Engineering of Dynamic Knowledge in Exact Sciences: First Results of Application of the Event Bush Method in Physics

Cyril A. Pshenichny¹, Roberto Carniel² and Paolo Diviacco³

¹ Geognosis Project, Intellectual Systems Laboratory, National Research University of Information Technologies, Mechanics and Optics, Kronverksky Prospect, 49, St. Petersburg 197101, Russia

² Laboratorio di misure e trattamento dei segnali, DICA, Università di Udine, Via delle Scienze, 206, 33100 Udine, Friuli, Italia

³ Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, Borgo Grotta Gigante 42/c, 34010 Trieste, Italia

cpshenichny@yandex.ru

roberto.carniel@uniud.it

pdiviacco@ogs.trieste.it

Abstract. The tools of knowledge engineering are commonly applied in poorly formalized information domains. However, there is a number of reasons to use them also in exact sciences along with mathematics and mathematical logic. In physics, the knowledge is largely dynamic, i.e. describes rather processes and changing objects than objects with fixed properties and states. Therefore, even ontologies in physics tend to include a “dynamic” component. Nevertheless, application of specific methods of dynamic knowledge engineering looks very promising and beneficial for physics. The paper reports first results of application of the event bush method in theoretic and applied (geophysical) contexts.

Keywords. Physics, knowledge engineering, dynamic knowledge, event bush method, ideal gas, site effect

1 Introduction

The sense of application of knowledge engineering tools in exact sciences may look arguable to many researchers. Indeed, the virtue of knowledge engineering is to shape up and structurize the semi-intuitive fields of knowledge, extract axioms and suggest inference where these are unknown. Physics and chemistry, being still far from as
formal as, say, geometry, are nevertheless well expressed in terms of mathematics and seem to require only mathematical logic for further (and perhaps complete) formalization.

Still, there are a few reasons to think differently.

1. To make physical or chemical issues computer-understandable, one needs to convey meaning to information systems, while the latter requires knowledge-engineering technologies. As was shown by Borst et al. [3], computer may readily recognize the mathematical operation of multiplication in an expression $F=ma$ but it is a highly non-trivial task to make it understand this expression as a part of Newtonian mechanics.

2. It often becomes necessary or desired to explain the matters of physics or chemistry to a wider community, e.g. to make geophysical models clear to other geoscientists [3; 11]. A method enabling one to share knowledge with a wider community (perhaps even without sharing it with a computer) is a pre-requisite for creation of computer-based collaboration environments.

3. The accuracy of existing and new physical models sometimes needs to be checked not only for mathematical errors but also to analyze whether the variables have been used adequately to reflect personal or collective vision of the phenomenon. This may strongly help to create new models and reconcile standpoints of physicists or chemists. Such reconciliation may appear especially helpful for physics. In physics, people often think first in terms of things and events (processes) and then convert their vision into variables and mathematical operations, and this passage is often determined by intuition or by existing bias in a given scientific school. Many authors, as for example Polson and Curtis [10] or Bond et al. [2] highlighted the role of previous experiences and preconceived notions that stem from their personal backgrounds. Diviacco [6] analyzed the relationship of creative (abductive) reasoning and social positioning of researchers and scientific institutes. Baddeley et al. [1] reported on the phenomenon of opinion shaping and herding, while suggesting paths to interrogate experts through elicitation.

4. Computer-aided engineering may benefit from “parsing” the physical models used in it [15].

5. Perhaps, knowledge engineering would offer new formalisms, which, along with existing mathematical and logical approaches, would also suit these fields well and be useful to them. Thus, one may expect that the method of event bush [8; 10] evolves into such formalism with time.

These considerations have led to a number of applications of knowledge engineering methods in physical and chemical domains.

