Knowledge-driven discovery of operational definitions

Jan M. Żytkow

Computer Science Department, UNC Charlotte, Charlotte, N.C. 28223
and Institute of Computer Science, Polish Academy of Sciences
zytkow@uncc.edu

Abstract. Knowledge representation which is internal to a computer lacks empirical meaning so that it is insufficient for the investigation of the external world. All intelligent systems, including robot-discoverers must interact with the physical world in complex, yet purposeful and accurate ways. We argue that operational definitions are necessary to provide empirical meaning of concepts, but they have been largely ignored by the research on automation of discovery and in AI. Individual operational definitions can be viewed as algorithms that operate in the real world. We explain why many operational definitions are needed for each concept and how different operational definitions of the same concept can be empirically and theoretically equivalent. We argue that all operational definitions of the same concept must form a coherent set and we define the meaning of coherence. No set of operational definitions is complete so that expanding the operational definitions is one of the key tasks in science. Among many possible expansions, only a very special few lead to a satisfactory growth of scientific knowledge. While our examples come from natural sciences, where the use of operational definitions is especially clear, operational definitions are needed for all empirical concepts. We briefly argue their role in database applications.

1 Operational definitions provide empirical meaning

Data about external world are obtained by observation and experiment. Sophisticated procedures and instruments are commonly used to reach data of scientific value. Yet we rarely think systematically about methods by which data have been procured, until problems occur. When a set of data is inconsistent with our expectations, we start asking: "How was this particular measurement obtained?", "What method has been used?", "How is this method justified?". Often it turns out that a method must be changed. Because data can be wrong in so many ways, sophisticated knowledge is required in order to examine and improve measurement methods.

It is critical to the growth of scientific knowledge to study new situations, for which no known method can measure a particular quantity. For instance, we may wish to measure temperatures lower than the capabilities of all existing instruments. Or we want to measure temperature change inside a living cell, as the cell undergoes a specific process.

When no known method applies, new methods must be discovered. New measurement methods must expand the existing concepts. For instance, a new thermometer must produce measurements on a publicly shared scale of temperature.

Discovery of new measurement methods, which we also call operational definitions, is the central problem in this paper. We provide an algorithm that demonstrates how empirical knowledge is used to construct new operational definitions, how new methods can be empirically verified and how choices can be made among competing methods.

We end each section with a few basic claims about operational definitions.

Claim 1: For each empirical concept, measurements must be obtained by repeatable methods that can be explained in detail and used in different laboratories.

Claim 2: The actual verification in empirical science is limited to empirical facts. Operational definitions determine facts; thus they determine the scope of scientific verification.

Claim 3: In contrast, scientific theories often make claims beyond the facts that can be empirically verified at a given time. Theoretical claims often apply to all physical situations, whether we can observe them or not.

In this paper we use examples of numerical properties of objects and their pairs. The numbers that result from measurements, for instance temperature or distance, we call *values* of empirical concepts.

Claim 4: Operational definitions can be classified in several dimensions: (a) they apply to objects, states, events, locations and other empirical entities; (b) they may define predicates of different arity, for instance, properties of individual objects, object pairs (distance) or triples (chemical affinity); (c) some operational definitions provide data while others prepare states that possess specific properties, such as the triple point of water.

2 The AI research has neglected operational definitions

Operational semantics links the terms used in scientific theories with direct observations and manipulations (Bridgman, 1927; Carnap, 1936). While important in empirical science, the mechanisms that produce high quality experiments have been neglected not only in the existing discovery systems but in the entire domain of artificial intelligence.

The distinction between formalism and its interpretation, also called semantics, has been applied to the study of science since 1920's and 1930's. Scientific theories have been analyzed as formal systems whose language is empirically interpreted by operational definitions.

A similar distinction applies to discovery systems and to knowledge they create. A discovery mechanism such as BACON (Langley, Simon, Bradshaw & Zytkow, 1987) can be treated as (1) a formal system that builds equations from data that are formally tuples in the space of the values of independent and dependent variables plus (2) a mechanism that procures data.

Similarly to scientists, BACON and other discovery systems use plans to propose experiments. Each experiment consists in selecting a list of values $x_1, ..., x_k$ of empirical variables $X_1, ..., X_k$, and in obtaining the value y of a dependent variable Y which provides the "world response" to the empirical situation characterized by $x_1, ..., x_k$. But instead of real experiments, the values of dependent variables are either typed by the user or computed in simulation, in response to the list of values of independent variables.

