From Theory to Technology: Rules versus Exemplars

Paper for the PSA'2004 Workshop on "Applying Science", Austin TX

Rens Bod
Institute for Logic, Language and Computation
University of Amsterdam, and
School of Computing, University of Leeds
rens@science.uva.nl

Abstract

How is scientific knowledge used, adapted and extended in deriving real-world systems and technological devices? This paper aims at developing a general model of "applying science" based on the Exemplar-Based Explanation (EBE) model. EBE embodies the view that a real-world system is derived not by solving theoretical laws for specific boundary conditions but by constructing the system out of previously derived systems that function as exemplars. I will discuss a number of artifacts from hydraulics and language technology, and develop an instantiation of EBE which generalizes over different disciplines. I argue that engineering practice is highly cumulative: new systems are almost literally built upon and constructed out of previous systems resulting into increasingly complex wholes.

1. Introduction

How do we get from theory to technology? That is to say, how is scientific knowledge used, adapted and extended in deriving real-world systems and technological devices? It has for a long time been assumed that technological devices are derived by solving the laws from pure science for specific boundary conditions (see Boon 2004 for an overview). According to this view, technology is taken as "applied science" (Bunge 1966). Yet it has become increasingly clear that in deriving a real-world system we do not simply solve theoretical laws for specific boundary conditions. Instead, we also add *non-theoretical* elements, such as intermediate models, corrections, normalizations and other adjustments, that stand in no deductive relation to laws (see Cartwright 1983, 1999; Boumans 1999; Morrison 1999 and the many examples therein). Applying a scientific theory to a concrete system is a matter of intricate approximation and de-idealization for which no general rules are known. How then do we derive a real-world system in engineering? In line with Ronald Giere, Thomas Nickles and others, I argue that engineers work in a model-based way. In deriving a new system they look for known systems that are in various ways similar to the new system. The models and techniques that successfully accounted for the known systems are extended and adapted to the new systems (see Giere 1988, 1999; Nickles 2003). Such known systems function as *exemplars* on which new systems are modeled.

The notion of exemplar is usually attributed to Thomas Kuhn in his account on normal science (Kuhn 1970). Kuhn urged that exemplars are "concrete problem solutions that students encounter from the start of their scientific education" (ibid. 1970, p. 187) and that "scientists solve puzzles by modeling them on previous puzzle-solutions" (ibid. p. 189). Scientists possess what Kuhn called "acquired similarity relations" that allow them "to see situations as like each other, as subjects for the application of the same scientific law or law-sketch" (ibid. p. 190). In Bod (2004), I proposed a formal, computational model of Exemplar-Based Explanation, termed EBE, which simulates this problem solving activity. EBE represents problem solutions by *derivation trees* which describe each step in linking laws to phenomena. Explanations of new phenomena are constructed out of largest possible derivational chunks from derivation trees of previous phenomena such that *derivational similarity* is maximized. I showed that EBE can solve a large variety of problems and phenomena. Yet the model runs into trouble as soon as non-theoretical elements such as corrections, approximations and normalizations are needed. For real-world systems and technological devices EBE is inadequate.

The current paper investigates what is involved in creating a computational model for deriving technological devices and real-world systems. In doing so, I will start off with the EBE model and explore how its shortcomings may be alleviated. I will show that as long as non-theoretical elements can be stated in terms of mathematical equations, they can be integrated in a derivation tree, resulting in a new EBE model. I argue that engineering practice is a highly cumulative enterprise: new systems are almost literally built upon and constructed out of previous systems resulting into increasingly complex wholes. I contend that this complexity is dealt with by taking derivations of previous systems as "given" and work from there, rather than working from theoretical laws.

This paper is organized as follows. In section 2, I will review EBE and show how it applies to problem solving in science. In section 3, I move to technology and discuss a number of concrete systems from hydraulics for which there are no formal derivations from higher-level laws but which involve empirical corrections and coefficients. I will develop a new EBE model which integrates theoretical and phenomenological modeling and which can derive a range of new hydraulic systems. In section 4, I will provide an excursion into a field at the other end of the technological spectrum, discussing some examples from language engineering. What counts for hydraulics also counts for language technology: real-world systems are derived not from theoretical laws, but from parts of derivations of previous systems. I contend that the new EBE model provides a general model of "applying science" for different technological disciplines.

2. Review of the exemplar-based explanation (EBE) model

Let's start by reviewing the EBE model in Bod (2004) with a simple idealized example. Consider the following derivation of the Earth's mass from the Moon's orbit in the textbook by Alonso and Finn (1996, p. 247):

> Suppose that a satellite of mass m describes, with a period P, a circular orbit of radius r around a planet of mass M. The force of attraction between the planet and the satellite is $F = GMm/r^2$. This force must be equal to m times the centripetal acceleration $v^2/r = 4\pi^2 r/P^2$ of the satellite. Thus,
>
> $$4\pi^2 mr/P^2 = GMm/r^2$$
>
> Canceling the common factor m and solving for M gives
>
> $$M = 4\pi^2 r^3/GP^2.$$

Figure 1. Derivation of the Earth's mass according to Alonso and Finn (1996)

This rather textual derivation can be formally represented by the derivation tree in figure 2.

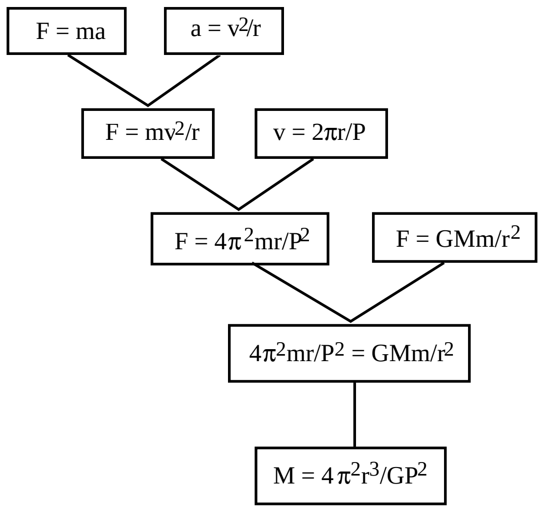

Figure 2. Derivation tree for the derivation in figure 1

This derivation tree represents the various derivation steps (insofar as they are mentioned in figure 1) from higher-level laws to an equation of the mass of a planet. In general, a derivation tree is a finite tree in which each node is labeled with a formula; the boxes are only convenient representations of these labels. The formulas at the top of each "vee" (i.e. each pair of connected branches) in the tree can be viewed as premises, and the formula at the bottom of each "vee" can be viewed as a conclusion, which in this tree is arrived at by simple term substitution. The last derivation step in the tree is not formed by a "vee" but consists in a unary branch which solves the directly preceding formula for a certain variable (in the tree above, for the mass M). The reader is referred to Baader and Nipkow (1999) for an overview on term rewriting and equational reasoning.

Note that a derivation tree captures the notion of a deductive-nomological (D-N) explanation of Hempel and Oppenheim (1948). In the D-N account, a phenomenon is explained by deducing it from general laws and antecedent conditions. But a derivation tree represents more than just a D-N explanation: there is also an implicit theoretical model in the tree in figure 2. A theoretical model is a representation of a phenomenon for which the laws of the theory are true (Suppes 1961, 1967; van Fraassen 1980). By equating the centripetal force of circular motion $4\pi^2mr/P^2$ with the gravitational force GMm/r^2 the model that is implied in figure 2 is a two particle model where one particle describes a circular orbit around the other one due to gravitational interaction and for which the mass of the first particle is negligible compared to the other.

EBE proposes that subtrees from derivation trees be productively reused to construct derivation trees for new phenomena. E.g. Kepler's third law, which states that r^3/P^2 is constant, can be derived by reusing the following subtree in figure 3 from the derivation tree in figure 2. This subtree reflects a theoretical model of a general planet-satellite or sun-planet system -- or any other orbiting system where the mass of the orbiting particle is negligible compared to the other.

