Measuring the unknown: knowledge-driven discovery of concept expansions

Jan M. Żytkow[1]

Abstract. Robot-discoverers and other intelligent systems must be able to interact with the world in complex, yet purposeful and accurate ways. Knowledge representation which is internal to a computer system lacks empirical meaning and thus it is insufficient for the investigation of the external world. We argue that operational definitions are necessary to provide empirical meaning of concepts. The research on automation of discovery has largely ignored operational definitions. In this paper we outline the scientific mechanism of operational definitions and the ways in which parts of the mechanism can be implemented. Individual operational definitions can be viewed as algorithms that interact with the real world. They can and they should be improved in the course of real-world interaction. 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 coherence of a set of definitions. No concrete set of operational definitions is complete. We demonstrate that expanding the operational definitions is a key task in science. We explain why among many possible expansions only a very special few lead to a satisfactory growth of scientific knowledge. Further, we demonstrate that the criteria of bootstrap verification which are very important in application to scientific laws, apply to operational definitions as well. While our examples come from natural sciences, where the use of operational definitions is especially clear, operational definitions apply in all cases of empirical concepts. We briefly argue their role in a robot-discoverer and in database applications.

1 Operational definitions provide empirical meaning of concepts

Data 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 the data have been procured, until problems occur. When a set of data is inconsistent with an accepted theory, 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 the data can be wrong in so many ways, measurement methods must be scrutinized closely in all branches of science. Separate scientific areas concentrate on the measurement methods, for instance metrology and analytical chemistry.

A common situation, critical to the growth of scientific knowledge, occurs when we want to investigate situations for which no known method can measure a particular quantity. For instance, we 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. Not only we need new measurement methods for new areas. We must be also able to determine that they expand the existing concepts. For instance, we must demonstrate that a new method produces temperature measurements on a publicly shared scale of temperature.

When no known method applies, new methods must be discovered. We use the term "discovery" rather than "invention" since measurement methods can be verified, as we will demonstrate later. Discovery of new methods, which we also call operational definitions, is the central problem in this paper. We present a systematic solution to the quest for new methods of measurement. 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.

At the end of each section we summarize, in the form of a few claims, the basic facts about measurement methods.

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

Claim 2: The actual verification in empirical science is limited to what can be empirically examined. The scope of operational definitions determines 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.

Among all scientific concepts, in this paper we restrict our attention to 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: In additions to methods that return values of properties, all other types of empirical concepts also require operational definitions, for instance relations between empirical objects and special situations, that may be recognized or created, such as the triple point of water.

[1] Computer Science Department, UNC Charlotte, Charlotte, N.C. 28223, U.S.A. and Institute of Computer Science, Polish Academy of Sciences; zytkow@uncc.edu

ECAI 98 Workshop Programme.

Claim 5: Some operational definitions provide data; other definitions prepare objects that possess specific properties.

In this paper we limit our attention to objects, but other entities can be treated in an analogous way.

Claim 6: Operational definitions apply to objects, states, events, locations and other empirical entities.

2 The research on automation of discovery 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 can be used 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 treated formally as tuples in the empirical space $\mathcal{E}$ of the values of independent and dependent variables plus (2) a mechanism of data procurement.

Similarly to scientists, BACON and other discovery systems use plans to propose sequences of experiments. Each experiment in $\mathcal{E}$ consists in preparing an empirical situation, described by a list of values $x_1, ..., x_k$ of empirical variables $X_1, ..., X_k$, and by measuring the value y of a dependent variable Y which provides the "world response" to the empirical situation characterized by $x_1, ..., x_k$.

Instead of real experiments, BACON uses two mechanisms to procure its data. In one mechanism the list of values of independent variables becomes the argument to a print statement to which a user must respond by typing the appropriate value of the dependent variable. The other mechanism substitutes simulation for the real data generation. The values of independent variables are passed on as the arguments to a function call that computes the simulated value of the dependent variable.

