cummins (1974), Nagel (1977*a*, 1977*b*), Gautier and Gruber (1993), and Far (1999).

4.2.3. Selection problem

Given a set of components, a selection problem is selecting a proper component set that, if used together, can achieve a defined function of the containing system. For example, in software system design, what should be the arrangement of objects to perform user authentication? Typical works are Freeman and Newell (1971), Stanfill (1983), Brady and Connell (1987), Gelsey (1987), Pu and Badler (1988), Qian and Gero (1996), Umeda and Tomiyama (1996), Far (1999), and Chakrabarti and Bligh (2001).

4.2.4. Verification problem

Verifying whether an item can exhibit a required function in a given situation is a verification problem. For example, can a particular software firewall protect user from identity theft? Typical works are Murakami and Nakajima (1988), Ulrich and Seering (1988), Umeda and Tomiyama (1996), Far (1999), and Chakrabarti and Bligh (2001).

FR problems can be evaluated against the abstraction hierarchy (Rasmussen 1983, 1985). In dealing with the identification and verification problems, one starts with a representation of structure and ends with a function. Selection, on the other hand, starts with a function and ends with a physical description of the item. Explanation can proceed in both directions.

5. FR THEORIES

In this section, a number of classical FR theories are reviewed. The focus is on their validity, expressive power, and engineering implementability.

5.1. Allan's theory

Allan (1952) states the following:

> For a system S, in environment E, y a valuable state of S occurs; what is x, a complex causal sequence, such that
>
> (a) if x (and other complementary conditions) occurs, then y will occur; and
>
> (b) if x, or its equivalent does not occur, y will not be brought about."
>
> This is interpreted as x, when found, occurs for the sake of y, meaning that y is more important causally or valuationally to the system S.

This is the intuitive form of functional ascription in terms of "means" (i.e., x) and "ends" (i.e., y), but not of much use in terms of validity and expressive power. The difficulty is that the above two conditions (a) and (b) do not represent a one–one mapping from the means to the ends sets. There might occur (not occur) many things other than y when x occurs (not occurs), and which one is the "end" for x is not clear. Revised versions of this theory are suggested below.

5.2. Beckner's theory

Beckner (1969) states the following: "The component $c \in C$ has function $f \in F$ in a system S if there is a set of circumstances in which f occurs when S has c, and f does not occur when S does not have c," where C is a set of components (c), S is a set of components comprising the system $(S \subset C)$, and F is a set of functions of the components (f).

This can be formulated logically as

$$\forall c \in C, \quad \exists f \in F, \quad \exists V$$
$$: \mathrm{HAS}(S, c)\text{: TRUE} \supset \mathrm{FUNCTION}(f, c)\text{: TRUE}. \quad (1)$$

$$\forall f \in F, \quad \exists c, \quad \exists V$$
$$: \mathrm{HAS}(S, c)\text{: FALSE} \supset \mathrm{FUNCTION}(f, c)\text{: FALSE}. \quad (2)$$

where V is a possible situation (in a logical sense) and FUNCTION and HAS are logical predicates.

A main criticism of this theory is that expression (2) cannot be easily verified. There might be situations $(\exists V)$ that $(\mathrm{HAS}(S,c)$; FALSE) but (FUNCTION (f,c): TRUE); and if limiting $S = C$, $\forall V$, $(\mathrm{HAS}(S,c)$: TRUE). Therefore, it is not necessary for S to have c to occur f. For instance, if the function of heart is to circulate blood in the body, it can be realized also without heart, may be using an artificial pump.

Beckner's (1969) theory is built based on the assumption that the function is a property of its host system, and the interactions with the external world are lumped together in the "circumstances." Therefore, it is difficult to use this theory to explain the function of a component other than its most frequent one. Furthermore, it cannot be used in situations when an arrangement of components is exhibiting a single function.

From the engineering implementation point of view, each component must be related to at least a function concept in the database, and each function concept has to be related to several components. Such data structures can be represented using a relational database, and would be manageable even when the number of components and associated functions grow. The only problem is that when generating FEs, every component–function pair must be associated with a list of conditions under which the component can exhibit the given function. The list may be incomplete (that is, conditions are not sufficiently specified) or become extensive (that is, too many conditions are given). Theoretically, this list can grow beyond control (what is usually called a "frame problem" in theoretical AI). Most of the first-generation FR-based systems implicitly have Beckner's (1969) theory as their underlying theory (see Section 3.2.1).

5.3. Canfield's theory

Canfield's (1964) theory states the following: "A function of the component *c* in system *S* is *f* means that *c* does *f*, and that *f* is useful to *S*."