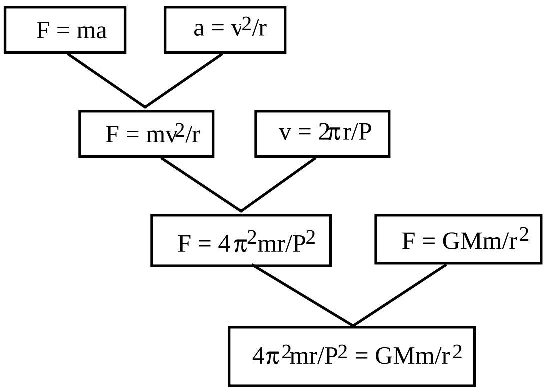

Figure 3. A subtree from figure 2 reflecting a theoretical model of a planet-satellite system

The subtree in figure 3 needs only to be extended with a derivation step that solves the last equation for r^3/P^2, as represented in figure 4.

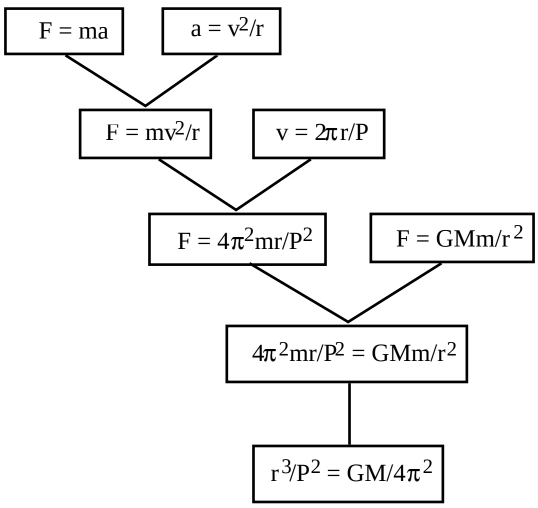

Figure 4. Derivation tree for Kepler's third law from the subtree in figure 3

Thus we can productively reuse partial derivations of previous phenomena to derive new phenomena. Instead of starting each time from scratch, we learn from previous derivations and reuse them for solving new problems. This is exactly what the exemplar-based view entails: a theory is viewed as a prior corpus of derivations or problem-solutions of exemplary phenomena (our body of physical knowledge, if you wish) by which new phenomena are predicted and explained. In a similar way we can derive the distance of a geostationary satellite, namely by solving the subtree in figure 3 for *r*.

However, it is not typically the case that derivations involve only one subtree. For example, in deriving the velocity of a satellite at a certain distance from a planet, we cannot directly use the large subtree in figure 3, but need to extract two smaller subtrees from figure 2 that are first combined by term substitution (represented by the operation "∘"[1]) and then solved for *v* in figure 5:

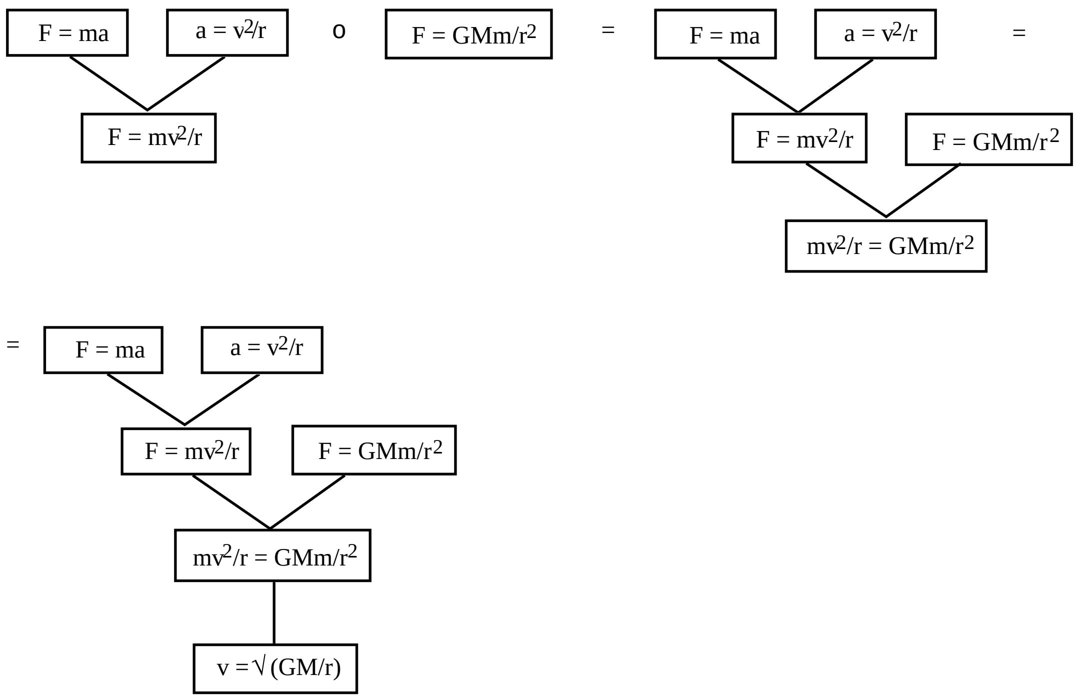

Figure 5. Constructing a derivation tree for a satellite's velocity by combining two subtrees from figure 2

[1] The *substitution operation* or *combination operation* "∘" is a partial function on pairs of labeled trees; its range is the set of labeled trees. The combination of tree *t* and tree *u*, written as $t \circ u$, is defined iff the equation at the root node of *u* can be substituted in the equation at the root node of *t* (i.e. iff the lefthandside of the equation at the root node of *u* literally appears in the equation at the root node of *t*). If $t \circ u$ is defined, it yields a tree that expands the root nodes of copies of *t* and *u* to a new root node where the righthandside of the equation at the root node of *u* is substituted in the equation at the root node of *t*. Note that the substitution operation can be iteratively applied to a sequence of trees, with the convention that ∘ is left-associative.

Although exceedingly simple, figure 5 shows that we can create new derivations by combining different parts from previous derivations, i.e. from exemplars. The result can in principle be used as an exemplar itself.

EBE can be summarized by the following two parameters: (1) a *corpus of derivation trees* representing exemplars and (2) a *matching procedure* that combines subtrees from the corpus by means of term substitution into a new derivation tree. (Of course there may be additional mathematical operations on the root node of each subtree, to which I will come back in the next section.) Note that subtrees can be of any size: from single equations to any combination of laws up to entire models and derivations. This reflects the continuum between rules, rule-schemes and entire exemplars in EBE.

Derivational Similarity

How does EBE "know" which subtrees from previous explanations can be reused in solving a new problem -- rather than trying out all combinations of subtrees? Kuhn's (and also Giere's) suggestion is that scientists see similarity relations or family resemblances between a new phenomenon and previous problem solutions and therefore know which law patterns can be applied to derive the new phenomenon. This similarity relationship remains a rather vague notion in most accounts and some believe it cannot be formalized or that seeking a formalization should be resisted (cf. Kuhn 1970, p. 192). How does EBE deal with this? Since EBE's matching procedure is entirely derivational, I interpret the notion of similarity in terms of *derivational* or *explanatory similarity*. Following Bod (2004), a new phenomenon is derivationally similar to a previously explained phenomenon if the derivation of the previous phenomenon can be (partially) reused to explain the new phenomenon, that is, if their resulting derivation trees share one or more subtrees. The larger the subtrees they share -- i.e. the larger the partial match between the two trees -- the more derivationally similar they are. Rather than trying to define similarity between phenomena, EBE focuses on the similarity between *derivations* of phenomena.

Still this similarity measure does not tell us which subtrees from previous derivation trees can be used to explain a newly presented phenomenon. Sure enough, EBE could exhaustively enumerate all possible combinations of subtrees that result in a derivation of the phenomenon, and next establish the similarity relations by determining the largest partial matches with previous derivations. But this is highly inefficient and unnecessary. In fact, there exist algorithms that can establish the largest partial match in one go by computing the so-called *shortest derivation* of a given phenomenon. The shortest derivation of a phenomenon is its derivation which consists of the smallest number of subtrees from the corpus. Since subtrees are allowed to be of arbitrary size, the shortest derivation corresponds to the derivation tree which consists of largest partial match(es) with previous derivation trees in the corpus. Given a set of subtrees and a (new) phenomenon, the shortest derivation can be computed by means of a so-called *best-first search* procedure which efficiently searches for the shortest path from root node (i.e. the phenomenon) to leaf nodes (i.e. laws or antecedent conditions) -- where path length is defined as the number of subtrees (see Bod 2000; Cormen et al. 2002). Thus derivational similarity can be maximized by minimizing derivation length.