Like in other domains, two tendencies can be recognized here, static and dynamic knowledge modeling. By static we mean the representation of information domain as a no-change environment sensu Pshenichny and Kanzheleva [13]. Static knowledge engineering proceeds in exact sciences, like in other fields, mainly by means of ontology design. The same time, the character of knowledge in exact sciences, especially in the physics which will be the focus henceforth in this paper, is such that even the properties of objects (e.g., mass, momentum, charge or energy of a body, orientation and intensity of a field) are considered mainly in relation to possible response to some
external impact, i.e., dynamic view is implicitly wrapped even in formally “static” descriptions. This is why it is not surprising at all that the ontologies used in this domain often tend to incorporate a “dynamic component”. Though, the ontology proper offers not too much to capture processes, the solution found by Borst et al. [3] is to create an ontology of processes as chains of flows, define flows as changes of particular kind of stuff (mass, energy, location, charge and so forth), link this stuff to components of physical system, associate these components with variables and decompose them into sub-components, which, in turn, would correspond to various values of variables. Proceeding from one value of variable to another under specified conditions (the latter represent physical function defined by a model or law) is visualized by a bond graph [4]. Though only variables and variable states are in the nodes of the bond graph, if the latter is interpreted in terms of subcomponents between which some stuff changes, it should be regarded as an event-based tool. As such, the bond graph is the counterpart to the essentially static PhySys ontology.

An approach demonstrated by Yoshioka et al. [15] is rather similar to that of Borst et al. [3]. In the physical concept ontology they develop in support of computer-aided engineering they suggest the following “conceptual categories”. Two of them describe the “static structure of knowledge” (entity, representing an “atomic physical object”, and relation, meaning the “relationship among entities to denote a static structure”). Other three relate to “dynamic” part of physical knowledge: attribute (“a concept attached to an entity that takes a value to indicate the state of the entity”), physical phenomenon that “designates physical laws or rules that govern behaviors” and physical law “representing a simple relationship among attributes”. Importantly, the passage from static to dynamic knowledge in both studies, in fact, means passage from names (of subcomponents in one case or attributes in the other) to variables, from the qualitative to quantitative vision. Only variables and their correctly built combinations (i.e., laws and models represented by formulae) are believed to be able to represent the behavior of what is called physical system by the quoted authors.

This is a rather common point in physical modeling resulting in the fact that when the event-based methods are applied in physics (bond graphs, equation graphs, flowcharts sensu stricto or sensu lato), quite often only variables, equations or, at best, questions occur in their nodes and meaning of arcs may be not defined at all. Perhaps from the point of view of physicists such semantics is perfectly organized as it gets rid of virtually anything qualitative. Still, if not to consider quantitative thinking as some upper stage of evolution of brain and take it as a part of existing scientific and cultural context, the semantics of these graphs and corresponding domains of physical knowledge looks rather groundless and needs substantiation in words, though, managed as strictly as variables.

In physics do exist visions that already imply their description by an event-based knowledge engineering method – e.g., branching process suggesting application of event trees (directed invariant change environment), various cyclic processes calling for causal loops (non-directed change environment), and so forth, up to the most complicated, alternative directed change environment.
However, these methods, as was argued by Pshenichny and Mouromtsev [13], contrary to the object-based (“static”) ones, need better rethink and formalization of grammar of event description in nodes.

One method of dynamic knowledge engineering, the event bush, has got the most evolved verbal grammar and, besides, addresses the most complicated and all-inclusive modeling environment, that of alternative directed changes [13]. In our paper we will examine an opportunity of formalization of record of events, processes and scenarios in physics by means of the event bush method.

To do this, first we consider a trivial task from theoretical physics, and then review a geophysical application of the event bush.

2 Theoretical Task

The theoretical task we considered for application of the event bush represents one of the fundamental physical laws known as Mendeleev-Clapeyron Law, or ideal gas equation. Being used in thermodynamics, statistical mechanics and kinetic theory of gases, it links pressure, volume, temperature and amount of gas. Its most frequently introduced form is

\[ pV=nRT \]  

(1)

where \( p \) is the pressure of the gas, \( V \) is the volume of the gas, \( n \) is the amount of substance of gas (also known as number of moles), \( T \) is the temperature of the gas and \( R \) is the ideal, or universal, gas constant, equal to the product of Boltzmann's constant and Avogadro's constant.

To put the knowledge from equation (1) into an event-based framework one needs to understand how the reasoning in terms of parameters corresponds to reasoning in terms of events. In particular, this means to relate variables of the equation to some propositions, their subjects and predicates.

For this, we shall conceptualize this equation as a possible scenario in some environment of directed alternative changes, in which we first of all have to define, as implied by the event bush method, some “key players” being primary, non-unique inputs, which would determine any further course of events, and external “actors” that may put some constraints on the behavior of “key players” [14]. Terms defining “key players” become subjects of primary internal, and those defining external actors, the subjects of primary external events in the event bush.