This can be formulated logically as

$$\forall c \in C, \quad \exists f \in F, \quad \exists V$$
$$: \mathrm{HAS}(S, c)\colon \mathrm{TRUE} \supset \mathrm{FUNCTION}(f, c)\colon \mathrm{TRUE}. \quad (3)$$

$$\forall f \in F, \quad \exists c, \quad \exists V$$
$$: \mathrm{FUNCTION}(f, c)\colon \mathrm{TRUE} \supset \mathrm{USEFUL}(f, S)\colon \mathrm{TRUE}. \quad (4)$$

where *C* is a set of components (c); *S* is a set of components comprising the system ($s \subset C$); *F* is a set of functions of the components (f); *V* is a possible situation (in a logical sense); and FUNCTION, HAS, and USEFUL are logical predicates.

It is argued that this theory is difficult to apply to explain functions of designed artifacts, mainly due to difficulties in identifying the system *S*. In addition, meeting (3) and (4) is neither necessary nor sufficient for something to be a function (Wright, 1973). It is not necessary because artifacts may be "designed" to have a certain function, even if they might be useless to a particular user. There might be some cases where *c* is designed to do *f* but cannot do it, except under certain circumstances (e.g., the function of a door knob is to maintain the door closed, but in case of fault this cannot be manifested). It is not sufficient because *c* might do some other useful things also, which is not considered as its function. Canfield (1964) also believes that a function is a property of its host system.

From an engineering implementation point of view, besides those problems mentioned for Beckner's (1969) theory, verifying whether *c* can do *f* may be straightforward (i.e., through using simulation and causal reasoning), but verifying its "usefulness" to *S* is not trivial. For each association between a component and a function in the database, the enabling conditions and an additional usefulness attribute (which may depend on the enabling condition) must be specified. Some of the modern FR-based systems with usability concerns use Canfield's (1964) theory as their underlying theory (see Section 3.2.3).

5.4. Wright's theory

Wright (1973, 1976) argues that explaining "natural" and "conscious" functions should follow the same pattern. In his etiological theory, the functional and causal explanations are considered together: a functional statement "function of the component *c* in system *S* is *f*" is equivalent to asserting "component *c* in system *S* in order to do *f*." In other words, the function of *c* in *S* is *f* if and only if

1. *f* is a consequence of *c*'s presence in *S* and
2. the component *c* is in *S* because *f* is a consequence of *c*'s presence in *S*.

The first statement addresses that *f* is a causal consequence of *c*, and the second statement indicates the component *c* is selected in *S* for the sake of *f*, that is, "why the component *c* is present in *S*."

One of the benefits of this theory is the ability to make distinction between success by function and success by accident (Wright, 1973; Millikan, 1989). In biological systems statement 2 can be evaluated with reference to natural selection. Many theorists who adopted Wright's (1973, 1976) theory have also extended it by introducing natural selection into their definitions (Walsh, 1996).

There are critics to this theory. First, it may not be possible to derive *f* as a causal consequence of *c* (at least for some biological systems whose causal mechanisms are not fully explored); and second, being in *S* may not be because *f* is a consequence of *c*, but because of a "belief" that it is so (Nagel, 1977*a*, 1977*b*) and the belief is a subjective external assertion not a genuine property of the system itself.

Matthen has extended Wright's (1973, 1976) theory by introducing a functional structure among the functions via utilizing a set of facts (model) used to subordinate one function to another as means to ends (Matthen, 1988).

Wright's (1973, 1976) theory is nevertheless an important FR theory in terms of binding causal and FA and paving the way for a model-based approach to FR.

5.5. Hempel and Nagel's theories

Hempel's (1959) formulation of FE is the following:

> "The function of component *c* in a system *S* during period *T* and in environmental setting *V* is to do *f*" is equivalent to "Component *c* in system *S* during period *T* and in environment *V* has the effects *f* that satisfy the conditions *n* which are necessary for the proper working of *S*."

The explanation has the following pattern:

1. During period *T* in environment *V*, the system *S* is in proper working order.
2. If the system *S* is in proper working order, condition *n* must be satisfied.

3. If component c is present in S, then the effect of component c's presence in S satisfies the condition n.

Hempel (1959) explains his theory along with an example: "The heartbeat in vertebrates has the function of circulating blood through the organism." Heartbeats may bring about other things like heart sound that does not serve for survival of the organism, and thus not considered as its function. In other words, what a component contributes to a system needs to serve for some good (e.g., survival in biological systems) to the system to be considered as a function. In addition, the heartbeat serves for survival of the organism only if many other conditions for proper working of the system are met such as that there are no ruptures of the aorta, that there is enough oxygen in the environment in which the organism is located, and so forth.