EBE embodies the hypothesis that scientists try to explain a new phenomenon by maximizing derivational similarity between the new phenomenon and previously derived phenomena. And the shortest derivation provides a possible way to attain this goal. The rationale behind maximizing derivational similarity is that it favors derivation trees which maximally overlap with previous derivation trees, such that *only minimal recourse to additional derivational steps needs to be made*. Note that EBE does not imply that scientists actually use a best-first search algorithm in explaining new phenomena. The existence of such an algorithm only shows that it is possible *in principle* to find an explanation which is derivationally most similar to previous explanations.

For example, the phenomenon known as Kepler's third law in figure 4 is maximally similar (modulo equivalence) to the problem of deriving the Earth's mass from the Moon's orbit in figure 2, because only *one* big subtree (figure 3) from the Earth's mass problem is needed to derive Kepler's law. Even if the two problems may seem different to the layman, for the physicist they are nearly equal, except for the final solution of a certain variable. The new problem can be solved by almost entirely reusing a previous problem solution, which is in fact based on the same model. Also the phenomenon of the velocity of a satellite in figure 5 is quite similar to the Earth-Moon system, though somewhat less than Kepler's third law since it involves *two* (smaller) subtrees that result in a smaller fraction of common derivation steps, as can be seen by comparing resp. figures 5 and 4 with figure 2.

In passing I should note that it is also possible for one phenomenon to have different derivation trees, in which case the phenomenon may be called derivationally redundant. For example, it is well known that the problem of deriving the Earth's mass can be solved not only from the Moon's orbit but also from the gravitational acceleration of an object at the Earth's surface. Thus our notion of similarity only tells us something about the similarity of the *derivations* of phenomena rather than of the phenomena themselves. If one phenomenon has two different derivations with no common subtrees, then these derivations refer to different underlying models. We will come back to this at the end of section 3.

We should keep in mind that the phenomena discussed so far are highly idealized and limited to textbook examples. There is *no* historical analog of using parts from the derivation of the Earth's mass to solve Kepler's third law (it rather happened the other way round). Yet in science education the two problems are treated as closely interconnected, and with good reason: it shows that the two problems can be solved by using the same law schemes and underlying model. In the next section I will investigate how the EBE approach can be extended to real-world systems and technological devices. As an intermediate step, I could also have dealt with idealized phenomena that are *not* exactly solvable. A typical example is the three-body problem in Newtonian dynamics. Even if we make the problem unrealistically simple (e.g. by assuming that the bodies are perfect spheres that lie in the same plane), the motion of three bodies due to their gravitational interaction can only be approximated by techniques such as perturbation calculus. However, in perturbation calculus every derivation step still follows numerically from higher-level laws. The actual challenge lies in real-world systems for which there are derivation steps that are *not* dictated by any higher-level law.

3. Extending EBE to real-world systems and technological devices

Derivations of real-world systems and technological devices are strikingly absent in physics textbooks. But they are abundant in engineering practice. As an example I will discuss a concrete system from hydraulics: the velocity of a jet through a small orifice, known as Torricelli's theorem, and which I will also refer to as an *orifice system*. I have chosen this system because it functions as a shared example in hydraulics on which several other systems are modeled, and yet it has no rigorous solution from higher-level laws but involves additional empirical coefficients. I will show how a "derivation" of the orifice system is used as an exemplar for deriving a range of other real-world systems, such as weirs, notches and water breaks.

The orifice system is usually derived from Daniel Bernoulli's famous equation, which in turn is derived from the Principle of Conservation of Energy.[2] According to the Principle of Conservation of Energy the total energy of a system of particles remains constant. The total energy is the sum of kinetic energy (E_k), internal potential energy ($E_{p,int}$) and external potential energy ($E_{p,ext}$):

$$\Sigma E = E_k + E_{p,int} + E_{p,ext} = constant$$

Applied to an incompressible fluid, the principle comes down to saying that the total energy per unit volume of a fluid in motion remains constant, which is expressed by Bernoulli's equation:

$$\rho gz + \rho v^2/2 + p = constant$$

The term ρgz is the external potential energy per unit volume due to gravity, where ρ is the fluid's density and z the height of the unit (note the analogy with mgh in classical mechanics). The term $\rho v^2/2$ is the kinetic energy per unit volume (which is analogous to $mv^2/2$ in classical mechanics). And p is the potential energy per unit volume associated with pressure. Bernoulli's equation is also written as

$$\rho gz_1 + \rho v_1{}^2/2 + p_1 = \rho gz_2 + \rho v_2{}^2/2 + p_2$$

which says that the total energy of a fluid in motion is the same at any two unit volumes along its path.

Here is how the engineering textbook *Advanced Design and Technology* presents the derivation of Torricelli's theorem from Bernoulli's equation (Norman et al. 1990, p. 497):

> We can use Bernoulli's equation to estimate the velocity of a jet emerging from a small circular hole or orifice in a tank, Fig. 12.12a. Suppose the subscripts 1 and 2 refer to a point in the surface of the liquid in the tank, and a section of the jet just

[2] Bernoulli used a precursor of this principle which was known as "Equality between the Potential Ascent and Actual Descent" (see Mikhailov 2002, p. 70).

outside the orifice. If the orifice is small we can assume that the velocity of the jet is *v* at all points in this section.

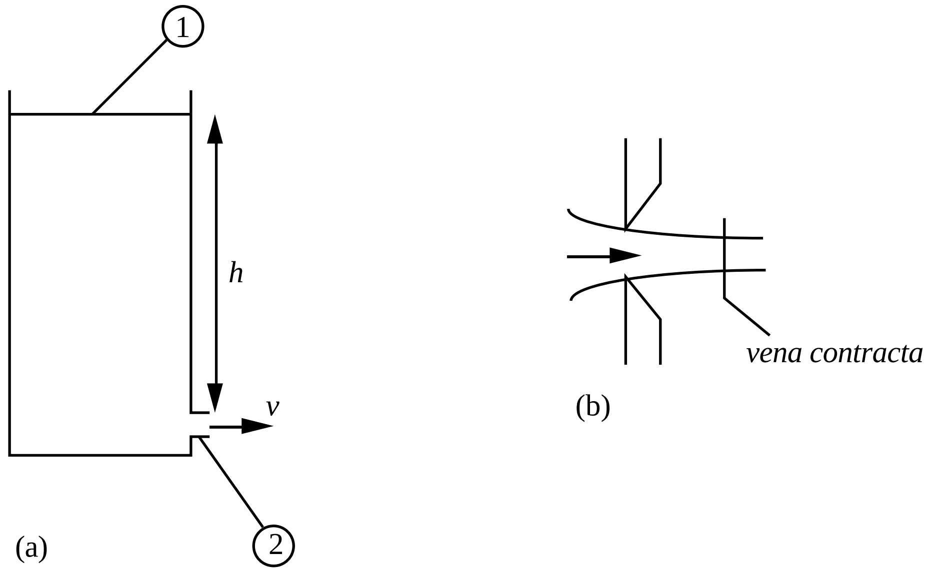

Figure 12.12

The pressure is atmospheric at points 1 and 2 and therefore $p_1 = p_2$. In addition the velocity v_1 is negligible, provided the liquid in the tank has a large surface area. Let the difference in level between 1 and 2 be h as shown, so that $z_1 - z_2 = h$. With these values, Bernoulli's equation becomes:

$h = v^2/2g$ from which $v = \sqrt{(2gh)}$

This result is known as Torricelli's theorem. If the area of the orifice is A the theoretical discharge is:

$Q(\text{theoretical}) = vA = A\sqrt{(2gh)}$

The actual discharge will be less than this. In practice the liquid in the tank converges on the orifice as shown in Fig. 12.12b. The flow does not become parallel until it is a short distance away from the orifice. The section at which this occurs has the Latin name *vena contracta* (*vena* = vein) and the diameter of the jet there is less than that of the orifice. The actual discharge can be written:

$Q(\text{actual}) = C_d A\sqrt{(2gh)}$

where C_d is the coefficient of discharge. Its value depends on the profile of the orifice. For a sharp-edged orifice, as shown in Fig. 12.12b, it is about 0.62.