This treatment bypasses real experimentation and measurements. Other papers and collections that consider many components of the scientific methods (Kulkarni & Simon, 1987; Sleeman, Stacey, Edwards & Gray, 1989; Shrager & Langley, 1990; Valdes-Perez, 1995) neglect operational definitions of concepts.

In the wake of robotic discovery systems, operational semantics must, at the minimum, provide realistic methods to acquire data. Żytkow, Zhu & Hussam (1990) used a robotic mechanisms which conducted automatically experiments under the control of FAHRENHEIT. In another robotic experiment, Żytkow, Zhu & Zembowicz (1992) used a discovery process to refine an operational definition of mass transfer. Huang & Zytkow (1997) developed a robotic system that repeats Galileo's experiment with objects rolling down an inclined plane. One operational definition controlled the robot arm so that it deposited a cylinder on the top of an inclined plane, while another measured the time in which the cylinder rolled to the bottom of the plane.

While operational semantics must accompany any formalism that applies to the real world, it has been unnoticed in AI. Jackson's claim (1990) is typical: "a well-defined semantics ... reveals the meaning of ... expressions by virtue of their form." But this simply passes on the same problem to a broader formalism, that includes all the terms used in formal semantics. Those terms also require real-world interpretation that must be provided by operational definitions.

Plenty of further research must be conducted to capture the mechanisms in which operational definitions are used in science and to make them applicable on intelligent robots.

Claim 5: Formal semantics are insufficient to provide empirical meaning.

Claim 6: Robotic discoverers must be equipped in operational definitions.

3 Operational definitions interact with the real world

Early analyses of operational definitions used the language of logic. For instance, a dispositional property "soluble in water" has been defined as

If x is in water then (x is soluble in water if and only if x dissolves)

But a more adequate account is algorithmic rather than descriptive:

```
Soluble (x)
  Put x in water!
  Does x dissolve?
```

As an algorithm, operational definition consists of instructions that prescribe manipulations, measurements and computations on the results of measurements. Iteration can enforce the requirements such as temperature stability, which can be preconditions for measurements. Iteration can be also used in making measurements. The loop exit condition such as the equilibrium of the balance, or a coincidence of a mark on a measuring rod with a given object, triggers the completion of a step in the measurement process.

Procedures that interpret independent and dependent variables can be contrasted as manipulation and measurement mechanisms. Each independent variable requires a manipulation mechanism which sets it to a specific value, while a response value of an dependent variable is obtained by a measurement mechanism. In this paper we focus on measurement procedures.

It happens that an instruction within procedure P does not work in a specific situation. In those cases P cannot be used. Each procedure may fail for many reasons. Some of these reasons may be systematic. For instance, a given thermometer cannot measure temperatures below -40C because the thermometric liquid freezes or above certain temperature, when it boils. Let us name the range of physical situations to which P applies by R_P.

Often, a property is measured indirectly. Consider distance measurement by sonar or laser. The time interval is measured between the emitted and the returned signal. Then the distance is calculated as a product of time and velocity. Let $C(x)$ be the quantity measured by procedure P. When P terminates, the returned value of C is $f(m_1, ..., m_k)$, where $m_1, ..., m_k$ are the values of different quantities of x or the empirical situation around x, measured or generated by instructions within P, and f is a computable function on those values.

Claim 7: Each operational definition should be treated as an algorithm.

Claim 8: The range of each procedure P is limited in many ways, thus each is merely a partial definition applicable in the range R_P.

Claim 9: An operational definition of concept C can measure different quantities and use empirical laws to determine the value of C: $C(x) = f(m_1, ..., m_k)$

Claim 10: An operational definition of a concept $C(x)$ can be represented by a descriptive statement: "If x is in R_P then $C(x) = f(m_1, ..., m_k)$"

4 Each concept requires many operational definitions

In everyday situations distance can be measured by a yard-stick or a tape. But a triangulation method may be needed for objects divided by a river. It can be extended to distance measurement from the Earth to the Sun and the Moon. Then, after we have measured the diameter of the Earth orbit around the Sun, we can use triangulation to measure distances to many stars.