Hempel (1959) pointed out that there are some problems with formulating FEs in accordance with his model, such as the *functional equivalence* problem, stating that what can be explained by FE is not that there is a component in a certain system at a certain time, but that there are at least one of some functionally equivalent components in a certain system at a certain time. According to his example, a FE "The heartbeat in vertebrates has the function of circulating blood through the organism" in fact explains the presence of (at least) one of functional alternatives in vertebrates that serve to circulate blood, not the presence of heart in vertebrates. In other words, this formulation is not logically sound. Statement 3 is not a necessary condition (although it is sufficient) for the performance of function. There may be some other components exhibiting the same function; therefore, the presence of c in S cannot be explained.

Nagel (1977*a*, 1977*b*) argued for changing statement 3 to a necessary condition "if and only if component c is present in S." Nagel's formulation of FE is the following:

4. During period T the system S is in environment V.
5. During period T and in stated circumstances, the system S does f.
6. If during period T the system S is in environment V, then if S performs f the component c is present in f.

Statement 6 indicates a necessary condition for performing f. Once again, a basic critique is that although the explanation is logically sound it may not be valid any more in certain cases, such as redundant components exhibiting a shared function, for example, two file servers used redundantly in a computer system.

These two theories each specify either necessary or sufficient conditions for functioning components. However, the above discussion suggest that what is often called *functionalism* is best viewed, not as a body of doctrine or theory advancing tremendously general principles such as the principle of universal functionalism, but rather as a program for research guided by certain heuristic maxims or "working hypotheses" (Hempel, 1959). In other words, FEs are useful to show scientists a course of their researches and help them discover new relationships between phenomena, but they have hardly any worth in terms of scientific explanation and prediction.

From practical applicability point of view, a one–one relation between a component and a functional concept is necessary. If such a database can be developed, the theory can be successfully used to explain functions of various components.

5.6. Cummins' theory

The basic assumptions in the conventional interpretation of functional ascription to objects are the following (Cummins, 1974):

1. The main purpose of functional characterization in commonsense world is to explain the "presence" of components that are functionally characterized.
2. The component is said to perform its function with respect to a containing system (or another component), if it affects the containing system (or another component) in the sense of either contributing to the performance of some activity of, or the maintenance of some condition in that containing system (or another component).

Cummins (1974) argues against those assumptions, and suggests an alternative FR scheme: functional ascription to an object is ascribing a "capacity" (or disposition) to the component that can be recognized by its role in an analysis of some capacity of a containing system. He believes that functional statements are actually disposition statements, and FE is actually explaining such dispositions:

If a function of component c in system S is f, then c has a disposition to f in S. To attribute a disposition d to a component c is to assert that the behavior of c is subject to (exhibits or would exhibit) a certain lawlike regularity. . . . To say that c has d is to say that c would manifest d were any of a certain range of events to occur.

> . . . Disposition requires explanation: if c has d, then c is subject to a regularity in behavior special to things having d, and such a fact needs to be explained.

In Cummins' (1974, 1975) theory, the capacity of a containing system is explained analytically by decomposing it into a number of other capacities and an analytical explanation of the capacities will lead to the function: If those capacities can be explained in terms of some general laws and together amount to the analyzed capacity, then each individual capacity can be interpreted as a function.

> Component c functions as a f in the system S (or the function of c in S is to do f) relative to an analytical account A of the S's capacity to ψ just in case c is capable

> of doing f in S and A appropriately adequately accounts for S's capacity to ψ by, in part, appealing to the capacity of c to do f in S.

Cummins's (1974, 1975) theory is the foundation of the model-based approach to FR, common in AI. Subsumption is the basis of structuring taxonomy; it can be viewed as a class hierarchy connecting concepts using "is-a" relations. The idea of subsumption and analytical explanation of capacities (dispositions) is used implicitly in many FR-based systems in planning, resource allocation, and explanation-based systems that use multilayer concepts for behavior and function.

6. DISCUSSION

6.1. Underlying assumptions

Having some ontological primitives (tokens or concepts) and representation and inferential schemes, any physical phenomena can, in principle, be explained in terms of *histories* and *episodes* (Hayes, 1985, 1990). Episodes are proper temporal slices of a history (Hayes, 1979). What is called *state* is an episode of zero duration. Using the notion of history and state, one can come up with the concept function.

