BenchCouncil Transactions on Benchmarks, Standards and Evaluations 4 (2024) 100184

Contents lists available at ScienceDirect

BenchCouncil Transactions on Benchmarks, Standards and Evaluations

journal homepage: www.keaipublishing.com/en/journals/benchcouncil-transactions-on-benchmarks-standards-and-evaluations/

Editorial

Five Axioms of Things

Jianfeng Zhan *

The International Open Benchmark Council, DE, USA
ICT, Chinese Academy of Sciences, Beijing, China
University of Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:
Thing
Model
Truth
Observation
Experiment
Evaluation
Measurement
Testing

ABSTRACT

This article explicitly defines several concepts, such as variables, models, and truth of a thing, that are fundamental to natural and social sciences. I present a generalized methodology for understaning a thing, categorically defining six foundational understanding approaches based on the nature of the thing and diverse perspectives: conjecture, observation, experiment, evaluation, measurement, and testing. I extend my previous work on the five axioms of evaluation to understanding a thing, which I call the five axioms of things. Also, I comment on five paradigms of science.

1. The formal definitions of fundamental concepts in understanding a thing

In this section, I rigorously delineate several fundamental concepts in understanding a thing, drawing upon definitions from key Refs. [1–5].

An individual can be defined as the object described by a given set of properties [5]. A *system* is a coherent entity comprising interacting or interdependent individuals, regardless of their likeness or diversity, culminating in a unified whole [1,2,5]. A system could be recursive. A *thing* could be an individual or a system. It could be a life, a natural phenomenon, an artifact, an abstract, or even a policy in natural or social sciences.

A *quantity* or *variable* is "a property of a thing whose instances can be compared by ratio or only by order [5,6]". One or more variables provide a partial understanding of a thing. The *truth* is a thing's facts or inherent properties that can be proven true or verified objectively.

A *model* serves as a streamlined representation of a thing that would otherwise be too intricate to analyze in exhaustive detail [5,7]. A model provides a full understanding of a thing, though it is simplified. A model can manifest as a physical, mathematical, or other construct. A mathematical model embodies a mathematical representation, frequently expressed through functions or equations, that captures the essence of a thing. In general, a model appears in a mathematical form. Regrettably, many things cannot be well-defined. When asserting that something is not well-defined, it implies that its structure and functions remain incompletely comprehended. For example, a human body is not well-defined.

Take the function as an instance. According to [3,5], "a function, denoted as f, is a rule that assigns a unique element, referred to as $f(x)$, from a set R to each element in a set D". In this context, "the domain, denoted as D, refers to the set of all possible values for which the function is defined [3]". On the other hand, "the range of the function, denoted as $f(x)$, consists of all the possible values that $f(x)$ can take as x varies within the domain [3]". The *independent variable* is represented by "a symbol that encompasses any arbitrary number within the domain of the function [3]". A *dependent variable*, represented by a symbol, "is used to denote a number within the range of the function [3]".

As a special instance of a model, a causal model is a *causal* explanation grounded in a model to understand a thing and infer its behavior [3,8].

A model is used to predict the outcome of understanding a thing. As these predicted outcomes increasingly align with the truth, the model is regarded as more precise, approaching a state of perfection.

2. Fundamental methodologies for understanding a thing

To obtain a model of a thing, it is essential to identify and isolate a system conducive to understanding the thing. This system must meet two criteria: first, it can operate autonomously. Second, it incorporates

* Correspondence to: The International Open Benchmark Council, DE, USA.
E-mail address: jianfengzhan.benchcouncil@gmail.com.
URL: http://www.zhanjianfeng.org.