This treatment bypasses the complexity of operational definitions and disregards their role in the scientific inquiry. In the wake of robotic discovery systems, operational semantics must, at the minimum, provide realistic methods to acquire data. Even papers and collections that consider many components of the scientific methods and their interrelations (Kulkarni & Simon, 1987; Sleeman, Stacey, Edwards & Gray, 1989; Shrager & Langley, 1990; Valdes-Perez, 1995) neglect operational definitions of concepts.

Both of the BACON's mechanisms bypass the complex issues involved in real experimentation and measurements. Little has been done thus far to remedy this deficiency of discovery systems. Żytkow, Zhu & Zembowicz (1992) presented a robotic mechanisms in which experiments have been conducted automatically by a buret, a timer and a balance under the control of FAHRENHEIT (Żytkow 1996). A discovery process that uses the primitive readings of balance and timer and primitive actions of buret has been applied to the refinement of 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 drove the robot arm so that it could deposit a cylinder at different, precisely prescribed locations on the top of an inclined plane, while another procedure measured the time interval in which the cylinder rolled to the bottom of the plane.

While operational semantics must be added to any formalism, so that it applies to the real world, it has been neglected in the entire domain of AI, which uses formal semantics, represented by symbolic structures. As Jackson (1990) puts it: "a well-defined semantics ... reveals the meaning of ... expressions by virtue of their form." But this simply expands the same problem to yet another, 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 on operational definitions must be conducted to capture the mechanisms in which they are used in science and to make them applicable on intelligent robots. In this paper we concentrate on the a set of definitions that prescribe measurements of an individual quantity. For convenience, we use the terms quantity, property and magnitude as synonyms.

Claim 7: Formal semantics are not insufficient to provide empirical meaning.

Claim 8: Robotic discoverers require operational definitions.

3 Operational definitions are algorithms that interact with the real world

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

(1) If x is in water then (x is soluble in water iff x dissolves)

where "iff" is a shorthand for "if and only if". But a more adequate account of the scientific way of determining solubility is operational rather than descriptive:

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

Statement (1) can be used as the descriptive representation of the procedure Soluble(x).

An operational definition, treated as an algorithm, consists of instructions that at the lowest level prescribe manipulations, measurements and computations on the results of measurements. Loop instructions can be used to enforce the loop exit conditions, 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 measuring rod with a given object, marks the end of the measurement process. The measured quantity is returned from such a loop.

We can distinguish two types of procedures that differentiate between independent and dependent variables. They can be contrasted as manipulation and measurement mechanisms. Each independent variable can be set by a manipulation mechanism to specific values, while the response values of each dependent variable are obtained by a measurement mechanism. In this paper we focus on the measurement procedures.

It may happen that a particular instruction within the procedure P would not work in a specific situation. The thermal equilibrium cannot be reached or thermometric liquid freezes. 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 and above 100C; it can only measure temperature of objects in direct thermal contact. Risking oversimplification, we can represent the range of a procedure as the Cartesian product of the ranges of feasibility of all instructions in the procedure. Let us call the range of procedure P 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 and then the distance is calculated as a product of time and velocity. Let $C(x)$ is 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 measured or generated by instructions within P, and f is a computable function on those values.

Claim 9: Each operational definition can be treated as an algorithm.

Claim 10: The range of each procedure is limited in many ways, thus each operational definition P is a partial definition applicable in the range R_P.

Claim 11: 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 12: An operational definition for 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 Many operational definitions are needed for each concept

In many everyday situations, distance can be measured by a yard-stick or a tape. But for objects divided by a river a triangulation method will work better. Measuring rod cannot be used to find the distance from the Earth to the Sun and to the Moon, while the triangulation method will work in both cases. 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. Some of the stars within the range of triangulation are called cefeids. They pulsate and through detailed modeling it has been determined that their absolute luminosity and their period of pulsation are linked by an empirical law. Another law, determined on Earth and applied to stars claims that the perceived brightness of a given 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 has been empirically discovered. It claims proportionality between the distance and so called red shift. The law of red shift has been used to determine the distance of the galaxies so distant that cefeids cannot be distinguished.