6.1.1. Functionality as a property of a pair (FIP)

There is a question concerning whether a function resides in an artifact (or its components) or it is an outcome of the interaction between artifacts (or two or more components). At first glance, it seems that humans have a database in which artifacts are associated with several functionalities. Many of the FR systems have assumed that a function is a property of its source artifact. Perhaps this is one of the sources of difficulty in both logical formulation (see, i.e., Wright, 1974) and actual implementation (see typical works of Tezza & Trucco, 1988; Keuneke, 1991; etc., for systems based on this assumption). Some other works argue that a function can be ascribed to a pair of artifacts instead of a single one (see, i.e., Forbus et al., 1987; Joskowicz, 1987; Faltings, 1990). In terms of histories of individual artifacts and states, it is almost impossible to explain how different functions can be attached to a single artifact. The "functionality in pair (FIP)" (Far, 1999) seems to be a central assumption stating that the at least a pair of artifacts (or components) are required to interact functionally and a function concept can be derived from their combined histories. Similar ideas are mentioned by the *locality of histories* (Hayes, 1985), *connectivity hypothesis* (Forbus et al., 1987) and *pairwise interaction of parts* (Faltings, 1990).

6.1.2. Functionality in state transition (FST)

Intuitively, the history that leads to a function should display a certain pattern (Bigelow & Pargetter, 1987). States, in the sense defined above, are useful to extract those patterns and define functional concepts. It is argued that a functional concept can be derived by identifying a persisted state or discovering an order in the sequence of states (Far, 1999). In other words, to say that a component c in a system S has a function f, f should be exhibited regularly when c is present in several configurations of the system S and in multiple setting of the environment in which system S performs. In biological systems, persistence is perhaps the most obvious regularity characteristic, and is believed to be governed by natural selection law. In designed artifacts other kinds of regularity may be discovered.

Joining the FST with the FIP leads to the function formation technique (Far, 1999). Interaction between components in a system is represented by their *inputs* and *outputs*. Inputs and outputs are described by a shared set of state variables for the components. In this sense, a component can be viewed as an n-bit processor, whose contribution to the functioning of the whole system depends on first, the active bits on the shared bus with the other components, and second, the other components having the same bits active. Finally, the function itself, is a regular pattern of states of their shared bits.

An FR technique based on these assumptions has many advantages. First, the problem of mapping from behavior to functions is removed and the functions describe the current nature of an artifact (see Bigelow, 1987, for a discussion on the importance of this factor in the explanation). Second, a function concept derived in this way can be explained in terms of system's structure and behavior without reference to any other intermediary concepts. Third, it provides a framework for comparing and evaluating functions of completely different systems with different structures. Fourth, it is an appropriate vehicle to explain the existence of certain components in a system, in terms of their contribution to persistence or a desired regularity of behavior in the containing system's behavior.

6.2. Explaining goal-directed behavior: Causal and FE

Although goals and functions share a big portion of their meaning spectrum, the explanation of goal-directed behavior includes two distinct components: causal and FEs (Nagel, 1977*a*). There are similarities between the two: both are supposed to have a reference to the context (DeKleer, 1984) and both refer to events that *usually* or *naturally* take place (Shoham, 1990). Despite the similarities, there are some important distinctions (Nagel, 1977*a*).

> Explanations proposed in connection with goal-directed (behavior) account for the presence of various items in two different ways. One way is the explanation of HOW the goal is realized in terms of assumed capacities of the system's various organs, the organization of the system's component parts, and a number of laws concerning the effects produced by the activities of those parts. . . . Explanations of this sort are often said to be causal. . . . Expla-

> nations of this type . . . are found in all branches of inquiry, and there is nothing teleological about them.

First, FE accounts for the presence of a component in a system in terms of certain effects it has on that system of which it is a member. Second, FE explains the "purpose" of the system in terms of either structure and behavior or functions of its components (Nagel, 1977*a*).

> Unlike causal explanations, those of this second type are often said to answer the question WHY at just the place and time it occupies . . . by stating certain consequences of the process or structure. Such explanations have traditionally been called teleological.

The first category of FE refers to an explanation of the presence of some component in the system (or state why the component exists) in terms of the contributions it makes to, in terms of certain effects the component produces in the system (Nagel, 1977*b*), or in terms of some capacity that the component has and contributes to the capacity of the containing system (Cummins, 1974, 1975). In the second category of FE, the traditional teleological process of *means* and *ends* are identified. End is that character of the system by virtue of which it functions or is capable of functioning. Means refer to a partial arrangement of such a whole to realize such an end.