Figure 6. Derivation of Torricelli's theorem in Norman et al. (1990)

Thus the theoretically derived discharge of the system differs substantially from the actual discharge and is corrected by a coefficient of discharge, C_d. This is mainly due to an additional phenomenon which occurs in any orifice system: the *vena contracta*. Although this phenomenon is known for more than three centuries (cf. Torricelli 1644), no rigorous derivation exists for it and it is taken care of by a correction factor. Note that the correction factor is not an adjustment of a few percent, but of almost 40%. The value of the factor varies however with the profile of the orifice and can range from 0.5 (the so-called Borda mouthpiece) to 0.97 (a rounded orifice).

Introductory engineering textbooks tell us that coefficients of discharge are experimentally derived corrections that need to be established for each orifice separately (see Norman et al. 1990; Douglas and Matthews 1996). While this is true for real-world three-dimensional orifices, it must be stressed that there are analytical solutions for idealized two-dimensional orifice models by using free-streamline theory (see Batchelor 1967, p. 497). Moreover, Sadri and Floryan (2002) have shown that the *vena contracta* can also be simulated by a numerical solution of the general Navier-Stokes equations which is, however, *again* based on a two-dimensional model. For three-dimensional orifice models there are no analytical or numerical solutions (Munson 2002; Graebel 2002). The coefficients of discharge are then derived by physical modeling, i.e. by experiment. This explains perhaps why *physics* textbooks usually neglect the *vena contracta* in dealing with Torricelli's theorem. And some physics textbooks don't deal with Torricelli's theorem at all. To the best of my knowledge, all engineering textbooks that cover Torricelli's theorem also deal with the coefficient of discharge. (One may claim that the *vena contracta* can still be qualitatively explained: because the liquid converges on the orifice, the area of the issuing jet is less than the area of the orifice. But there exists no quantitative explanation of C_d for a three-dimensional jet.)

Although no analytical or numerical derivations exist for real-world orifice systems, engineering textbooks still link such systems via experimentally derived corrections to the theoretical law of Bernoulli, as if there were some deductive scheme. Why do they do that? One reason for enforcing such a link is that theory does explain some important features of orifice systems: the derivation in figure 6, albeit not fully rigorous, explains why the discharge of the system is proportional to the square-root of the height h of the tank, and it also generalizes over different heights h and orifice areas *A*. Another reason for enforcing a link to higher-level laws is that the resulting derivation can be used as an exemplar for solving new problems and systems. To show this, I will first turn the derivation in figure 6 into its corresponding derivation tree. But how can we create such a derivation tree if the coefficient of discharge is not derived from any higher-level equation? The orifice system indicates that there can be phenomenological models that are not derived from the theoretical model of the system. Yet, when we write the coefficient of discharge as the empirical generalization $Q(actual) = C_d Q(theoretical)$, which is in fact implicit in the derivation in figure 6, we can again create a derivation tree and "save" the phenomenon. This is shown in figure 7 (where we also added the principle of conservation of energy).

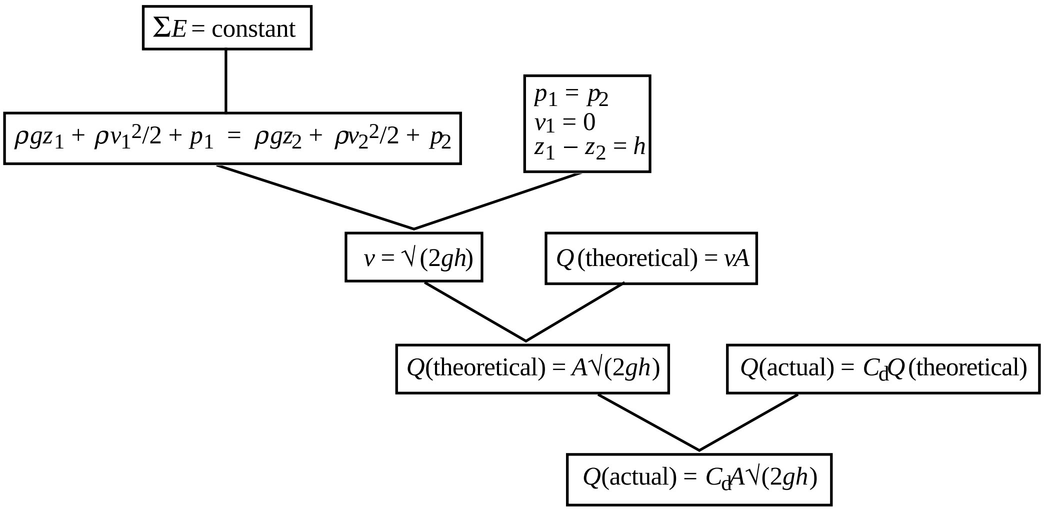

Figure 7. Derivation tree for the derivation in figure 6

The tree in figure 7 closely follows the derivation given in figure 6, where the initial conditions for p_1, p_2, v_1, z_1 and z_2 are represented by a separate label in the tree. The coefficient of discharge is introduced in the tree by the equation $Q(actual) = C_d Q(theoretical)$. Although this equation does not follow from any higher-level law or principle, we can use it *as if* it were a law. Of course it is not a law in the universal sense; it is a correction, a rule of thumb, but it can be reused for a range of other hydraulic systems, ranging from nozzles, notches, weirs, open channels and many pipeline problems -- see Douglas and Matthews (1996).

Does the derivation tree in figure 7 again represent a deductive-nomological (D-N) explanation? Different from the derivation trees in section 2, the final result Q(actual) = $C_d A\sqrt{(2gh)}$ in figure 7 is not deduced from general laws and antecedent conditions only. Additional knowledge in the form of a correction is needed to enforce a link. While this correction can be expressed in terms of a mathematical equation, it clearly goes beyond the notion of fundamental law or antecedent condition that are said to be essential to a D-N explanation (see Hempel 1965, p. 337).[3] Of course, the correction can be viewed as an *auxiliary hypothesis*, but it should be kept in mind that it is not derived from any higher-level law.

It is also difficult to frame the derivation tree in figure 7 into the semantic notion of theoretical model, since the formula Q(actual) = $C_d A\sqrt{(2gh)}$ is not true in the theoretical model of the system, except if C_d were equal to 1, which never occurs. However, the derivation tree does seem to concur in the notion of a *partial* model (or *partial* structure) since such a model allows for accommodating only a partial satisfaction of (some of) the laws in the theory (see da Costa and French 2003, p. 60). We can imagine a hierarchy of models: a theoretical model, a phenomenological model and a data model that are connected in terms of partial isomorphisms. The derivation tree in figure 7 implies two of such models: a theoretical model of discharge and a phenomenological model of discharge which are connected by C_d.

A new EBE model

What does this all mean for EBE? By using the derivation tree in figure 7 as an exemplar and by using the same substitution mechanism for combining subtrees from exemplars as in section 2, together with a mathematical procedure that can solve an equation, we obtain an exemplar-based model for fluid mechanics that can explain a range of new real-world systems. For example, the three subtrees in figure 8 can be extracted from the derivation tree in figure 7 and be reused in deriving the rate of flow of a rectangular *weir* of width B and height h (see e.g. Norman et al. 1990, p. 498).

---

3 Note that the correction factor does neither fit the notion of what Hempel called a *proviso* (Hempel 1988). A proviso would consist of the condition that there are no additional phenomena and thus *no vena contracta*. Under this proviso, the derivation of the discharge would need no correction factor, but it would be far from the truth.