In our case, it is reasonable to look for those “key players” and external “actors” whose properties can be expressed as variables in the considered equation. \( R \) should be excluded from consideration because it is a constant. What remains is pressure \( p \), volume \( V \), number of molecules (moles) \( n \) and temperature \( T \). If to take molecules of gas as “key players”, then their amount is obviously their property, temperature is a characteristic of their chaotic motion, and pressure is also their characteristic, of the overall impact they exert on something. Likely, exactly this “something” is an external “actor”, whose appearance gives sense to the pressure. And, obviously as well, its
property must be “to have volume” or, to put it more correct, to have finite volume, \( V \). However, as all the variables have been included into consideration, it will be good if neither we have more “key players”, nor other external “actors” are needed, nor any other properties (predicates) are needed for this only external “actor” we are trying to define. In fact, “to have volume” and “to be finite” is all we need from it. Then, whatever it be, we can call it just “volume” and attribute the predicate “to be finite” to it. However, “to be finite” may be axiomatically taken equivalent to “to have walls”. This predicate can be used to constrain the definition of “key players” before and after the action of the external “actor”.

Then, the structure of the event bush appears as follows: (ia) “Molecules of gas move chaotically in infinite volume and do not hit volume-limiting walls”; (ib) “Volume is limited”. The resulting bush is shown in Fig. 1, a. Importantly, when external “actor” comes to play, i.e., the volume in which molecules of gas reside becomes limited (no matter in what way), we cannot say that molecules of gas immediately exert pressure on the volume’s walls. No, this takes some, however short, time, though happens inevitably and without any additional influence. Semantically this is expressed in the change (ii) “Molecules of gas move chaotically in finite volume and do not hit volume-limiting walls” FLOW (ii) “Molecules of gas move chaotically in finite volume and hit volume-limiting walls”.

When the bush is ready, it looks more or less straightforward to relate variables to its events (see Fig. 1, b). Each event in event bush is a statement consisting of one subject and whatever number of predicates. If a subject or predicate are quantifiable (e.g., a subject “molecules of gas” may imply number of molecules, predicate “to move chaotically” implies temperature that characterizes this movement), this quantitative parameter may be readily added to the formulation of the event right after the corresponding subject or predicate – “molecules of gas in amount of \( n \) moles” or “move chaotically with temperature \( T \)”. We put the parameter in italics for clarity. Thus a qualitative parameter both formally and stylistically “stems from” the qualitative formulation of event. In some cases, moreover, event’s subject or predicate may be a parameter to itself, e.g., “volume”. Once attributed to a subject or predicate, a parameter is traced throughout the bush in all events in which this subject is present or this predicate is asserted. If the predicate is negated, no parameter is associated with it.

When parameters are associated with events, each connective of the bush is being attributed a computational sense (Fig. 1, c) serving as a step in construction of some formula (one of the formulae constructed by the bush represents the considered law). Either a connective adds one mathematical operation over the variables present in the nodes (events) it connects, or it presents an assumption relevant for the formula. Herewith, it becomes important again that semantic peculiarity in change between two (ii) events fixed by the flux connective, which was quoted one paragraph above. Now it is laden with an assumption that puts in explicit form the fact that we ignore the time elapsing after the volume becomes closed and before the pressure appears in it.
Following the two flows of the event bush, we arrive either to a “nothing happened” case or to the scenario that is described by the Mendeleev-Clapeyron law and see the relevant physical formula is being constructed.

**Diagram:**

- **Flow 1:**
  - Volume $V$ is limited
  - Molecules of gas move chaotically in limited volume and hit the volume-limiting walls
  - Molecules of gas move chaotically in limited volume and do not hit the volume-limiting walls

- **Flow 2:**
  - Volume $V$ is limited
  - Molecules of gas in amount of $n$ make movement chaotically with temperature $T$ in initial volume and do not hit the volume-limiting walls
  - Molecules of gas in amount of $n$ make movement chaotically with temperature $T$ in limited volume and do not hit the volume-limiting walls
  - Molecules of gas in amount of $n$ make movement chaotically with temperature $T$ in limited volume and hit with pressure $p$ the volume-limiting walls
  - Molecules of gas in amount of $n$ make movement chaotically with temperature $T$ in limited volume and hit with pressure $p$ the volume-limiting walls
Fig. 1. An event bush describing the environment, whose one scenario is modeled by the Mendeleev-Clapeyron law: (a) environment proper, (b) environment with variables attributed to events, (c) environment with variables, assumptions and mathematical operations. See how the formula of the law is built step by step in event bush.