Cummins (1974) has created two strategies for FE: subsumption and analytical. In the subsumption strategy, the elements of explanation are certain kinds of events, e, that would cause the artifact i to manifest the function f. Here, the explanation clarifies the connection between the events e, and the manifestations f, as instances of one or more general laws that are not special to the artifact i. In the analytical strategy function f is decomposed into a number of other functions, $f_1 \dots f_n$, each manifested by the artifact i or a pair of its components, in a way that the f_js result in or amount to a manifestation of f. The decomposition of f is pushed down until the explanation can be developed for all f_j using the subsumption strategy.

6.3. Is FR useful?

The eliminativism viewpoint on functions argues that interpretation of functional attributions of the artifacts are considered as doubtful, because of two reasons. The first is the lack of understanding of how various internal mechanisms operate and how various artifacts interact, and human's evaluation or hypotheses about relevant or irrelevant causes is a necessity. "Therefore ascription of function to artifacts cannot be taken literally, as objective assertions. They must be construed as statements that have only a heuristic value in guiding inquiry into the mechanisms." (Nagel, 1977*b*). In other words, functional (and causal) explanations are supposed to be assumption based and may not be considered as descriptions of genuine and lawful property of artifacts. That is, changes of the assumption set will affect the functionality of an artifact (Shoham, 1990). For example, a buzzer cannot exhibit its supposed function when the assumption that "the clapper can be lifted by the magnetic field, against the spring's restoring force" is violated, or even if the assumption of being located in air with atmospheric pressure is removed.

The second reason is that functions (and causes) may not be considered *scientific* because they cannot be defined by lists of objective attributes (Russell, 1913), and they do not play any significant role in explaining the nature of the items (Bigelow & Pargetter, 1987).

Ignoring FIP is a source of misinterpretation of the first kind. The FIP assumption states that the context (or environment or constraints, etc.) is expressible in terms of a pair of items or components. For the clapper and coil pair in the buzzer example, either alternative lifting and releasing function, or remaining in the released state is deducible.

Rasmussen (1991) has argued against the second reason.

> The quantitative, relational representation of physical objects, (based on a selection of practically isolated relationships) considers the objects to be a well defined microworld in which the relationships of the physical laws are undisturbed by external factors . . . and their behavior can be described with no reference to internal physical functioning . . . This is not applicable for analysis of the courses of events when the structure of the technical systems break down.

Causal or FR is more common in cases that the objects are studied along with the interaction with their environments, such as analysis of accidents, using prototypical categories of causes and functions. They are useful in the way they contribute to widening understanding and predicting hypothetical courses of events, even if they might not be considered scientific in restrict terms. In other words, an explanation at a functional level is just a description of the phenomena, and details of the physical laws and causal chain underlying these phenomena gives a real, sufficient, and complete explanation (Rosenberg, 2001).

7. CONCLUSION

FR research has been carried out for a long time in various fields including biology, sociology, and engineering design. Numerous FR-based systems have been built using a variety of theories, assumptions, and ontological primitives. There is no single FR theory or technique that can fulfilled the purpose function of the researchers and developers in these fields. In this paper a comparative survey of the FR theories was presented. The FR research in a variety of disciplines was reviewed and the common core and basic problems were identified. Instead of using eliminative approach we intentionally presented a spectrum of definitions, theories, and methods with the emphasis on their differences. A main conclusion is that the developers of FR-based systems in engineering design should clarify their basic assumptions and the theory they are built upon. This will help the users to have a better understanding of the methods, completeness, and/or soundness of the theory and validity of the results.

ACKNOWLEDGMENTS

The authors are grateful to the editors and anonymous referees for their valuable and constructive comments on earlier versions of this manuscript.

REFERENCES

Abu-Hanna, A., Benjamins, R., & Jansweijer, W. (1991). Device understanding and modeling for diagnosis. *IEEE Expert 6(2)*, 26–32.

Acar, L., & Ozguner, U. (1990). Design of knowledge-rich hierarchical controllers for large functional systems. *IEEE Transactions on Systems, Man, and Cybernetics 20(4)*, 791–803.

AIEDAM. (1996). Special Issue. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 10(4)*.

AI Magazine. (1990). Special Issue. *AI Magazine 11(4)*.

Allan, D.M. (1952). Towards a natural teleology. *The Journal of Philosophy 49*, 449–459.

Asai, K., Uenaka, J., Kambayashi, S., Higuchi, K., Kume, E., Otani, T., Fujii, M., & Fujisaki, M. (1990). HASP: human acts simulation program. *Proc. 1st Int. Conf. Supercomputing in Nuclear Applications* (SNA '90), Mito, Japan, March 1990.

Ayala, F.J. (1970). Teleological explanation in evolutionary biology. *Philosophy of Science 37(1)*, 1–15.

