A Process-Ontological Model for Software Engineering

Jari Palomäki and Harri Keto

Tampere University of Technology/Pori
Department of Information Technology
Pohjoisranta, P.O.Box 300, FIN-28101 Pori, Finland
{jari.palomaki, harri.keto}@tut.fi

Abstract. The term of a "process" is used in Software Engineering (SE) theories and practices in many different ways, which cause confusion. In this paper we will firstly give a general overview how the term "process" is used in SE context. Secondly, we describe how ambiguous the way of using the term "process" in SE discipline is compared to practical SE management. Then, thirdly, we will consider a process ontology in which everything is in a process. Our overall view is that everything in the world is composed of processes. Lastly, we propose a topological model of process in SE context.

1 Introduction

The term of a "process" is used in many widely known theories and essential practices of Software Engineering (SE). Earlier computers were used to solve mathematical problems. The same people who implemented the programs were customers of the process of a software development. Later on it was noticed that this "process in the small" was not well enough understood to deal with the emerging problems of producing a large and complex software system. Nowadays the scope of interest is widened to an engineering discipline, which concerns all aspects of software production including management and the improvement of processes. "Process thinking" has become one of the major efforts to make software engineering to an engineering, which has a qualitative value.

This paper is composed as follows. Firstly we give a general overview how the term "process" is used in SE context. Secondly, we describe how ambiguous the way of using the term of a "process" in SE discipline is compared to practical SE management. Then, thirdly, we will consider a process ontology in which everything is in a process. Our overall view is that everything in the world is composed of processes. Lastly, we propose a topological model of process in SE context.

2 Process Domains in Software Engineering

We will present four domain areas how the term "process" is used in SE context. These domain areas overlap and interact with each other.

The first domain area is a traditional *software engineering*, where a process is considered a production process, which includes support and organizational processes. For example a *software life circle model* and a *process model* are instances of these domains (see e.g. ISO/IEC 1995). We will include also a *meta level process model* to create a software life cycle process (see e.g. IEEE 1997 and OMG 2005).

The second domain area is a *software process engineering*. It is considered as a part of an *enterprise level process engineering*, which, according to Davenport (1993, 142), can be divided into a *process innovation* to implement radical changes and a *process improvement* to create and maintain continuous improvement of existing processes. Those enterprise level processes, which are closely related to customers, are called *key processes*. In the field of software there has been a growing interest on software process improvement (SPI), which has become an independent professional field of its own. We refer to ISO 15504 standard series (see e.g. ISO/IEC 2003) and SEI's CMMI model (see SEI 2002) for further SPI information. This domain area includes meta level models for creating and maintaining improvement processes.

The third domain area is a *SE knowledge model*, e.g. a framework like SWEBOK, (see IEEE 2004). It models the key competence areas needed in SE. Lastly, the forth process domain area is a *process modeling*. The information system is considered a model of the real world. Certain modeling methods are used to model abstractions of the processes of the real world.

3 The Ambiguity of a Process in Software Engineering

In the four domain areas where the term "process" is used in SE context both the intension and extension of the term "process" vary. Also, SWEBOK recognizes the ambiguous role of the term "software engineering process" (see IEEE 2004, Chapter 9, 9-1): it can refer for instance to the "right way of performing tasks", the "general discussion of process related to SE" or "the actual set of activities performed within an organization". There is an agreement on SE standards. That is, a software engineering process should be considered as "processes", which means that "there are many processes involved".

The term "process" is used in a different way in a SE discipline community compared to a practical SE management. SE management tends to make the process straightforward and as simple as possible by using the viewpoint of a key process. This approach has its roots in the definition of an organizational process and enterprise process engineering. Davenport (1993, 5) defines the process as a specific order of work activities across time and place. There is the beginning, and an end, which are clearly identified into inputs and outputs, and a structure for action. This is a very structural approach and emphasizes the order of activities.

Software engineering as a discipline has become to a phase, where a domain is recognized as a set of complex processes, where their relations, interactions, improvements, and implementations to reality are profoundly analyzed. This is due to the young nature of the SE discipline. Shaw in (1990, 21) illustrates how software engineering models develop as a result of the interaction between science and engineering. As there is ambiguity of the term of a "process" used both in discipline community and a practical SE management, we shall introduce the theoretical backgrounds of a process-ontology and a topological model to the term of a "process".

3 A Process-Ontology

The most famous work of process philosophy is Alfred North Whitehead's *Process and Reality*, (1929). He believed that all events are related to one another and to the environments in which they occur. The world can best be understood as interrelated systems of larger and smaller events, some of which are relatively stable. Whitehead's metaphysics was also a philosophy of process. Events are always changing. Change represents the actualization of certain potentialities and the disappearances of others. The world does not simply exist, it is always becoming.