3 Geophysical Application

The same approach of construction of “parsing” the physical formulae, or models, was used to explicate one seismological model of the so-called site effect.

Many earthquakes have indicated that the presence of deposits of soft soil over an underlying harder rock can increase dramatically and/or concentrate locally damages and life losses. Soft soils amplify shear waves and, thus, amplify ground shaking. This amplification of motion over soft sediments is called site effect and takes place mainly due to the trapping of seismic waves associated to the impedance contrast between the more superficial sediments and underlying bedrock.

The horizontal-to-vertical (H/V) spectral ratio approach, i.e. the study of the ratio between the amplitude spectra of the horizontal and the vertical component of seismic noise, was first introduced by Nogoshi and Igarashi [9] but became widely known only with the work of Nakamura [7]. The same Nakamura then clarified his method in a more recent paper [8].

Horizontal and vertical spectra on the surface ground of the sedimentary basin ($H_f$, $V_f$) can then be written as follows:

$$H_f = A_hH_b + H_s; \quad V_f = A_vV_b + V_s$$

(2)
where $A_H$ and $A_V$ are the spectral amplification factors of the horizontal and vertical motion of vertically incident body waves, correspondingly, while $H_b$ and $V_b$ are the spectra of horizontal and vertical motion as they would be acquired in the bedrock under the basin. $H_s$ and $V_s$ are the spectra of horizontal and vertical components respectively of surface waves traveling along the earth surface. Thus two types of seismic signals are involved in producing the final site effect, called by us for simplicity “interior spectrum” (i.e. a seismic signal with given – horizontal or vertical – spectrum, originating in the earth’s interior) and “exterior spectrum” (i.e. a seismic signal with given – horizontal or vertical – spectrum, coming to the surface from anywhere on or above it).

So the entire model of site effect on seismic spectra as suggested by Nogoshi and Igarashi [9] can be represented as $H_f/V_f=(H_b/V_b)*((A_H+H_s/H_b)/(A_V+V_s/V_b))$ – being is a purely mathematical derive from two equations, $H_f=A_HH_b+H_s$ and $V_f=A_VV_b+V_s$, which are actually the model. Exactly these two equations will be addressed by the event bush.

For more detail, the reader is referred to the paper by Carniel et al. [5].

To parse equation (2), an event bush was composed that describes how the propagating seismic spectra are transformed by heterogeneous geologic substrate up to the earth surface (Fig. 2) based on the qualitative understanding of this phenomenon.

Fig. 2. Event bush describing how the propagating seismic spectra are affected by heterogeneous substrate
As is evident from the title of the bush, its main “players” should be the interior and exterior couples of horizontal and vertical seismic spectra (initially placed on the left, i.e., being classified as primary internal events) and geologic substrate (bedrock and sedimentary bed) plus the Earth surface, which are put on top of the bush as primary external ones.

As could be concluded from the Nogoshi – Igarashi – Nakamura model, the interior couple is considered to come to the bedrock but is not generated in it, because the opposite would cause other, side effects that could appear influential on propagation of spectra and had to be accounted for. To avoid this unnecessary complication, as well as for the sake of semantic clarity, we consider interior couple of spectra per se as a primary internal event. Its full formulation is “Interior horizontal spectrum and vertical spectrum go unaffected by bedrock, unamplified by sedimentary bed and unaffected by Earth surface”. By “affected” we mean any kind of influence, i.e. a very general concept of (possibly nonlinear) filtering. The model says that the sedimentary bed amplifies spectra (and this is its influence on the latter, mathematically representable with a linear filter transfer function); however, we may not exclude that other two primary external “players”, bedrock and surface, also may affect (or not) somehow. So we decide that the predicate “to affect” is attributable to any subject including “Bedrock” and “Earth surface”, while the predicate “to amplify” is attributed to “Sedimentary bed” from the very beginning.