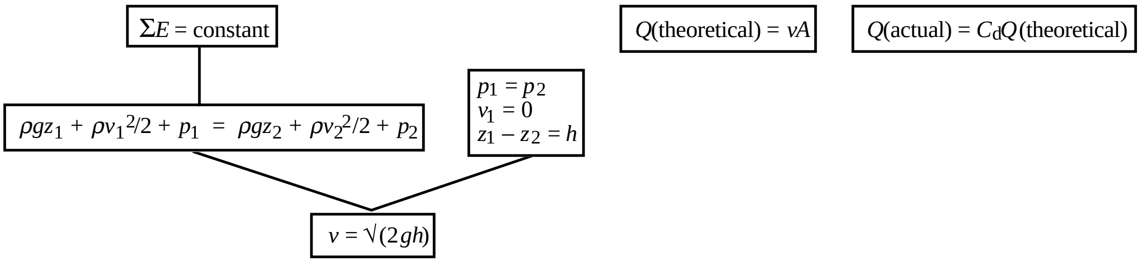

Figure 8. Three subtrees from figure 7 that can be reused to derive a weir

By adding the mathematical equivalence $vA = \int v \mathrm{d}A$ and the equation $\mathrm{d}A = B\mathrm{d}h$, which follows from the definition of a rectangular weir, we can create the derivation tree in figure 9 for the discharge of a weir.

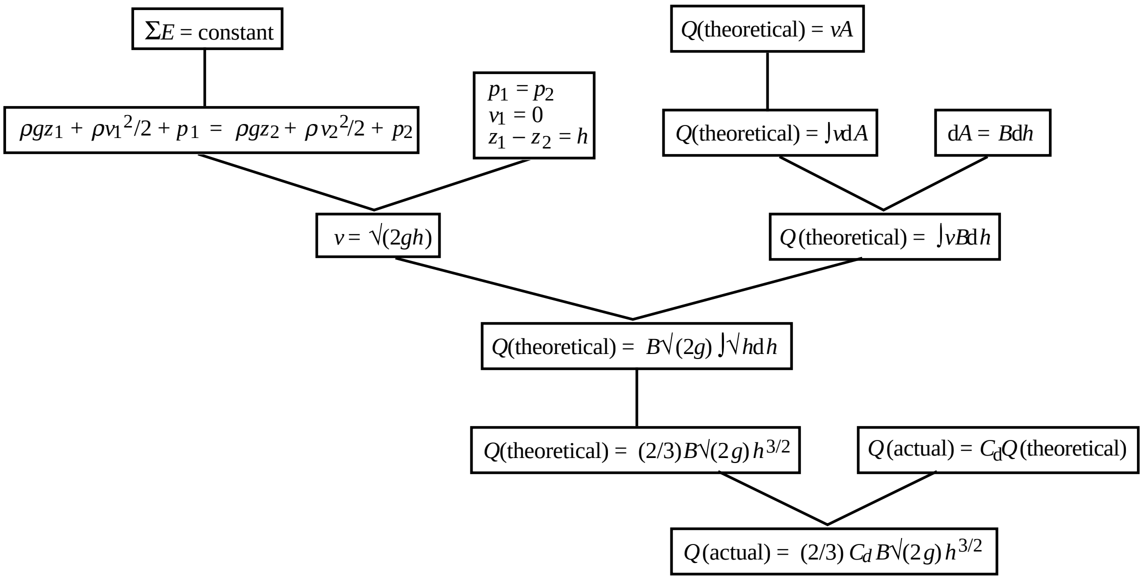

Figure 9. Derivation tree for a weir constructed by combining the subtrees from figure 8

The derivation tree in figure 9 closely follows the derivations given in Norman et al. (1990, p. 498) and Douglas and Matthews (1996, p. 117), where a weir system is constructed out of an orifice system. This corresponds to engineering practice where new systems are almost literally built upon or constructed out of similar previous systems. To give an historical example, the earliest known derivation of the weir system by Jean-Baptiste Bélanger (1828, p. 37) takes the orifice system as given and reuses it to derive the formula for the weir system in figure 9. That is, rather than deriving the weir system from scratch (i.e. from the principle of energy conservation), Bélanger reused not only Bernoulli's derivation of Torricelli's theorem (as reflected by the leftmost subtree in figure 8) but also the empirical coefficient of discharge (the rightmost subtree in figure 8) and the equation for discharge itself (the intermediate subtree in figure 8). Bélanger thus modeled the weir system by reusing and extending (parts of) a previously derived system in such a way that only minimal recourse to additional derivational steps is needed. EBE simulates this exemplar-based modeling by combining those derivational chunks that maximize derivational similarity or, equivalently, minimize derivation length. Note that the three subtrees in figure 8 indeed correspond to the smallest number of subtrees from the orifice system that are needed to construct, via some intermediate mathematical steps, a derivation for the weir system.

The resulting derivation in figure 9 effectively becomes an exemplar itself (which in EBE means that it is added to the corpus) and is reused and extended to derive a so-called *V-notch*. The V-notch, in turn, is extended to derive a so-called *trapezoidal notch*, which is again further extended to derive a *Cipolletti weir*, etc. (see e.g. Douglas and Matthews 1996; Chanson 2002). Modeling in engineering is thus highly cumulative: new systems are built upon or constructed out of previous systems and their derivations form increasingly complex wholes. We can handle this complexity by taking large(st) partial derivations from previous systems as "given" (as in EBE) and work from there, rather than deriving a system all the way down from laws.

The example of building new systems out of largest possible derivational chunks from previous systems is not only interesting for simulating engineering practice. I urge that we actually *need* exemplar-based knowledge to derive new real-world systems. Just compare the derivation of the weir in figure 9 with figure 5 in section 2 where we also constructed a new derivation tree by combining subtrees from a previous derivation tree -- i.e. a satellite's velocity from a derivation of the Earth's mass. But there is a very important difference: while the phenomenon represented in figure 5 can just as well be derived from theoretical laws rather than from previous subtrees, this is *not* the case for the phenomenon represented in figure 9. For deriving the weir system, we need to make recourse to the empirical rule $Q(actual) = C_d Q(theoretical)$ which is taken from a previously explained phenomenon that functions as an exemplar.

This use of theory-external knowledge in engineering modeling is also discussed by Margaret Morrison and Mary Morgan who call it "additional 'outside' elements" (Morrison and Morgan 1999, p. 11). They argue that models are autonomous entities which mediate between theory and the world. EBE is consonant with this view, in that previous models or exemplars can be reused almost independently of theory for deriving new systems. But in addition, EBE provides an exact computational model for dealing with "additional 'outside' elements". EBE is also in agreement with Boumans (1999) who argues that models are used as "recipes" for constructing other models. According to Boumans, models in economics integrate a broad range of ingredients such as analogies, metaphors, theoretical notions, mathematical concepts and techniques, stylized facts, empirical data and even policy views. In the next section, I will discuss models from language technology that also include various kinds of other elements such as idiosyncratic and idiomatic expressions. EBE is also congenial to the accounts by Nancy Cartwright and Ronald Giere. Cartwright gives a wealth of examples for her claim that "approximations and adjustments are required whenever theory treats reality" (Cartwright 1983, p. 13). In similar vein, Giere states that "strictly speaking, most purported laws of nature seem clearly to be false" (Giere 1999, p. 90). But while all these accounts stress the necessity of using theory-external elements in modeling, none of the accounts propose a computational mechanism which describes how theory-internal and theory-external knowledge can be integrated and (re)used for deriving new systems.

The surplus value of EBE, as I see it, is that (1) it proposes a computational approach to modeling in engineering, explicating which parts of previous systems and models can be used where, and (2) it integrates both theory-internal and theory-external knowledge; that is, as long as empirical rules, corrections, approximations, normalizations and the like can be stated in terms of mathematical equations they can be integrated by a derivation tree, and be reused to solve new problems by maximizing derivational similarity. EBE bypasses the notorious problem of defining "similarity of phenomena" by using a more precise (though perhaps more limited) notion of "similarity of derivations" of phenomena. EBE may therefore also suggest a formal alternative to case-based reasoning where previous cases are used to explain other, "similar" cases (cf. VanLehn 1998).