But when galaxies are even more remote, the cefeids cannot be distinguished. The Hubble Law of proportionality between the red shift in the lines of the hydrogen spectrum and the distance from Earth has been used to determine the distance of those galaxies.

Similarly, while a gas thermometer applies to a large range of states, there are low temperature states in which 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 states of high temperature in which no vessel can hold the gas, and states in which the inertia of gas thermometer has the unacceptable influence on the measured temperature. The measurements of thermal radiation and other methods can determine temperature in many such cases.

Claim 13: Empirical meaning of a concept is defined by a set of operational definitions.

Claim 14: Each concrete set is limited and new methods must be constructed for objects beyond those limits.

5 Operational definitions can be empirically and theoretically equivalent

Consider two operational definitions P_1 and P_2 that measure the same quantity C. When they apply to the same objects their results should be empirically equivalent within the accuracy of measurement. If, applied to the same objects, P_1 and P_2 provide different results, one or both methods 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 the direct use of measuring rod or tape. But in addition to the empirical study of equivalence, the procedures can be compared with the use of empirical theories and their results may be proven to be the same.

Triangulation uses a basic theorem of Euclidean geometry: one side and two adjacent angles in a triangle uniquely determine the remaining sides. This justifies theoretically the consistency of two methods: by the use of yard-stick and by triangulation. Since, or 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 rod or a tape while the other using triangulation, the results are consistent.

Even if a procedure returns a value, that value may not be physically adequate. The measurement process affects the measured quantity. This principle has been justified by quantum mechanics, but it must be considered even in the range of classical measurements. When a thermometer reaches thermal equilibrium with the measured body b, the temperature of b can be changed drastically. The thermal inertia of a thermometer should be adequate to the task.

Claim 15: Methods can differ by their accuracy and the degree to which they influence the measured quantity.

Claim 16: Inadequacy of the measured results further limits the range of operational definitions.

Claim 17: When two operational definitions define the same property and apply to the same objects, their results should be empirically equivalent.

Claim 18: 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 19: 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 20: 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 th actual equivalence of P_1 and P_2.

Claim 21: Proving the equivalence of two procedures is desired, because often the empirical verification of equivalence is costly or practically impossible.

6 Operational definitions of a concept must form a coherent set

We have considered several examples of 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? The meaning 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 $\phi\text{-}i_1$,...,$\phi\text{-}i_k$, such that $\phi\text{-}i_1 = \phi_i$, $\phi\text{-}i_k = \phi_j$, and for each $m = 2, ..., k$ the ranges of $\phi\text{-}i_m$ and $\phi\text{-}i_{m+1}$ 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 red shift and Hubble Law.

Similarly, the measurements with gas thermometer have been used to calibrate the alcohol and mercury thermometers in order to enforce their equivalence 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 also overlap with, and are empirically equivalent to measurements with the use of gas thermometer.

Claim 22: Each empirical concept should be defined by a coherent set of operational definitions. Historically there are cases when the coherence is missing, but the discovery of a missing link is perceived as a challenge.

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

Claim 23: By examining a coherent set of operational definitions we can demonstrate that the values measured by any two procedure in the set are on the same scale.

Claim 24: In the case of conflict between different methods, additional data are collected and methods are adjusted in order to restore the coherence.

Claim 25: Operational definitions provide means to expand to new areas the range of the laws they use.

7 Expanding the operational definitions is a key task in science

Whoever wants to discover a pattern in the distribution of galaxies, must be able to measure the distances between them. A scientist who wants to empirically examine the shift of tectonic plates may do so by comparing the distances on the order of several tens of kilometers over the time period of a year if the accuracy of measurement is below a millimeter.

For each concept, operational definitions can be expanded in several obvious directions, to reach very small values, very large values, and values that are very precise. But the directions are far more numerous. Within the range of "room" temperatures, consider the temperature inside a cell, or temperature of a state that is fast varying and must be measured every second. Or consider the measurement of temperature on the surface of Mars. Each of these cases requires different methods.