But there are stars for which the difference between the "winter angle" and the "summer angle" measured on the Earth, is non-measurably small, so another method of distance measurement is needed. Cefeids are some of the stars within the range of triangulation. They pulsate and their maximum brightness varies according to the logarithm of periodicity. Another law, determined on Earth and applied to stars claims that the perceived brightness of a constant light source diminishes with distance as $1/d^2$. This law jointly with the law for cefeids allows us to determine the distance to galaxies in which individual cefeids are visible.

For such galaxies the Hubble Law was empirically discovered. It claims proportionality between the distance and red shift in the lines of hydrogen spectrum. The Hubble Law is used to determine the distance of the galaxies so distant that cefeids cannot be distinguished.

Similarly, while a gas thermometer applies to a large range of states, in very low temperatures any gas freezes or gas pressure becomes non-measurably small. A thermometer applied in those situations measures magnetic susceptibility of paramagnetic salts and uses Curie-Weiss Law to compute temperature. There are high temperatures in which no vessel can hold a gas, or states in which the inertia of gas thermometer has unacceptable influence on the measured temperature. Measurements of thermal radiation and other methods can be used in such cases.

Claim 11: Empirical meaning of a concept is defined by a set of operational definitions.

Claim 12: Each concrete set is limited and new methods must be constructed for objects beyond those limits.

5 Methods should be linked by equivalence

Consider two operational definitions P_1 and P_2 that measure the same quantity C. When applied to the same objects their results should be empirically equivalent within the accuracy of measurement. If P_1 and P_2 provide different results, one or both must be adjusted until the empirical equivalence is regained.

From the antiquity it has been known that triangulation provides the same results, within the limits of measurement error, as a direct use of measuring rod or tape. But in addition to the empirical study of equivalence, procedures can be compared with the use of empirical theories and equality of their results may be proven.

Triangulation uses a basic theorem of Euclidean geometry that justifies theoretically the consistency of two methods: by the use of yard-stick and by triangulation. To the extent in which Euclidean geometry is valid in the physical world, whenever we make two measurements of the same distance, one using a tape while the other using triangulation, the results are consistent.

Claim 13: Methods can differ by their accuracy and by degree to which they influence the measured quantity.

Claim 14: When two operational definitions define the same property and apply to the same objects, their results should be empirically equivalent. If they are not, additional data are collected and methods are adjusted in order to restore their equivalence.

Claim 15: When two operational definitions define the same concept $C(x)$, it is possible to prove their equivalence. The prove consists in deducing from a verified empirical theory that the statements that represent them are equivalent, that is, $f_1(m_1, ..., m_k) = f_2(n_1, ..., n_l)$

Claim 16: When the statements that represent two procedures use empirical laws $C(x) = f_1(m_1, ..., m_k)$, $C(x) = f_2(n_1, ..., n_l)$, theoretical equivalence of both procedures follows from those laws.

Claim 17: The more general and better verified are the theories that justify the equivalence of two procedures P_1 and P_2, the stronger are our reasons to believe in the equivalence of P_1 and P_2.

Claim 18: Proving the equivalence of two procedures is desired, because the empirical verification of equivalence is limited.

6 Operational definitions of a concept form a coherent set

We have considered several procedures that measure distance. But distance can be measured in many other ways. Even the same method, when applied in different laboratories, varies in details. How can we determine that different measurements define the same physical concept? Procedures can be coordinated by the requirements of empirical and theoretical equivalence in the areas of common application. However, we must also require that each method overlaps with some other methods and further, that each two methods are connected by a chain of overlapping methods.

Definition: A set $\Phi = \{\phi_1, ..., \phi_n\}$ of operational definitions is coherent **iff** for each i, j = 1,...,n

(1) ϕ_i is empirically equivalent with ϕ_j. Notice that this condition is trivially satisfied when the ranges of both operational definitions do not overlap;

(2) there is a sequence of definitions ϕ-i_1,...,ϕ-i_k, such that ϕ-$i_1 = \phi_i$, ϕ-$i_k = \phi_j$, and for each $m = 2, ..., k$ the ranges of ϕ-i_m and ϕ-i_{m+1} intersect.