In doing all this, we need to slightly extend our previous definition of EBE given in section 2. There we stated that the leaf nodes of a derivation tree should refer to either general laws or antecedent conditions. In the new

EBE model, the leaf nodes may also be empirical rules -- or any other equations that are not deduced from higher-level laws. We may lump these three kinds of knowledge (laws, antecedent conditions and empirical rules) together as "knowledge that is not derived from higher-level knowledge". The definition of derivational similarity remains the same. We conjecture that scientists try to derive a new phenomenon by maximizing derivational similarity with previously derived phenomena, i.e. by using the largest partial matches from previous derivation trees, such that minimal recourse to additional derivation steps is needed.

The final formula in figure 9 is widely used in hydraulic engineering, where the coefficient C_d is often established experimentally. Yet it should be stressed that C_d is not a meaningless fudge factor. Instead, C_d has been defined in terms of other meaningful variables for various types of systems. For example, for the class of rectangular weirs there exists an empirical generalization which computes C_d from two other quantities. This generalization was first formulated by Henry Bazin, the assistant of the celebrated hydaulician Henry Darcy (Darcy and Bazin 1865), and is commonly referred to as Bazin formula (also called "Bazin weir formula", to distinguish it from "Bazin open channel formula" -- see Douglas and Matthews 1996, p. 119):

$$C_d = (0.607 + 0.00451/H) \cdot (1 + 0.55(H/(P + H)^2))$$

In this formula H = head over sill in metres, and P = height of sill above floor in metres of the weir. Bazin formula is an empirical regularity derived from a number of concrete weir systems. Although the regularity is known for more than 150 years, there exists no derivation from higher-level laws. Yet this does not prevent us from using and reusing the regularity in designing real world systems that have to work accurately and reliably, and it is easy to see that the formula can be integrated in the derivation tree of figure 9. Hydraulics is replete with formulas like Bazin's, each describing particular regularities within a certain flow system. There are, for example, Francis formula, Rehbock formula, Kutter formula, Manning formula, Chezy formula, Darcy formula, Keulegan formula, to name a few (see Chanson 2002 for an overview). Many of these formulas are known for more than a century but none of them has been deduced from higher-level laws. They are entirely based on previous systems and form the lubricant that makes new systems work.

In passing it is noteworthy that we cannot derive the phenomena from classical mechanics in section 2 by means of the derivations from fluid mechanics given in this section. For instance, Kepler's third law cannot be derived by subtrees from the orifice system. While this may seem trivial, there are many classical phenomena that *can* be derived by subtrees from the derivations given in this section. An example is the (idealized) velocity that an object attains in free fall from a height h in Newtonian mechanics, which is $v = \sqrt{(2gh)}$. This is "equal" to Torricelli's theorem, which gives the (idealized) velocity of the jet from a tank of height h, $v = \sqrt{(2gh)}$. So what happens if we use EBE to construct a derivation for a phenomenon which is merely described by $v = \sqrt{(2gh)}$ on the basis of a corpus which contains both the Bernoullian derivation of Torricelli's theorem and the Newtonian derivation of the velocity of a falling object? Then EBE obtains two different derivations for this phenomenon: one derived from Bernoulli's law and one from Newton's laws. Since the derivations are both maximally similar to a derivation in the corpus, which of the two should be chosen? If no distinction is made between the velocity of a fluid and that of a point mass, $v = \sqrt{(2gh)}$ is inherently ambiguous (or semantically undetermined) and two different models and derivations apply to it. This is not as problematic as it seems, since historically Daniel Bernoulli solved the problem of the velocity of water from an orifice by analogically treating a flow in terms of Newtonian-like particles, which makes the two phenomena indeed "equivalent". But if we want to avoid EBE mixing up derivations from different fields, we should introduce different variables for point mass velocity and fluid velocity. This can be accomplished by using subcategorizations, e.g. v_p for the velocity of a particle and v_f for a fluid. This way, the two velocities cannot be substituted, and the phenomena $v_p = \sqrt{(2gh)}$ and $v_f = \sqrt{(2gh)}$ get each a different derivation. But Bernoulli's historical example suggests that mixing up terms may also be illuminating, opening the door to analogical modeling.

4. EBE in other disciplines: an excursion into language technology

What counts for hydraulics also counts for many other technological disciplines: real-world systems and phenomena are derived not from general laws, but from parts of derivations of previous systems and phenomena. As an example from the other end of the technological spectrum, I will give a brief excursion into language technology.

While language *theory* is permeated by the idea that a language is aptly described by a formal grammar, i.e. a finite and succinct set of rules which can derive an infinite set of well-formed utterances, language *technology* does not work that way. As soon as a natural language processing system, such as a dialog understanding system or a machine translation system, needs to deal with a non-trivial fragment of a language, say English, formal grammars turn out to be severely inadequate. Grammars either *undergenerate*, which means that they provide no derivation for otherwise well-formed utterances, or they *overgenerate*, which means that they provide too many derivations for well-formed utterances (cf. Manning and Schütze 1999). "All grammars leak", is the well-known dictum of Edward Sapir (Sapir 1921, p. 38). In fact, there are so many idiosyncratic and idiomatic phenomena in natural language that only an approach which takes into account previously produced sentences can accurately model a language. After unsuccessful attempts to apply formal grammars to automatic linguistic analysis, a different paradigm has been developed since the 1980s in language technology: new sentences are derived not by using a concise set of rules, but by using a large corpus of previously derived sentences together with a "learning procedure" (see Manning and Schütze 1999 for an historical overview).[4]

Before going into the details of this learning procedure, let me first explain what derivations of sentences look like. It is by now widely acknowledged that sentence derivations can be represented by tree structures, similar to derivation trees in physics in the previous sections. The first linguistic tree structure was (most likely) proposed by Wilhelm Wundt in his *Logik* (Wundt 1880). But it was Noam Chomsky who made the notion of *syntactic phrase-structure tree* more widely accepted (Chomsky 1957). Although richer structures have also been proposed in the meantime (ranging from feature structures to attribute-value matrices), there is ample agreement that tree structures form the backbone of sentence analysis, sometimes enriched with phonological, morphological and semantic representations (see Sag, Wasow and Bender 2003; Bresnan 2000;

[4] Even if the notion of "grammar" is still used by many systems it is not succinct but consists of (tens of) thousands of rules that are derived from actual language corpora (see e.g. Knuuttila and Voutilainen 2002).

Goldberg 1995). In this section, I will focus on syntactic representations only.

So what does a syntactic phrase-structure tree look like? Figure 10 gives two tree structures for respectively the sentences *She wanted the dress on the rack* and *She saw the dog with the telescope*.

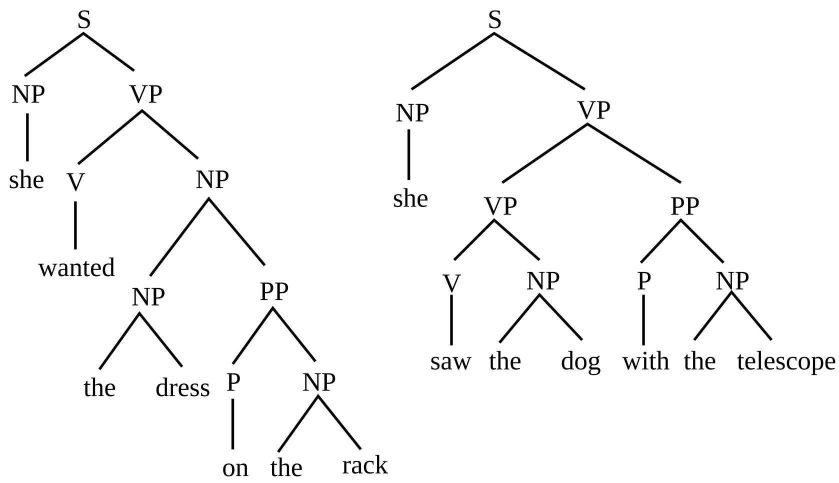

Figure 10. Two sentences with their phrase-structure trees

A phrase-structure tree describes how parts of a sentence combine into constituents and how these constituents combine into a representation for the whole sentence. The constituents in a phrase-structure tree are labeled with syntactic categories such as *NP* for noun phrase, *PP* for prepositional phrase, *VP* for verb phrase and *S* for the whole sentence. To keep the example simple, we have left out some low-level labels for Noun and Article. The two trees in figure 10 are structurally different in that in the first sentence the prepositional phrase *on the rack* forms a noun phrase with *the dress*, whereas in the second sentence the prepositional phrase *with the telescope* forms a verb phrase with *saw the dog*. Both sentences are "structurally ambiguous", to which I will come back below.