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

(1) we can observe objects in a new range R for which C cannot be measured with the sufficient 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 at all;

(3) some of the measured properties 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 the laws to compute the value of C. As an example, consider

```
The task:
   determine distance D from Earth
      to each in a set R of galaxies,
   given some of the measured properties of R:
      A1, A2, ..., An;

   (this presupposed operational definitions
      for A1,...,An in the range R)
   For instance, let A2 be the quantity of
      the redshift of hydrogen spectrum.
   Let D=h(A2) 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 known
      method of spectral analysis
   Compute the distance D(Earth, g) as h(A2(g))
```

While various other laws apply to galaxies, some cannot be used. Consider the law $D = a/\sqrt{brightness}$. This law applies even to the most remote sources of light. But the brightness that is used in the law is the absolute brightness at the source, not the brightness perceived by an observer. Should we have a way of determining the absolute brightness, we could use the inversed square root law to determine the distance to galaxies. When the distance to galaxies can be determined, however, we know from observations that that galaxies at the same distance have different brightness.

A schema similar to the application of the Hubble Law applies to other operational definitions that determine distance. Observable properties measurable in a new range can be yearly parallax, perceived brightness, shape, electromagnetic spectrum, and so on. The same 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

Seek 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:

```
Seek all alternative laws that provide operational
     definitions of C
  For each such law, build a procedure for 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 the 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 the set of such possible expansions is huge. The use of an existing law narrows down the scope of possible concept expansions to the number of laws for which the above algorithm works. 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 objectively existing properties of physical objects, then the new values which fit a law indicate concept expansion which has a potential for the right choice.

Claim 26: Whenever the empirical inquiry expands to new territories, new discoveries follow. New procedures are instrumental to that growth.

Claim 27: Each new procedure expands the existing law it uses to a new range of applications.

If a number of procedures provide alternative concept expansions, various selection criteria can be used, depending on the goal of research. Operational definitions can differ by their range, accuracy, the degree to which they can be verified in the new area, and so on.

Claim 28: Among two methods, one which has a broader range is preferred, for it better justifies concept expansion by a broader expansion of an existing law.

Claim 29: Among two methods, one which has a higher accuracy is preferred, for it provides more accurate data and thus a stronger empirical foundation for the expansion of empirical theories.

Claim 30: The range of verification of each empirical law is limited to the area where objects to which the law applies are empirically available. We can claim that the law applies to all objects described in the law, but we cannot verify the law until we find a way to observe such objects.

Claim 31: Methods must be verified in their new area of application or else, the empirical laws they apply would be mere definitions.

Claim 32: If two 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.

The last two claims bring about the subject of bootstrap confirmation. In the next two sections we will explain how bootstrap verification applies to empirical laws and then we expand bootstrap verification to operational definitions.

8 Bootstrap verification applies to scientific laws

When a scientific law is used in an operational definitions, is it reduced to a definition? This has been considered a serious problem and led to conventionalism, according to which the laws of science are merely conventions and can be evaluated by their utility rather than truth.

The definitional use of scientific laws has been analyzed by philosophers of science. They focused on the situation which is very common in the scientific inquiry. When a scientific hypothesis H is tested, often H itself is applied in the process (Glymour, 1980, Simon 1970). This practice is common because scientific theories use intrinsic variables (Simon 1970), which represent properties that are not observable. The numerical values of intrinsic variables are determined indirectly through the empirical laws applied to observational data. Those numerical values are then used to verify the laws through which they have been determined.

Glymour (1980) named this process "bootstrap confirmation". Even though bootstrap confirmation is commonly applied by scientists, from the logical perspective the practice may look circular, and a number of people participating in the discussion of bootstrap confirmation expressed this concern (Edidin, 1988, Christensen 1990).

We will use the term bootstrap verification and argue that this is a non-circular, sound method of hypotheses verification. Let us start from making a few distinctions.

Since hypotheses are seldom tested in isolation, a realistic definition of hypothesis verification must include a background theory T. T can include some, none, or all additional knowledge, that can be used in operational definitions that transform empirical facts from their raw form to the form expressed in the language of the tested hypothesis. Scientists make notorious use of the previously accepted knowledge in construction of measuring and manipulating instruments, and in converting raw facts to theoretically meaningful facts. In the measurement of any scientific magnitude there are situations in which the values are hidden from direct observation, and must be obtained by other measurements and application of our knowledge.

With the background knowledge T included, bootstrap confirmation is a three argument relation: "Evidence E confirms hypothesis H with respect to theory T", allowing to speak about the use of H and established background knowledge T in the process of testing H.

In his 1980 paper Glymour confronts circularity with a simple intuitive criterion for valid bootstrap confirmation:

> hypothesis H is bootstrap confirmed by evidence E when a variation E' of E is possible that would falsify H, that is, should the experimentation produce results E' different from E, H would be disconfirmed by E',

where any piece of evidence E is a conjunction of ionic statements (atomic and their negations), while E' is a modification of E, that is, E' can be obtained from E by negating some of facts in E. Cases of redundant conjunctions such as E=$Pa\&Pa$, which lead to inconsistent E'=$Pa\&\neg Pa$ must be excluded.

Because serious and justified objections against Glymour's definition have been raised by Christensen (1983, 1990) we will introduce another definition, which eliminates those objections.

We need some more terminology: Take an alternative F of ionic facts. F' is a modification of F if it is an alternative in which one or more of the facts in F are negated. For instance if F is $(Pa \vee Qa)$, then F' can be $(\neg Pa \vee Qa)$. Notice that the spirit of the modification from an alternative F to F' is the same as from a conjunction E to E'. Alternatives of ionic facts are important because they can be used to express typical observational conclusions of theories. $(x)(Rx \rightarrow Bx)$ implies $(Ra \rightarrow Ba)$, that is $(\neg Ra \vee Ba)$.