Beckner, M. (1969). Function and teleology. *Journal of the History of Biology 2*, 151–164.

Bigelow, J., & Pargetter, R. (1987). Functions. *The Journal of Philosophy 84(4)*, 181–197.

Bobrow, D.G. (1984). Qualitative reasoning about physical systems: an introduction. *Artificial Intelligence 24*, 1–5.

Brady, M., & Connell, J. (1987). Generating and generalizing models of visual objects. *Artificial Intelligence 31*, 159–183.

Bradshaw, J.A., & Young, R.M. (1991). Evaluating design using knowledge of purpose and knowledge of structure. *IEEE Expert 6(2)*, 33–40.

Bylander, T. (1988). A critique of qualitative simulation from a consolidation viewpoint. *IEEE Transactions on Systems, Man, and Cybernetics 18(2)*, 252–263.

Canfield, J. (1964). Teleological explanations in biology. *The British Journal of Philosophy of Science 14*, 285–295.

Chandrasekaran, B. (1990). Design problem solving: a task analysis. *AI Magazine 11(4)*, 59–71.

Chandrasekaran, B., & Josephson, J.R. (2000). Function in device representation. *Engineering with Computers 16*, 162–177.

Chakrabarti, A. (1998). Supporting two views of function in mechanical designs. *AAAI Workshops on Reasoning about Function and on Functional Modeling and Teleological Reasoning*, Madison, WI.

Chakrabarti, A., & Bligh, T.P. (2001). A scheme for functional reasoning in conceptual design. *Design Studies 22(6)*, 493–517.

Chittaro L., & Kumar, A.N. (1998). Reasoning about function and its applications to engineering. *Artificial Intelligence in Engineering 12(4)*, 331–336.

Cockburn, A. (2001). *Writing Effective Use Cases*. Reading, MA: Addison–Wesley.

Constantine, L.L., & Lockwood, L.A.D. (1999). *Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design*. Reading, MA: Addison–Wesley.

Cummins, R. (1974). Dispositions, states and causes. *Analysis 34.6(162)*, 194–204.

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy 72(20)*, 741–765.

DeKleer, J. (1984). How circuits work. *Artificial Intelligence 24*, 205–280.

DeKleer, J., & Brown, J.S. (1984). A qualitative physics based on confluences. *Artificial Intelligence 24*, 7–83.

DeJong, H.L. (2003). Causal and functional explanations. *Theory and Psychology 13(33)*, 291–317.

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

Dormoy, J.L., & Raiman, O. (1988). Assembling a device. *Proc. 7th National Conf. Artificial Intelligence (AAAI-88)*, pp. 330–335.

Ellman, T. (1989). Explanation-based learning: a survey of programs and perspectives. *ACM Computing Surveys 21(2)*.

Elster, J. (1983). Functional explanation in social science. In *Explaining Technical Change*, pp. 55–68. New York: Cambridge University Press.

Faltings, B. (1987). Qualitative kinematics in mechanisms. *Proc. 10th Int. Joint Conf. Artificial Intelligence (IJCAI '87)*, pp. 436–442.

Faltings, B. (1990). Qualitative kinematics in mechanisms. *Artificial Intelligence 44*, 89–119.

Far, B.H. (1992). *Functional Reasoning, Explanation and Analysis*. Technical Report JAERI-M 91-225.

Far, B.H. (1999). How to learn function of devices. *IEEE Int. Conf. Systems, Man, and Cybernetics, IEEE SMC '99*, Vol. 5, pp. 860–865.

Far, B.H., & Nakamichi, M. (1993). Qualitative fault diagnosis in systems with nonintermittent concurrent faults: a subjective approach. *IEEE Transactions on Systems, Man, and Cybernetics 23(1)*, 14–30.

Fink, P.K. (1985). Control and integration of diverse knowledge in a diagnostic expert system. *Proc. 9th Int. Joint Conf. Artificial Intelligence (IJCAI '85)*, pp. 426–431.

Fink, P.K., & Lusth, J.C. (1987). Expert systems and diagnosis expertise in the mechanical and electrical domains. *IEEE Transactions on Systems, Man, and Cybernetics 17(3)*, 340–349.

Forbus, K., Nielsen, P., & Faltings, B. (1987). Qualitative kinematics: a framework. *Proc. 10th Int. Joint Conf. Artificial Intelligence (IJCAI '87)*, pp. 430–435.

Franke, D.W. (1991). Deriving and using descriptions of purpose. *IEEE Expert 6(2)*, 41–47.

Freeman, P., & Newell, A. (1971). A model for functional reasoning in design. *Proc. 2nd Int. Joint Conf. Artificial Intelligence (IJCAI '71)*, pp. 621–640.