Although phrase-structure trees are not labeled with equations, they are compositionally built up as in physics derivation trees: each category is defined in terms of its underlying subcategories (and if we enrich each syntactic label with its logical-semantic interpretation, we would again obtain derivation trees with equations). Note that phrase-structure trees are represented upside down: the root is at the top instead of at the bottom. This is pure convention.

How can these sentences be used to derive new sentences, i.e. what does a "learning procedure" look like? There is not one way to do this. One straightforward but not very succesful method is to read off the "grammar rules" that are implicit in the trees, such as S => NP VP, VP => V NP, NP => NP PP, NP => *the dress*, etc. in figure 10 (Charniak 1996).[5] Another, more successful method is by reading off for every single word a subtree including that word (Chiang 2000). Yet another and still more successful method is by first enriching each syntactic label with its so-called "headword" and by next reading off the rules from the trees (Collins 1997). We know the relative successfulness of these methods as they have been evaluated on the same benchmark, the so-called Penn Treebank corpus consisting of 50,000+ sentences (Marcus et al. 1993). We will not go into further details of these different methods (but see Bod 1998 or Bod et al. 2003).

While these methods may seem rather disparate, they are based on the same underlying idea: new sentences are derived by parts of previously derived sentences. The distinctive feature of each method is their definition of what are to be considered the *relevant* parts. Yet it is also possible to generalize over these different methods by taking *all* partial trees as "relevant" parts. This general model is known as *Data-Oriented Parsing* or *DOP* (Bod 1998). By putting restrictions on the parts, other models and methods can be instantiated (see Charniak 1997).

The following example illustrates how the general DOP model works. If we take the sentences in figure 10 as our (unrealistically small) corpus, we can derive the new sentence *She saw the dress with the telescope* by extracting two subtrees from the trees in figure 10 and by combining them by means of *label substitution*:

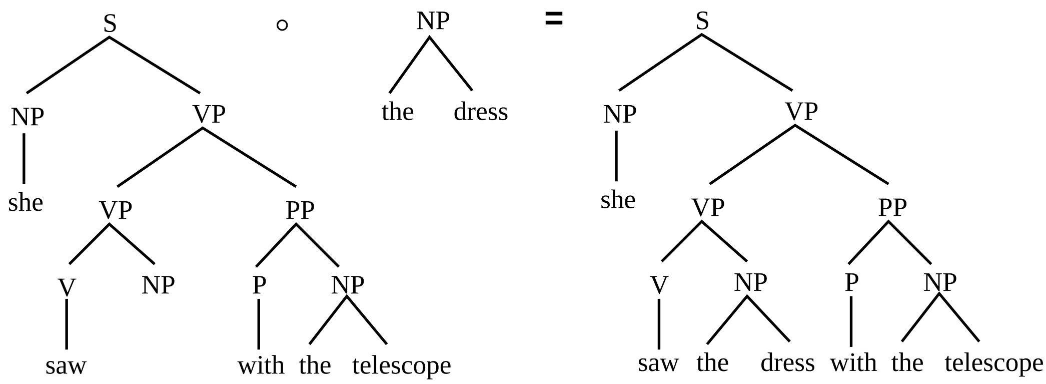

Figure 11. Deriving a new sentence by combining subtrees from figure 10

[5] A grammar rule like S => NP VP says that a sentence (S) consists of a noun phrase (NP) followed by a verb phrase (VP).

Note the similarity of label substitution with term substitution in the previous sections. Like term substitution, label substitution is a partial function on pairs of labeled trees and its range is the set of labeled trees. The label substitution of tree *t* and tree *u*, written as *t* ∘ *u*, is defined iff the root node of *u* is equal to the leftmost syntactic leaf node of *t* (which in figure 11 is an *NP*). If label substitution is defined, it yields a tree where *u* is substituted in the leftmost syntactic leaf node of *t*. As in EBE, the underlying idea of DOP is that new trees are constructed by combining partial trees from a prior corpus. It is easy to see that we can create an EBE model for DOP by properly instantiating the two parameters given in section 2.

In figure 11, the new sentence *She saw the dress with the telescope* is interpreted analogous to the corpus sentence *She saw the dog with the telescope*: both sentences receive roughly the same phrase structure. Yet we can also derive an alternative phrase structure for this new sentence, namely by combining three (rather than two) subtrees from figure 10, as shown in figure 12.

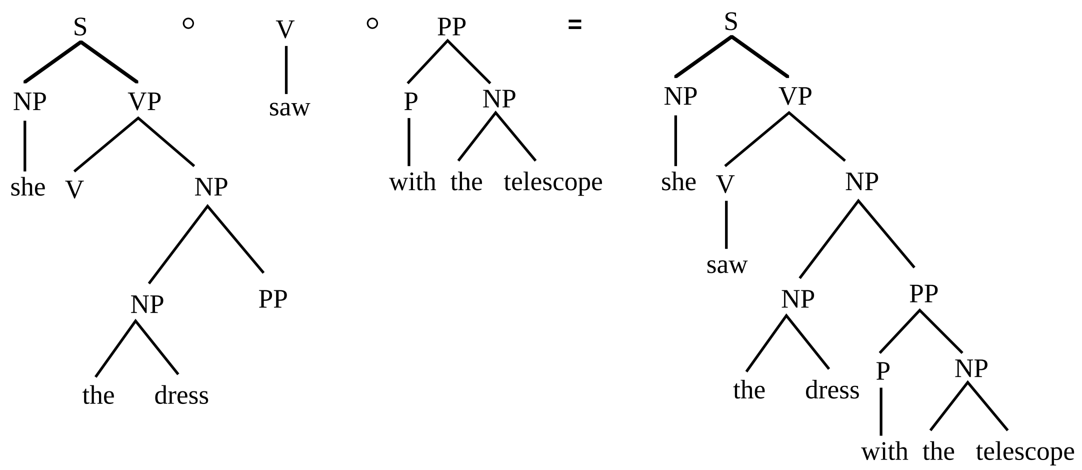

Figure 12. A different derivation for *She saw the dress with the telescope*

Thus the sentence *She saw the dress with the telescope* can be derived in (at least) two different ways: either analogous to the first tree in figure 10 or analogous to the second tree in figure 10. Which one should be chosen? As in EBE, it is hypothesized that humans derive a new sentence by maximizing derivational similarity -- or equivalently, minimizing derivation length -- with respect to previously derived sentences.[6] For our example sentence *She saw the dress with the telescope*, the shortest derivation (which maximizes derivational similarity) is represented by figure 11: only two subtrees from the corpus are needed to construct this tree, while at least three corpus-subtrees are needed to construct the tree in figure 12.

As before, the notion of derivational similarity favors new trees that are most similar to previous trees. Of course, the corpus in figure 10 is far too small for simulating actual language processing. More realistic experiments use corpora of hundreds of thousands of (manually) constructed phrase-structure trees. By using largest possible derivational chunks from such corpora, we can also take into account arbitrary multi-word expressions or idiom chunks such as *to take advantage of* and fixed phrases such as *What time is it?* (note that one does not say *How late is it?* in English).

While DOP generalizes over a large number of models in natural language technology, EBE is even more general: it allows in principle for corpora of any sort of trees -- be they physical, linguistic, musical or of any other kind. As long as we can construct a corpus of exemplary derivations for a certain discipline, we can create an EBE model for it and use it to derive new phenomena without making recourse to an axiomatic system of rules. Sure enough, the exemplary derivations in the corpus do include general laws or rules, such as $F = m a$ in physics or $S \Rightarrow NP\ VP$ in linguistics, but they also include very particularist information ranging from empirical coefficients in hydraulics to idiomatic expressions in natural language that do not follow from these laws or rules. The world may be full of nomothetic elements like laws. But it is also full of idiographic, particularist elements such as ad hoc corrections. EBE does justice to both.