The measurements of distance in our examples form such a coherent set. Rod measurements overlap with measurements by triangulation. Different versions of triangulation overlap with one another. The triangulation applied to stars overlaps with the method that uses cefeids, which in turn overlaps with the method that uses Hubble Law.

Similarly, the measurements with gas thermometer have been used to calibrate the alcohol and mercury thermometers in their areas of joint application. For high temperatures, measurements based on the Planck Law of black body radiation overlap with the measurements based on gas thermometers. For very low temperatures, the measurements based on magnetic susceptibility of paramagnetic salts overlap with measurements with the use of gas thermometer.

Claim 19: Each empirical concept should be defined by a coherent set of operational definitions. When the coherence is missing, the discovery of a missing link becomes a challenge.

For instance, the experiment of Millikan provided a link between the charge of electron and electric charges measured by macroscopic methods.

Claim 20: By examining theoretical equivalence in a coherent set Φ of operational definitions we can demonstrate that the values measured by all procedure in Φ are on the same scale.

Claim 21: Operational definitions provide means to expand to new areas the range of the laws they use.

7 Laws can be used to form new operational definitions

Operational definitions can expand each concept in several obvious directions, towards smaller values, larger values, and values that are more precise. But the directions are far more numerous. Within the range of "room" temperatures, consider the temperature inside a cell, temperature of a state that is fast varying and must be measured every second, or temperature on the surface of Mars. Each of these cases requires different methods. A scientist may examine the shift of tectonic plates by comparing the distances on the order of tens of kilometers over the time period of a year, when the accuracy is below a millimeter.

Whenever we consider expansion of operational definitions for an empirical concept C to a new range R, the situation is similar:

(1) we can observe objects in R for which C cannot be measured with the needed accuracy;

(2) some other attributes $A_1, ..., A_n$ of objects in R can be measured, or else those objects would not be empirically available;

(3) some of $A_1, ..., A_n$ are linked to C by empirical laws or theories. We can use one or more of those laws in a new method: measure some of $A_1, ..., A_n$ and then use laws to compute the value of C.

Consider the task: determine distance D from Earth to each in a set R of galaxies, given some of the measured properties of R: $A_1, A_2, ..., A_n$. Operational definitions for $A_1, ..., A_n$ are available in the range R. For instance, let A_2 measure the redshift of hydrogen spectrum. Let $D = h(A_2)$ be Hubble Law. The new method is:

```
For a galaxy g, when no individual cefeids can be distinguished:
Measure A2 of the light coming from g by a method of spectral analysis
Compute the distance D(Earth, g) as h(A2(g))
```

The same schema can yield other operational definitions that determine distance by properties measurable in a new range, such as yearly parallax, perceived brightness or electromagnetic spectrum.

Some laws cannot be used even though they apply to galaxies. Consider $D = a/\sqrt{B}$ (B is brightness). It applies even to the most remote sources of light. But B used in the law is the absolute brightness at the source, not the brightness perceived by an observer. Only when we could determine the absolute brightness, we could determine the distance to galaxies by $D = a/\sqrt{B}$.

The following algorithm can be used in many applications:

```
Algorithm:
Input: set of objects observed in range R
       attribute C that cannot be measured in R
       set of attributes A1,...,Ak that can be measured in R
       set {F1,...,Fp} of known operational definitions for C
       set LAWS of known empirical laws
Output: a method by which the values of C can be determined in R

Find in LAWS a law L in which C occurs
   Let B1,...,Bm be the remaining attributes that occur in L
   Verify that C can be computed from L, and the values of B1,...,Bm
   Verify that {B1,...,Bm} is subset of {A1,...,Ak},
      that is, B1,...,Bm can be measured in at least some situations in R
   Use L and B1,...,Bm to create new procedure F for C
   Make F consistent with procedures in {F1,...,Fp}
```

After the first such procedure has been found, the search may continue for each law that involves C.

In set-theoretic terms, each expansion of concept C to a new range R can be viewed as a mapping from the set of distinguishable classes of equivalence with respect to C for objects in R to a set of possible new values of C, for instance, the values larger than those that have been observed with the use of the previous methods. But possible expansions are unlimited. The use of an existing law narrows down the scope of possible concept expansions to the number of laws for which the above algorithm succeeds. But the use of an existing law does not merely reduce the choices, it also justifies them. Which of the many values that can be assigned to a given state corresponds to its temperature? If laws reveal the real properties of physical objects, then the new values which fit a law indicate concept expansion which has a potential for the right choice.