Gautier, P.O., & Gruber, T.R. (1993). Generating explanations of device behavior using compositional modeling and causal ordering. *Eleventh National Conf. Artificial Intelligence*.

Gelsey, A. (1987). Automated reasoning about machine geometry and kinematics. *Proc. 3rd IEEE Conf. AI Applications*, Orlando, FL.

Hayes, P. (1979). Naive physics 1: ontology for liquids. In *Formal Theories of the Common Sense World* (Hobbs, J.R., & Moore, R.C., Eds.), pp. 71–107. Norwood, NJ: Ablex Publishing Co.

Hayes, P. (1985). The second naive physics manifesto. In *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence* (Collins, A., & Smith, E.E., Eds.), pp. 468–485. San Mateo, CA: Morgan Kaufman.

Hayes, P. (1990). The naive physics manifesto. In *The Philosophy of Artificial Intelligence* (Boden, M.A., Ed.), Chap. 8. New York: Oxford University Press.

Hempel, C.G. (1959). The logic of functional analysis. In *Readings in the Philosophy of Social Science* (Martin, M., & McIntyre, L.C., Eds.). Cambridge, MA: MIT Press.

Hubka, V., & Eder, W.E. (2001). Functions revisited. *Proc. ICED '01, Glasgow.*

IEEE Intelligent Systems. (1997). Special Issue. *IEEE Intelligent Systems 12(2)*.

International Journal of Applied AI. (1994). Special Issue. *International Journal of Applied AI 8(2)*.

Iwasaki, Y. (1989). *Causality and Model Abstraction*. Knowledge Systems Laboratory Report TR KSL-89-80.

Jacobson, I., Christerson, M., Jonsson, P., & Övergaard. G. (1992). *Object-Oriented Software Engineering: A Use Case Driven Approach*. Harlow, England: Addison–Wesley.

Joskowicz, L. (1987). Shape and function in mechanical devices. *Proc. 6th National Conf. Artificial Intelligence (AAAI-87)*, pp. 611–615.

Joskowicz, L., & Addanki, S. (1988). From kinematics to shape: an approach to innovative design. *Proc. 7th National Conf. Artificial Intelligence (AAAI-88)*, pp. 347–352.

Kara, L.B., & Stahovich, T.F. (2002). Spatial reasoning about mechanical behaviors. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 16(5)*, 363–384.

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

Kuipers, B. (1986). Qualitative simulation. *Artificial Intelligence 29*, 289–338.

Lehman, H. (1965). Functional explanations in biology. *Philosophy of Science 32(1)*, 1–20.

Lind, M. (1988). System concepts and the design of man–machine interfaces for supervisory control. In *Tasks, Errors and Mental Models* (Goodstein, L.P., Andersen, H.B., & Olsen, S.E., Eds.), pp. 269–277. London: Taylor & Francis.

Matthen, M. (1988). Biological functions and perceptual content. *The Journal of Philosophy 85*, 5–27.

McMenamin, S.M., & Palmer J.F. (1984). *Essential Systems Analysis*. Englewood Cliffs, NJ: Prentice Hall.

Merriam–Webster. (2002). *Webster's Third New International Dictionary, Unabridged*. Springfield, MA: Merriam–Webster.

Millikan, R.G. (1989). In defense of proper functions. *Philosophy of Science 56*, 288–302.

Murakami, T., & Nakajima, N. (1988). Computer-aided design—diagnosis using feature description. In *Artificial Intelligence in Engineering: Diagnosis and Learning* (Gero, J.S., Ed.), pp. 199–226. New York: Elsevier.

Nagel, E. (1977*a*). Teleology revisited: goal directed processes in biology. *The Journal of Philosophy 74(5)*, 261–279.

Nagel, E. (1977*b*). Teleology revisited: functional explanation in biology. *The Journal of Philosophy 74(5)*, 280–301.

Nielsen, J. (2000). *Designing Web Usability*. Berkeley, CA: New Riders Publishing.

Oxford University Press. (2003). *Oxford English Dictionary*. New York: Oxford University Press.

Pegah, M., Sticklen, J., & Bond, W. (1993). Functional representation and reasoning about the F/A-18 aircraft fuel system. *IEEE Expert 2*, 65–71.

Pu, P., & Badler, N.I. (1988). Design knowledge capturing for device behavior reasoning. In *Artificial Intelligence in Engineering Design* (Gero, J.S., Ed.), pp. 37–56. New York: Elsevier.