https://doi.org/10.1016/j.tbench.2024.100184

Available online 16 December 2024

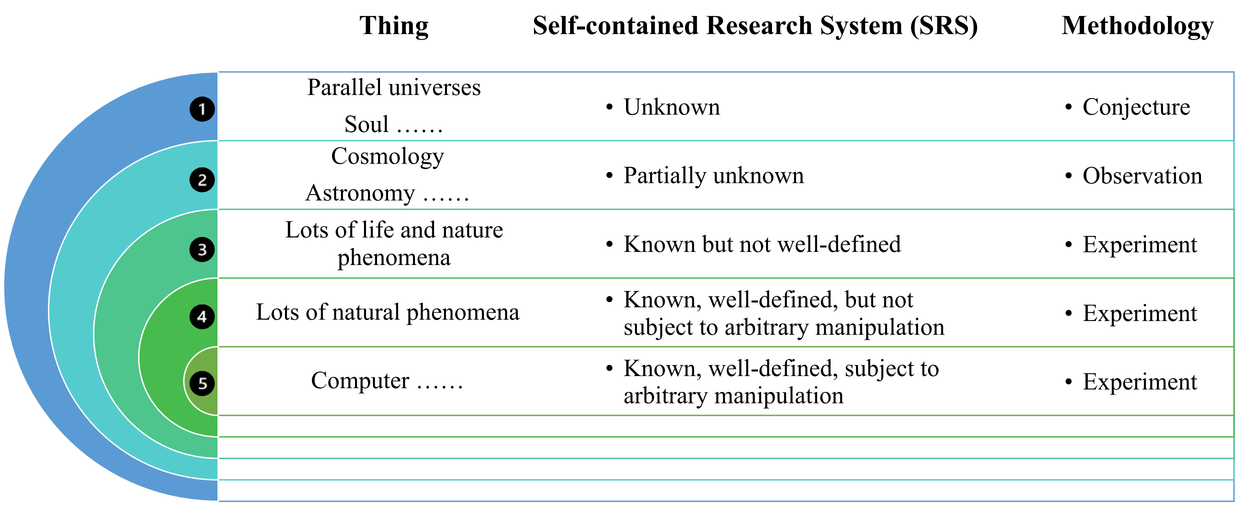

Fig. 1. Five Categories of Self-Contained Research Systems (SRS) and Their Corresponding Methodologies.

the primary factors that determine the outcomes of understanding a thing, which I refer to as essential factors. I denote this system as a Self-contained Research System (abbreviated as SRS). If it is impossible to isolate an SRS, the impacts of other external factors will affect the outcomes of understanding a thing. After identifing and isolating an SRS, it is plausiable to investigate the effect of the essencial factors on the thing. I refer to the proposed methodology as SRS.

In comparison with the causal model methodology proposed in [8], the SRS methodology offers several advantages. Firstly, even if a thing's SRS is known, it may not be well defined, making it challenging to derive its causal models. Secondly, comprehending an SRS serves as a fundamental and robust basis for establishing its causal model.

Dealing with the diverse nature of an SRS presents various challenges as shown in Fig. 1. The first kind is when an SRS is unknown, e.g., in the case of investigating parallel universes, or soul. The second kind is when an SRS is only partially known, e.g., in the case of investigating a thing in cosmology and astronomy. The third kind is when an SRS is known but cannot be well defined, e. g., a human body. The fourth kind is when an SRS is known and well-defined but not subject to arbitrary manipulation. If a system can be modeled in a function, arbitrary manipulation entails setting its independent variables to any arbitrary number within the function's domain. The fifth kind is when an SRS is known, well-defined, and subject to arbitrary manipulation. For example, a computer nearly falls into this category. Moving from the fifth kind to the first kind of SRS, the challenge level increases.

I formally define six methodologies for understanding a thing: conjecture, observation, experiment, evaluation, measurement, and testing.

The conjecture is obtaining the model of a thing when it is impossible to identify and isolate an SRS. The observation aims to derive a model of a thing in cases where there is only a partial understanding of an SRS. The experiment aims to derive a model of a thing in cases where it is feasible to identify and isolate an SRS. In the context of an experiment, there are subtly different scenarios: an SRS may not be well-defined; an SRS is well-defined but not subject to arbitrary manipulation; and an SRS is well-defined and subject to arbitrary manipulation.