According to Whitehead, the world is a process which is the becoming of *actual entities* (or actual occasions). They endure only a short time, and they are processes of their own self creation. There are also *eternal objects* to be understood as conceptual objects. They enter into the actual entity becoming concrete without being actualities themselves. Although novel actual entities are progressively added to the world, there are no new eternal objects. They are the same for all actual entities.

However, we will not make any detour into Whitehead's "process philosophy" here. Nor will we make any study of the *Process of Reality*, such a complex and a difficult book as it is.[1] Instead of that we will just adopt an idea that everything consists of processes, and that these processes are divided into eternal processes interpreted as *concepts*, and actual processes, which we will interpret to be events occupying a finite amount of a four dimensional space-time. Thus, the world is constructed out of *events*. Every event in space-time is overlapped by other events, i.e., events are not impenetrable. A space-time order results from a relation between events. Also, in terms of these events spatio-temporal point-instants, lines, surfaces, and regions can be defined by using the *Method of Extensive Abstraction* as follows, (see Russell 1927, Chapters XXVIII and XXIX, and Whitehead 1929, Part IV).

A fundamental relation in construction of point-instants in a four dimensional space-time is a five-term relation of *co-punctuality*, which holds between five events having a common area to all of them. A set of five or more events is called *co-punctual* if every quintet chosen out of the set has the relation of co-punctuality. A *point-instant* is a co-punctual set which cannot be enlarged without ceasing to be a co-punctual. The existence of point-instant so defined is provided if all events can be well-ordered, i.e. if the Axiom of Choice is true, (cf. Russell 1927, 299)

Given two point-instants κ and λ, we denote by $\kappa\lambda$ their *logical product*, i.e., the events which are members of both. If $\kappa\lambda$ is non-empty, then κ and λ are said to be *connected*. A set of point-instants is defined to be *collinear*, if every pair of point-instants are connected, and every triad of point-instants α, β, γ are such that either $\alpha\beta$ is contained in γ, or $\alpha\gamma$ is contained in β. A set of point-instants is defined to be a *line*,

[1] For more recent study of process philosophy in this line of thought we refer to *After Whitehead: Rescher on Process Metaphysics*, 2004, edited by Michael Weber.

if it is collinear, and it is not contained in any larger collinear set. The lines so defined are not supposed to be straight.

That definition of a line is analogous to that of a point-instant. It is possible to extend this method to obtain surfaces and regions, as well, (see Russell 1927, 311 *ff.*). A set of lines is called *co-superficial*, if any two lines intersect, i.e. they have a common point-instant, but there is no point common to all the lines of the set. A *surface* is a co-superficial set of lines which cannot be extended without ceasing to be co-superficial. A set of surfaces is called *co-regional*, if any two surfaces have a line in common, but no line is common to all the surfaces of the set. A *region* is a co-regional set of surfaces which cannot be extended without ceasing to be co-regional.

A *space-time order* is constructed out of the relation between events as follows. Two events are said to be *compresent* when they overlap in space-time. With respect to a given event it is possible to divide events into *zones* as follows: In the first zone there are those events that are compresent with a given event. The, in the second zone, there are those events which are not compresent with a given event, but compresent with an event compresent with it, and so on. The *n*th zone will consist of events that can be reached in n steps, but not in $n–1$ steps, in which a step is taken to be as the passage from an event to another which is compresent with it. Assuming a minimum size of events, it is possible to pass from one event to another by a finite number of steps. Two point-instants are connected, if there is an event which is a member of both. Thus, point-instants can be collected into zones as well, and the passage from event to event by the relation of compresence can be replaced by the passage from point-instant to point-instant by the relation of connection. Accordingly, suppose there are n events, $e_1, e_2, \dots, e_n$, and suppose e_1 is compresent only with e_2, e_2 is compresent with e_1 and e_3, e_3 with e_2 and e_4, and so on. We can then construct the order $e_1, e_2, \dots, e_n$. The relation of connection is a causal relation between events, where the cause of an event occurs earlier than its effect.

We shall also distinguish events in a living brain from events elsewhere, (Russell 1948, 246). So thoughts should be among the events of which the brains consist, i.e., each region of the brain is a set of events. These events are called *mental events*. Mental events can be known without inferences and they consist of bundles of compresent qualities. Events, which are not mental, are called *physical events*, and they, if known at all, are known only by inference so far as their space-time structure is concerned.

Accordingly, in our view, ontologically, everything consists of processes. Among processes, firstly, there are eternal processes and actual processes. Eternal processes are interpreted as concepts, whereas actual processes are interpreted as space-time events. Eternal processes are instantiated in actual processes. Secondly, among actual processes there are mental events and physical events. Mental events consist of bundles of compresent qualities which can be known without inferences, whereas physical events, if known, are known only by inference as regards to their space-time structure.