Claim 22: Whenever the empirical methods expands to new territories, new discoveries follow. New procedures are instrumental to that growth.

Claim 23: Each new procedure expands the law it uses to a new range. If procedures P_1 and P_2 use laws L_1 and L_2 respectively, and produce empirically inconsistent results for new objects in range R, the choice of P_1 will make L_2 false in R.

If a number of procedures provide alternative concept expansions, various selection criteria can be used, depending on the goal of research.

Claim 24: Among two methods, prefer the one which has a broader range, for it justifies concept expansion by a broader expansion of an existing law.

Claim 25: Among two methods, prefer the one which has a higher accuracy, since it provides more accurate data for the expansion of empirical theories.

Claim 26: Methods must and can be verified in their new area of application or else, the empirical laws they apply would be mere definitions.

8 Operational definitions apply to all empirical concepts

While explicit operational definitions are rarely formed by experimental scientists, they become necessary in autonomous robots. A robot explorer can also benefit from mechanisms for generation of new procedures.

Operational meaning applies to databases. They are repositories of facts that should be shared as a major resource for knowledge discovery and verification. But data and knowledge can be only useful for those who understand their meaning. Operational definitions describe how the values of all fields were produced.

Similarly to our science examples, operational definitions can be generated from data and applied in different databases. Consider a regularity L, discovered in a data table D, which provides accurate predictions of attribute C from known values of $A_1, ..., A_n$. L can be used as a method that determines values of C.

Consider now another table D_1, that covers situations similar to D, but differs in some attributes. Instead of test C, tests $B_1, ..., B_m$ are provided, which may or may not be compatible with C. Suppose that a doctor who has been familiar with test C at his previous workplace, issues a query against D_1 that includes attribute C which is not in D_1. A regular query answering mechanism would fail, but a mechanism that can expand operational meaning of concepts may handle such a query (Ras, 1997). A quest Q for operational definition of concept C with the use of $B_1, ..., B_m$ will be send to other databases. If an operational definition is found, it is used to compute the values of C in the doctor's query.

References

Bridgman, P.W. 1927. *The Logic of Modern Physics*. The Macmillan Company.

Carnap, R. 1936. Testability and Meaning, *Philosophy of Science, 3*.

Huang, K. & Zytkow, J. 1997. Discovering Empirical Equations from Robot-Collected Data, Ras Z. & Skowron A eds. *Foundations of Intelligent Systems*, Springer, 287-97.

Jackson, P. 1990. *Introduction to Expert Systems*, Addison-Wesley.

Kulkarni, D., & Simon, H.A. 1987. The Process of Scientific Discovery: The Strategy of Experimentation, *Cognitive Science, 12*, 139-175.

Langley, P.W., Simon, H.A., Bradshaw, G., & Zytkow J.M. 1987. *Scientific Discovery; An Account of the Creative Processes*. Boston, MA: MIT Press.

Ras, Z. 1997. Resolving queries through cooperation in multi-agent systems, in eds. T.Y. Lin & N. Cercone, *Rough Sets and Data Mining*, Kluwer Acad. Publ. 239-258.

Shrager, J. & Langley, P. eds. 1990. *Computational Models of Scientific Discovery and Theory Formation*, Morgan Kaufmann Publ.

Sleeman, D.H., Stacey, M.K., Edwards, P., & Gray, N.A.B., 1989. An Architecture for Theory-Driven Scientific Discovery, *Proceedings of EWSL-89.*

Valdes-Perez, R. 1995. Generic Tasks of scientific discovery, *Working notes of the Spring Symposium on Systematic Methods of Scientific Discovery*, AAAI Technical Reports.

Żytkow, J.M., Zhu, J. & Hussam A. 1990. Automated Discovery in a Chemistry Laboratory, in: *Proc. Eight National Conf. on Artificial Intelligence*, AAAI Press, 889–894.

Żytkow, J., Zhu, J. & Zembowicz, R. 1992. Operational Definition Refinement: a Discovery Process, *Proc. 10th Nat'l Conf. on Artificial Intelligence*, AAAI Press, 76-81.