Experiments and observations fall into two primary categories: those conducted independently and those involving stakeholders. Involving stakeholders classifies an experiment or observation as an evaluation, while those without stakeholders are categorized as natural experiments or observations.

My previous work [9] distinguished between the concepts of evaluation, measurement, and testing. Measurement is experimentally obtaining one or more values attributed to a quantity of a thing [6]. A test oracle is a fact or inherent property of a thing and its SRS. Testing is a verification process to determine whether (1) a thing conforms to the test oracles (the first category) and/or (2) When a thing operates within an SRS, both the thing and its SRS conform to the test oracles (the second category) [5,9]". As per the findings of [9], evaluation entails causal infering the impact and value of a thing within an SRS tailored to meet the evaluation requirements of stakeholders, relying on measurements and/or testing of the SRS.

Measurements and testing offer a foundational methodology for gaining partial insights into a thing by focusing on specific properties or facts. In contrast, observation, experiments, and evaluation endeavor to achieve a full understanding of the thing.

Viewed from another perspective, understanding a thing can be seen as intentionally imposing a research condition (RC) upon it in order to establish an SRS. Building on the previous discussion, an RC can be envisioned as the SRS from which the thing under investigation is removed. I formally define an RC as the context that is applied to the thing, playing a crucial role in guaranteeing independent operation and incorporating the essential factors. Within a particular methodology for understanding a thing, I designate a specific RC, such as an evaluation condition for evaluation, an observation condition for observation, or an experimental condition for experiment.

3. Five axioms of things

Derived from the essence of understanding a thing, I propose five axioms focusing on key aspects of the outcomes of understanding a thing, including observation, experiment, evaluation, measurement, and testing, as the foundational theory. These axioms serve as the bedrock upon which universal theories and methodologies are built for understanding a thing.

The Axiom of Metric Essence asserts that in the absence of stakeholder involvement, the essence of a metric holds intrinsic physical significance. Alternatively, when stakeholders are engaged, the essence of the metric may possess intrinsic physical significance or be solely determined by the value function. In the latter scenario, a value function establishes a composite metric that normalizes metrics of different dimensions based on stakeholder perspectives.

The Axiom of True Outcomes declares that when an SRS is known, the outcome of understanding a thing has true value.

The Axiom of Traceable Outcomes declares that when its SRSes are known, the divergence in the outcomes of understanding a thing can be attributed to disparities in RCs, thereby establishing traceability.

The Axiom of Consistent Outcomes posits that when a thing is examined under various samples from a known RC population, the outcomes tend to converge towards the true value under the RC population.

The Axiom of Comparable Outcomes declares when equipped with equivalent well-defined experimental conditions, the outcomes of understanding different things are comparable.

4. Comments to five paradigms of science

There have been several insightful discussions on the diverse paradigms of science in previous works [10,11].

Ioannidis [10] elegantly encapsulated the evolving paradigm shifts in science, delineating the transition from the traditional paradigms of empirical/experimental science (practiced for millennia) to theoretical model science (spanning centuries), followed by computational science (over decades), and the emergence of data-driven science as envisioned by Jim Gray (over the past 15 years), ultimately culminating in the advent of the 5th paradigm: AI-driven science.

Building upon the precise definitions of fundamental concepts and methodologies outlined in Sections 1 and 2, I offer concise reflections on Ioannidis's narrative regarding the paradigm shifts within the domain of science.

Following the formal definition of an ''experiment'', I express reservations about the conflation of ''empirical and experimental''. I propose the usage of ''empirical practice'' instead of ''empirical science''. Empirical practice, akin to observation and conjecture, operates within a context where the SRS is either unknown or only partially known. In such scenarios, essential factors may be absent, impeding the attainment of truth.