Qian L., & Gero, J.S. (1996). Function–behavior–structure paths and their role in analogy-based design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 10(4)*, 289–312.

Rasmussen, J. (1983). Skills, rules and knowledge: signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics 13(3)*, 257–266.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man, and Cybernetics 15(2)*, 234–243.

Rasmussen, J. (1990). Human error and the problem of causality in analysis of accidents. *Philosophical Transactions of the Royal Society of London B327*, 449–462.

Rasmussen, J. (1991). Event analysis and the problem of causality. In *Distributed Decision Making: Cognitive Models for Cooperative Work* (Rasmussen, J., Brehmer, B., & Leplat, J., Eds.). New York: Wiley.

Rosenberg, A. (2001). Reductionism in a historical science. *Philosophy of Science 68*, 135–163.

Russell, B. (1913). On the notion of cause. *Proceedings of the Aristotelean Society 13*, 1–25.

Ruse, M. (1971). Function statements in biology. *Philosophy of Science 38(1)*, 87–95.

Russumanno, D., & Bonnell, R.D. (1996). Functional reasoning in a failure modes and effect analysis (FEMA) expert system. *Proc. IEEE Reliability and Maintainability Symp.*, pp. 339–347.

Sembugamoorthy, V., & Chandrasekaran, B. (1986). Functional representation of devices and compilation of diagnostic problem solving systems. In *Experience, Memory and Reasoning* (Kolodner, J., & Reisbeck, C., Eds.), pp. 47–73. Mahwah, NJ: Erlbaum.

Shank, R., & Abelson, R. (1977). *Scripts, Plans, Goals and Understanding: An Inquiry Into Human Knowledge Structure*. Mahwah, NJ: Erlbaum.

Shekar, B. (1990). Temporal sequencing of concepts in a functional knowledge base. *Proc. Pacific Rim Int. Conf. Artificial Intelligence (PRICAI '90)*, pp. 699–704.

Shoham, Y. (1990). Non-monotonic reasoning and causation. *Cognitive Science 14*, 213–252.

Simon, H.A. (1969). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Sorabji, R. (1964). Function. *Philosophy Quarterly 14(57)*, 289–302.

Stanfill, C. (1983). The decomposition of a large domain: reasoning about machines. *Proc. National Conf. Artificial Intelligence (AAAI-83)*, pp. 387–390.

Sticklen, J., & Bond, W.E. (1991). Functional reasoning and functional modeling. *IEEE Expert 6(2)*, pp. 20–21.

Tezza, T., & Trucco, E. (1988). Functional reasoning for flexible robots. In *AI in Engineering: Robotics and Processes* (Gero, J.S., Ed.), pp. 3–19. New York: Elsevier.

Ulrich, K.T., & Seering, W.P. (1988). Function sharing in mechanical design. *Proc. 7th National Conf. Artificial Intelligence (AAAI-88)*, pp. 342–346.

Umeda, Y., & Tomiyama, T. (1996). Supporting conceptual design based on the function–behavior–state modeler. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 10(4)*, 275–288.

Umeda, Y., & Tomiyama, T. (1997). Functional reasoning in design. *IEEE Intelligent Systems 12(2)*, 42–48.

Walsh, D.M. (1996). A taxonomy of functions. *Canadian Journal of Philosophy 26(4)*, 493–514.

Wiegers, K.E. (2003). *Software Requirements*, 2nd ed. Redmond, WA: Microsoft Press.

Wright, L. (1973). Functions. *Philosophy Review 82(2)*, 139–168.

Wright, L. (1976). *Teleological Explanations*. Berkeley, CA: University of California Press.

Behrouz Homayoun Far received his BS and MS degrees in electronic engineering from Tehran University, Iran, in 1983 and 1986, respectively. He received his PhD degree from Chiba University, Chiba, Japan, in 1990. Dr. Far is currently an Associate Professor in the Department of Electrical and Computer Engineering and the Director of the Intelligent Software Systems Laboratory, both at the University of Calgary. His research fields of interest are qualitative reasoning, automatic programming, and distributed AI. Dr. Far is a member of ACM, IEEE Computer Society, JSAI, IEICE, and IPSJ.

A. Halim Elamy received a BS in applied computer science and statistics (double major) and a certificate in remote sensing and geographic information systems from the University of Winnipeg in 2004. He also holds a BS (Honors) in engineering. Mr. Elamy has been working in software development and consultancy for more than 15 years. Halim received many international and national awards for his outstanding research and academic achievements. He is currently a graduate student in the Department of Electrical and Computer Engineering at the University of Calgary. Halim's research interests include AI, multiagent systems, experimental design, and data warehousing.