In principle, interior spectra may never meet the bedrock meant in the model (say, traveling yet deeper than, or far away from, this bedrock body all along). Exactly this is what the same-formulated tertiary event caused by this primary internal one is introduced for. Hence, the encounter of that bedrock by the spectra is another, secondary event “Interior horizontal spectrum and vertical spectrum go affected by bedrock, unamplified by sedimentary bed and unaffected by Earth surface”. Naturally, this secondary event is caused by the said primary internal and a primary external one, “Bedrock exists”.

From this point, other possible events involving interior spectra develop. These spectra can either travel in the considered area within the bedrock and not leave it (this is expressed by a tertiary statement formulated the same way as the above secondary one), or come to sedimentary bed, or come to the surface. Two latter options are secondary events that result from a combination of the secondary event “Interior horizontal spectrum and vertical spectrum go affected by bedrock, unamplified by sedimentary bed and unaffected by Earth surface” with a primary external, “Sedimentary bed exists and amplifies”, in one case, and “Earth surface exists”, in the other. Exactly these two secondary events represent the contrasting cases analyzed by Nakamura. Each of them can be clearly documented, i.e., represent some end results of propagation of the spectra, and this is reflected in the bush by corresponding tertiary events.

Now, another independent “actor” comes to play. This is a purely superficial, or atmospheric, or anthropogenic, event that may also generate vertical and horizontal seismic spectra, henceforth denoted as “exterior”. This is another primary internal event of the bush. In principle, it may pass unrelated to the interior spectra, just “meaning itself” and resulting in a tertiary event. Nevertheless, if a portion of it
comes to the Earth surface (i.e., a combination of this primary internal event with the primary external one “Earth surface exists” takes place), this results in a couple of superficial spectra spreading close to the ground and being somehow affected by the latter. This is expressed by the secondary event, “Exterior horizontal spectrum and vertical spectrum go affected by Earth surface”, and the end result of this, by the same-named tertiary event.

If two different couples of spectra, one coming from the interior, the other purely exterior, meet at the Earth surface, this naturally leads to two interior-exterior “couples from couples”, one interior-exterior horizontal and one interior-exterior vertical. These “couples from couples” will be denoted composite spectra. Their Interior components bear the history of previous transformations (“affected by bedrock, amplified or not by sedimentary bed and affected by Earth surface”), while exterior ones may only be affected by Earth surface. In the event bush, this is expressed as confluence of the event two events – one of these in both cases is in one case, “Exterior horizontal spectrum and vertical spectrum go affected by Earth surface”, and the other, depending on what kind of interior spectra are involved (or, in other words, where the seismometer is located), either “Interior horizontal spectrum and vertical spectrum go affected by bedrock, unamplified by sedimentary bed and affected by Earth surface”, or “Interior horizontal spectrum and vertical spectrum go affected by bedrock, amplified by sedimentary bed and affected by Earth surface”. Both events describing the composite spectra lead to tertiary events, which document the two scenarios captured by the Nogoshi – Igarashi – Nakamura model.

Some important notes can be made on the overall structure of the event bush.

Its primary internal events include two different types of spectra couples, interior and exterior (see above). In the model examined, no other seismic signals are considered.

Primary external events include the members of the geological sequence (bedrock and sedimentary bed) and the earth surface. Again, in the considered case this is surely the full set of opportunities.

Secondary events (ii) fall into four classes:
1. those formed by combination of ia and ib;
2. those formed by combination of the 1st class type (ii) events and one event ib (namely, “Earth surface exists”);
3. those formed by confluence of the 2nd class type (ii) events and one of the 1st class type (ii) events (namely, “Exterior horizontal spectrum and vertical spectrum go affected by Earth surface”);
4. those formed by simple cause-effect relation from the 3rd class type (ii) events.

Tertiary events, as supposed by the rules of event bush composition, were generated by primary internal ones and secondary events except those of the 3rd class. Their formulation repeats that of the events they originate from.

We do not ascertain that this bush is the only possible one describing the transformation of seismic signals in heterogeneous geological environment. It would be interesting to try to build other bushes in a different semantics and look at their interrelation. Moreover, a “vice versa” bush can be created describing the way the geolog-
ic bodies (bedrock, sedimentary bed and, finally, the earth surface) are being affected
by seismic spectra and, in turn, transform the latter.