It is crucial to underscore that in an experiment, an SRS is known; irrespective of its manifestation, understanding its behavior within a controlled environment enables the pursuit of truth through experimentation.

As elucidated in Section 1, a model serves as the culmination of observations or experiments. On the other side, computational, data-driven, and AI-driven sciences predominantly function as novel tools or methodologies that complement observations or experiments. From this perspective, it is arguable that they do not constitute an independent methodology for understanding a thing.

References

[1] System, 2024, https://www.merriam-webster.com/dictionary/system. (Accessed: 6 February 2024).

[2] A. Backlund, The definition of system, Kybernetes 29 (4) (2000) 444–451.

[3] J. Stewart, Single Variable Calculus: Concepts and Contexts, Cengage Learning, 2018.

[4] D.S. Starnes, D. Yates, D.S. Moore, The Practice of Statistics, Macmillan, 2010.

[5] J. Zhan, L. Wang, W. Gao, H. Li, C. Wang, Y. Huang, Y. Li, Z. Yang, G. Kang, C. Luo, H. Ye, S. Dai, Z. Zhang, Evaluatology: The science and engineering of evaluation, BenchCouncil Trans. Benchmarks Stand. Eval. 4 (1) (2024) 100162.

[6] I. BiPM, I. IFCC, I. IUPAC, O. ISO, The international vocabulary of metrology—basic and general concepts and associated terms (VIM), JCGM 200 (2012) 2012.

[7] H.D. Young, R.A. Freedman, L.A. Ford, University Physics with Modern Physics, 2020.

[8] J. Pearl, D. Mackenzie, The book of why: the new science of cause and effect, Basic books, 2018.

[9] J. Zhan, A short summary of evaluatology: The science and engineering of evaluation, BenchCouncil Trans. Benchmarks Stand. Eval. (2024) 100175.

[10] Y. Ioannidis, The 5th paradigm: AI-driven scientific discovery, Commun. ACM (2024).

[11] T. Hey, S. Tansley, K.M. Tolle, et al., The Fourth Paradigm: Data-Intensive Scientific Discovery, vol. 1, Microsoft research Redmond, WA, 2009.

Dr. Jianfeng Zhan is a Full Professor at the Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), and University of Chinese Academy of Sciences (UCAS), the director of the Research Center for Distributed Systems, ICT, CAS. He received his B.E. in Civil Engineering and MSc in Solid Mechanics from Southwest Jiaotong University in 1996 and 1999 and his Ph.D. in Computer Science from the Institute of Software, CAS, and UCAS in 2002. His research areas focus on evaluatology, evaluatology-based design automation, and optimization automation. His exceptional expertise is exemplified by his introduction to the discipline of evaluatology, an endeavor that encompasses the science and engineering of evaluation; within this discipline, his proposition of a universal framework for evaluation encompasses essential concepts, terminologies, theories, and methodologies for application across various disciplines. He has made substantial and effective efforts to transfer his academic research into advanced technology to impact general-purpose production systems. Several technical innovations and research results, including 35 patents from his team, have been adopted in benchmarks, operating systems, and cluster and cloud system software with direct contributions to advancing parallel and distributed systems in China or worldwide. Over the past two decades, he has supervised over ninety graduate students, post-doctors, and engineers. Dr. Jianfeng Zhan is the founder and chairman of BenchCouncil. He also holds the role of Co-EIC of BenchCouncil Transactions on Benchmark, Standards and Evaluations, alongside Prof. Tony Hey. Dr. Zhan has served as an Associate Editor for IEEE TPDS (IEEE Transactions on Parallel and Distributed Systems) from 2018 to 2022. In recognition of his exceptional contributions, he has been honored with several prestigious awards. These include the second-class Chinese National Technology Promotion Prize in 2006, the Distinguished Achievement Award of the Chinese Academy of Sciences in 2005, the IISWC Best Paper Award in 2013, and the Test of Time Paper Award from the Journal of Frontier of Computer Science.