```
Definition of Bootstrap Confirmation:
   E bootstrap confirms H with respect to
   (additional) background knowledge T
   iff
   the following conditions 1, 2, 3 and 4 are
   jointly satisfied:
   1. E is an observational evidence (true,
      conjunction of simple facts)
   2. E, H, and T are jointly consistent;
   3. there is a non-tautological consequence F of
      T&H, such that
      a. E->F is valid, and
      b. no non-tautological modification F' of F
         is in T&H;
   4. a modification E' of E is logically possible
      such that
      a. E' is inconsistent with T&H, and
      b. E' is not inconsistent with T
```

This definition is different from the previous attempts in several ways. First, even if H is used in testing H, H does not have to be a part of T. In fact, it is wiser not to put H in T. Second, by 3.b we eliminate observational consequences of H and/or T, of the form $Pa \vee Qa$, where Q has nothing to do with H or T, because then both Qa and $\neg Qa$ can be added to Pa without inconsistency. Third, 3.a captures the normal way in which facts are confronted with elementary predictions of a theory. Facts are conjunctions, while predictions take the form of elementary implications or alternatives, which should follow from facts. Fourth, 4.b excludes all Christensen's counterexamples (1983, 1990), while retaining all his positive examples.

9 Bootstrap verification applies to operational definitions

We can now apply the idea of bootstrap verification to operational definitions. When we expand a concept to the new range of values, is any operational definition acceptable? If we do not have verification criteria how can we chose among competing, mutually inconsistent operational definitions $P_1, ..., P_n$? We will now use bootstrap verification to provide answers to those questions.

In each of our examples a concept measurement is extended to a new range with the use of an existing law and methods of measurement for other concepts that occur in that law. Some of these laws can be used in many ways, leading to bootstrap verification.

Consider the use of the Law of Momentum Conservation as a definition of mass:

$$m(b)/m(a) = [V(a,i) - V(a,j)]/[V(b,j) - V(b,i)]$$

where masses of objects a, b, ... are determined by their velocities observed at time i, j, ... After one object is chosen as a unit of mass, the masses of other objects can be determined by collision experiments. Many experiments with the use of the same objects a and b collided with different initial velocities can lead to different ratios of $m(b/m(a)$, providing tests required by bootstrap verification. Experiments demonstrate, however, that those ratios are approximately equal. Similar experiments with the use of further objects c, d, and so forth, lead to further cases of verification.

As our second example, consider the Curie-Weiss law $\chi = A/(T-B)$ used in order to expand the measurements of temperature T to very low values with the help of measurements of magnetic susceptibility χ. A and B are two constants characteristic for a given paramagnetic salt. They have been measured in the range of temperatures that can be determined with the use of gas thermometer and can be used in lower temperatures. Since different paramagnetic salts can be used as thermometric substances, the values of temperature used by each paramagnetic thermometer can be different. When compared, independently measured values provide bootstrap verification of the method based on the Curie-Weiss Law.

10 Operational definitions apply in all cases of empirical concepts

10.1 Robot-discoverer applications

While operational definitions are rarely formed in explicit form by human experimental scientists, they become necessary when we want to collect data by robotic equipment.