5. Conclusion

I have argued for a general model of "applying science", termed EBE. This model explains new systems and phenomena by recombining fragments or chunks from previously derived systems and phenomena. Examples from hydraulics and language technology suggest that EBE can operate with any kind of corpus as long as we have a precise notion of "derivation". This results in the following general methodology for applying science: (1) construct a prior corpus of derivations of exemplary phenomena, and (2)

[6] Most models also take into account the frequency of occurrence of derivational chunks in the corpus (see Manning and Schütze 1999), but I will not go into this here.

combine as large derivational chunks as possible from the corpus in deriving new phenomena. Newly constructed derivations are added to the corpus, and may be reused as exemplars themselves. I contend that science and technology should be understood not in terms of a "minimalist" system of laws or rules, but in terms of a "maximalist", dynamically updated ensemble of derivations.

References

Alonso, Marcelo and Edward Finn (1996), *Physics*, Addison-Wesley.

Ankeny, Rachel (2003), "Cases as explanations: modeling in the biomedical and human sciences", *Volume of abstracts, 12th International Congress of Logic, Methodology and Philosophy of Science*.

Baader, Franz and Tobias Nipkow (1998), *Term Rewriting and All That*, Cambridge University Press.

Batchelor, G. (1967), *An Introduction to Fluid Dynamics*, Cambridge University Press.

Bélanger, Jean-Baptiste (1828), *Essai sur la Solution Numérique de quelques Problèmes Relatifs au Mouvement Permanent des Eaux Courantes*, Carilian-Goeury, Paris.

Bod, Rens (1998), *Beyond Grammar*, CSLI Publications.

Bod, Rens (1999). "Context-Sensitive Spoken Dialogue Processing with the DOP Model", *Journal of Natural Language Engineering* 5(4), 306-323.

Bod, Rens (2000), "Parsing with the shortest derivation", *Proceedings COLING'2000*, 69-76.

Bod, Rens (2004), "Exemplar-Based Explanation", *Proceedings E_CAP'2004*, Pavia.

Bod, Rens, Jennifer Hay and Stefanie Jannedy (2003), *Probabilistic Linguistics*, The MIT Press.

Boon, Mieke (2004), "Explaining Basic Sciences in the Engineering Sciences", *Proceedings PSA'2004 Workshop on "Applying Science"*, Austin, TX.

Boumans, Marcel (1999), "Built-in justification", in M. Morgan and M. Morrison (eds.), *Models as Mediators*, Cambridge University Press, 66-96.

Bresnan, Joan (2000), *Lexical-Functional Syntax*, Blackwell.

Bunge, Mario (1966), "Technology as Applied Science", *Technology and Culture*, 7, 329-347.

Cartwright, Nancy (1983), *How the Laws of Physics Lie*, Oxford University Press.

Cartwright, Nancy (1999), *The Dappled World*, Cambridge University Press.

Chanson, Hubert (2002), *The Hydraulics of Open Channel Flow*, Butterworth-Heinemann.

Chiang, David, (2000), "Statistical parsing with an automatically extracted tree adjoining grammar", *Proceedings ACL'2000*, Hong Kong, China.

Charniak, Eugene (1996), "Tree-bank Grammars", *Proceedings AAAI'96*, Menlo Park, Ca.

Charniak, Eugene (1997), "Statistical Techniques for Natural Language Parsing", *AI Magazine*, Winter.

Chomsky, Noam (1957), *Syntactic Structures*, Mouton, The Hague.

Collins, Michael (1997), "Three Generative Lexicalised Models for Statistical Parsing", *Proceedings ACL'97*, Madrid, Spain.

Cormen, Thomas, Charles Leiserson and Ronald Rivest (2002), *Introduction to Algorithms*, The MIT Press.
Da Costa, Newton and Steven French (2003). *Science and Partial Truth - A Unitary Approach to Models and Scientific Reasoning*, Oxford University Press.
Darcy, Henry and Henry Bazin (1865), *Recherches Hydrauliques*, Imprimerie Nationale, Paris.
Douglas, J. and R. Matthews (1996), *Fluid Mechanics*, Vol. 1, 3rd edition, Longman.
Giere, Ronald (1988), *Explaining Science: A Cognitive Approach*, University of Chicago Press.
Giere, Ronald (1999), *Science without Laws*, The University of Chicago Press.
Goldberg, Adele (1995), *Constructions*, The University of Chicago Press.
Graebel, W. (2001), *Engineering Fluid Mechanics*, Taylor & Francis.
Hempel, Carl and Paul Oppenheim (1948), "Studies in the Logic of Explanation", *Philosophy of Science*, 15, 135-175.
Hempel, Carl (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, The Free Press, New York.
Hempel, Carl (1988), "Provisoes: a problem concerning the inferential function of scientific theories", *Erkenntnis*, 28, 147-164.
Kitcher, Philip (1989), "Explanatory unification and the causal structure of the world", in Kitcher, Philip, and Salmon, Wesley (eds.), *Scientific Explanation*, University of Minnesota Press, 410-505.
Knuuttila, Tarja and Atro Voutilainen (2002), "A parser as an epistemic artefact: a material view on models", *Proceedings PSA 2002*.
Kuhn, Thomas (1970), *The Structure of Scientific Revolutions*, 2nd edition, University of Chicago Press.
Manning, Chris and Hinrich Schütze (1999), *Foundations of Statistical Natural Language Processing*, The MIT Press.
Marcus, M., B. Santorini and M. Marcinkiewicz, (1993), "Building a Large Annotated Corpus of English: the Penn Treebank", *Computational Linguistics*, 19, 313-330.
Mikhailov, Gleb (2002), *Die Werke von Daniel Bernoulli*, Band 5, Birkhäuser Verlag, Basel.
Morgan, Mary and Margaret Morrison (eds.) (1999), *Models as Mediators*, Cambridge University Press.
Morrison, Margaret (1999), "Models as autonomous agents", in M. Morgan and M. Morrison (eds.), *Models as Mediators*, Cambridge University Press, 38-65.
Morrison and Morgan (1999), "Models as mediating instruments", in M. Morgan and M. Morrison (eds.), *Models as Mediators*, Cambridge University Press, 10-37.
Munson, Bruce (2002). *Fundamentals of Fluid Mechanics*, Wiley.
Nickles, Thomas (2003), "Normal science: from logic to case-based and model-based reasoning", in Thomas Nickles (ed.), *Thomas Kuhn*, Cambridge University Press, 142-177.
Norman, Eddie, Joyce Riley and Mike Whittaker (1990), *Advanced Design and Technology*, Longman.
Sadri, R. and Floryan, J. M. (2002), "Entry flow in a channel", *Computers and Fluids*, 31, 133-157.

Sag, Ivan, Thomas Wasow and Emily Bender (2003), *Syntactic Theory: A Formal Introduction*, CSLI Publications.

Sapir, Edward (1921), *Language*, Hartcourt Brace & Company.

Suppes, Patrick (1961), "A comparison of the meaning and use of models in the mathematical and empirical sciences", in H. Freudenthal (ed.), *The Concept and Role of the Model in Mathematics and Natural and Social Sciences*, Reidel, 163-177.

Suppes, Patrick (1967), "What is a scientific theory?", in S. Morgenbesser (ed.), *Philosophy of Science Today*, New York: Basic Books, 55-67.

Torricelli, Evangelista (1644), *De Motu Gravium Naturaliter Descendentium et Proiectorum Libri duo*, Florentiae.

van Fraassen, Bas (1980), *The Scientific Image*, Oxford University Press.

VanLehn, Kurt (1998), "Analogy Events: How Examples are Used During Problem Solving", *Cognitive Science*, 22(3), 347-388.

Wundt, Wilhelm (1880), *Logik. Eine Untersuchung der Prinzipien der Erkenntnis und der Methoden Wissenschaftlicher Forschung*, Enke, Stuttgart.