Like in the case of Mendeleev-Clapeyron equation, the succession of steps for
building the bush that fits physical model is the same: building the bush proper, at-
tributing variables to the events, attributing mathematical operations or assumptions
to the event bush connectives. Aiming to “parse” and clarify the Nogoshi-Igarashi-
Nakamura model, we found several implicit assumptions that were meant but not
mentioned by the scientists relating the amplification at linear contact and Earth sur-
face. Some of the assumptions like “Bedrock and linear contact exist; sedimentary
bed and Earth surface do not” are obviously non-realistic (i.e., there is no contact
without sedimentary bed) but formally should be mentioned in general framework.
Interestingly, those assumptions linked with flow connectives, which are not associat-
ed with influx, seem to have a kind of “a posteriori” meaning fixing the pre-requisites
that have been used to infer an event. Also, we found it necessary to introduce a new
and more elaborated system of designations of seismic spectra to avoid confusion
present in the authors’ formulae (e.g., it is not clear that the authors’ $H_s$ and $V_s$ are
meant to include the history of passing the bedrock, which is designated by other
members of the formulae, $H_b$ and $V_b$, and also include the “surface history” of some
of the spectra, while this cannot be concluded from the formulae). Also, the event
bush clearly discerns the cases of “qualitative zero” (e.g., “no amplification at linear
contact” and, hence, no corresponding variable introduced in the bush) and “quantita-
tive zero” (seismic spectrum on the surface may be present or not, so the correspond-
ing variable is present but may be equal to zero). However, whether this can be con-
sidered a method and, if not, what should be done to make it such, will be discussed
below.

4 Discussion

The two physical tasks considered above show that the proposed approach may work
under some conditions, but this is definitely not enough to postulate that it is a method
ready for extensive application.

The following issues, to our mind, deserve particular attention: whether this good,
bad or irrelevant for physical modeling if not all subjects/predicates are attributed
variables, whether different variables can be attributed to similar subject and predi-
cates (i.e., one event bush gives birth to seemingly incompatible physical models),
how to formally define those flows in event bush which really lead to construction of
a formula (see Fig. 1c). A very important dimension of research is further elaboration
of event description grammar that would hopefully allow one to avoid nonrealistic
assumptions, which have been required by formal reasons. Perhaps, existing formali-
zation is still not strong enough to easily formulate the rules of assignment of varia-
tles to events and operations/assumptions to connectives.

More such tentative studies should be carried out by various research groups to ex-
amine the conditions under which this approach works well. It should be noted here-
with that the reported event bush failed to incorporate other models of the site effect phenomenon suggested in the literature. Obviously, there remain unresolved problems in this approach and not all of them are even well understood.

However, one of such problems is possible multiple qualitative interpretations of similar formula/model, i.e., not only one event bush may “host” several models but the opposite may also be possible.

Another point that seems noteworthy is that different mathematical expressions are often known for a similar law. Probably these should be considered as different laws with independent qualitative interpretations or there a “qualitatively preferred” form of each formula – by analogy to that among many mathematical expressions of the same law there are those which make physical sense and those that represent merely mathematical reformulation.

Importantly, this approach should work not only to explicate the existing knowledge but also to build new models and suggest physical laws.

Further research is needed to clarify these points and introduce dynamic knowledge engineering as a method of physical research. Nevertheless, when developed, this method could work well not only in physics but also in other exact sciences where reasoning proceeds partly or entirely in terms of variables, e.g., economics, or technical design. Finally, when the semantics of event description is developed equally well as semantics of objects and relations in object-based methods of knowledge engineering, not only the event bush but other relevant approaches will become applicable in “parsing”/creation of physical formulae that would better capture particular patterns of reasoning – e.g., event tree or its ramifications, to address branching processes, causal loops, to mimic cyclic ones, and so forth.

5 Conclusions

1. Only the event-based methods may allow “parsing” of existing physical laws and models and creation of new ones based on commonly-shared qualitative reasoning.

2. To develop a method of “parsing”/creation of physical models and laws, there should be a semantics of events as strict and formalized as semantics of objects and relations in objects-based methods of knowledge engineering.

3. So far, the method of event bush seems to be the only candidate that fits the above requirements and can at least efficiently “parse” physical models and laws in some cases.

4. Further research is needed to find out its limits of applicability and formulate the rules of its use in physics.
References