When a robot explorer realizes that a given operational procedure P cannot be applied in a given situation, the robot should be able to use its knowledge and replace P by another equivalent procedure P_1. Using our computational mechanism, the robot can find out that P and P_1 are not only theoretically equivalent with regard to robot's knowledge, but are also empirically equivalent in the test situations when both can be used.

10.2 Database applications

Operational meaning is necessary to utilize a database. Databases are repositories of facts. They should be shared publicly or among allied institutions as a major resource for knowledge discovery and verification. But data and knowledge can be only useful for those who understand their meaning. That includes the semantics of all the attributes and understanding of the situations described by data: "By what procedure have the values in a given field been produced?"

Operational definitions can be generated from data and applied in different databases. Some regularities discovered in data are especially strong, and can provide unique predictions of some attributes. Consider such a regularity L, discovered in a data table D, which provides predictions of the values of C from known values of $A_1, ..., A_n$. L can be used as a method that determines values of C. That method is consistent with the original method through which the values of C have been determined, because L is a regularity that allows to compute accurately the known values of C.

Consider now another data table D_1, that contains data which describe the same situations but uses slightly different attributes than D. Instead of test C, some other 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, uses D_1, and issues a query that includes attribute C which is unknown in D_1. A regular query answering mechanism cannot answer such a query, but a mechanism that can expand operational meaning of concepts will handle such a query (Ras, 1997). A quest for operational definition will be send to other databases:

(Q1) "Discover a definition of concept C with the use of $B_1, ..., B_m$".

In response to that quest, database D will (1) verify that attributes $B_1, ..., B_m$ have the same operational meaning in both databases, (2) use the quest (Q1) to invoke a discovery mechanism on its data, (3) determine that a regularity L has been discovered that can be used as the requested operational definition, and (4) send back L to database D_1. Such relation can take on a form of a system of rules, an equation, a taxonomy and so on.

This mechanism can be used even in the original database D_1, if some values of C are available, to compute the missing values of C.

References

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

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

Christensen, D. 1983, Glymour on Evidential Relevance, *Philosophy of Science*, 50:471-81.

Christensen, D. 1990, The Irrelevance of Bootstrapping, *Philosophy of Science*, 57:644-62.

Edidin, A. 1988. From Relative Confirmation to Real Confirmation, *Philosophy of Science*, 55, pp. 265-271.

Glymour, C. 1980, Theory and Evidence, Princeton, New Jersey: Princeton University Press.

Huang, K-M. & Zytkow, J.M. 1997. Discovering Empirical Equations from Robot-Collected Data, Ras Z. & Skowron A eds. *Foundations of Intelligent Systems*, Springer, p. 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 Academic Publishers, pp. 239-258

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

Simon, H.A. 1970. The Axiomatization of Physical Theories, *Philosophy of Science*, 37:16-26.

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. 1996. Automated Discovery of Empirical Laws, *Fundamenta Informaticae, 27*, p.299-318.

Żytkow, J.; and Zembowicz, R. 1993. Database Exploration in Search of Regularities. Journal of Intelligent Information Systems 2:39-81.

Żytkow, J.M., Zhu, & Zembowicz, 1992. Operational Definition Refinement: a Discovery Process, *Proc. Tenth National Conference on Artificial Intelligence*, AAAI Press, 76-81.