5 A Topological Model for a Process

We shall give a topological model for a process, in which events are interpreted as *open sets*, i.e., events will have a one-one correspondence with open sets. To get an idea, a few topological concepts are defined as follows. Consider a set *T*. Let $\{O_{i\in I}\}$ to be a set of open subsets of *T* satisfying the following axioms:

A1 The union of any number of open sets is an open set.
A2 The intersection of two open sets is an open set.
A3 *T* itself and the empty set ∅ are open sets.

A *topology* on a set *T* is then the specification of open subsets of *T* which satisfy these axioms, and this set *T* is called a *topological space*.

A set of open subsets $\{O_{i\in I}\}$ of *T* is said to be an *open covering* of *T*, if the union of $O_{i\in I}$ contains *T*. An open covering $\{V_{j\in I}\}$ of a space *T* is said to be a *refinement of an open covering* $\{O_{i\in I}\}$, if for each element V_j of $\{V_{j\in I}\}$ there is an element O_i of $\{O_{i\in I}\}$ such that $V_j \subseteq O_i$. If $\{O_{i\in I}\}$ is any open covering of *T*, and there is some finite subset $\{O_{i1}, O_{i2}, \dots, O_{in}\}$ of $\{O_{i\in I}\}$, then a space *T* is called a *compact*.

A topological space *T* is *separated*, if it is the union of two disjoint, non-empty open sets. A space *T* is *connected*, if it is not separated. A space *T* is said to be *path-connected* if for any two points *x* and *y* in *T* there exists a continuous function *f* from the unit interval [0, 1] to *T* with $f(0) = x$ and $f(1) = y$. This function is called a *path* from *x* to *y*. A space *T* is *simply connected* if only if it is a path connected, and it has no "holes".[2] A space *T*, which is *connected,* but not simply connected, is called *multiply connected.*

Given two points *a* and *b* of a space *T*, a set $\{O_1, O_2, \dots, O_n\}$ of open sets is a *simple chain from* a *to* b provided that O_1 (and only O_1) contains *a*, O_n (and only O_n) contains *b*, and $O_i \cap O_j$ is non-empty if and only if $|i - j| \leq 1$. That is, each link intersects just the one before it and the one after it, and, of course, itself. It can be proved that if *a* and *b* are two points of connected space *T*, and $\{O_{i\in I}\}$ is a set of open sets covering *T*, then there is a simple chain of elements of $\{O_{i\in I}\}$ from *a* to *b*, (for the proof, see the Theorem 3-4 in Hocking & Young, 1961). Moreover, let $C_1 = \{O_{11}, O_{12}, \dots, O_{1n}\}$ and $C_2 = \{O_{21}, O_{22}, \dots, O_{2m}\}$ be simple chains from a point *a* to a point *b* in a space *T*. The chain C_2 will be said to *go straight through* C_1 provided that i) every set O_{2i} is contained to some set O_{1j} and ii) if O_{2i} and O_{2k}, $i < k$, both lie in a set O_{1r}, then for every integer *j*, $i < j < k$, O_{2j} also lies in O_{1r}. Accordingly, the finer chain C_2 goes straight through the coarser chain C_1.

[2] More formally, a path-connected space *T* is *simply connected* if given two points *a* and *b* in *T* and two paths $p : [0,1] \to T$ and $q : [0,1] \to T$ joining *a* and *b*, i.e., $p(0) = q(0) = a$ and $p(1) = q(1) = b$, there exists a homotopy in *T* between *p* and *q*. Two maps $p, q : X \to Y$ are said to be *homotopic* if there is a map $H : [0, 1] \times X \to Y$ such that for each point *x* in *X*, $H(0, x) = p(x)$ and $H(1, x) = q(x)$. The map *H* is called a *homotopy* between *p* and *q*. Intuitively, maps *p* and *q* are homotopic, if *p* can be continuously deformed to get *q* while keeping the endpoints fixed, and a path-connected space *T* is simply connected, if every closed path in *T* can be continuously deformed into a point.

A topological model is used as follows: a process as a whole is interpreted as a topological space *T*, which, at least for empirical reasons, is compact and, depending on the number of parallel processes, is either a simply- or a multiply connected. The space *T* contains a start point *a* and an endpoint *b* of the process. The start point *a* is an event, which is included in the open set O_1, and, similarly, the endpoint *b* is an event, which is included in the open set O_n. The simple chain from *a* to *b* consists of sequences of events interpreted as a set $\{O_1, O_2, \ldots, O_n\}$ of open sets. Moreover, it is possible to get as coarse or as fine a chain from *a* to *b* as necessary. In a case there are parallel processes, i.e., processes which we want to keep distinct in a certain moment, for example feedbacks, we just add “holes” to our space *T*. This prevents the parallel processes from deforming to each other. The space *T* will then be multiply connected.

6 Concluding Remarks

We have given a general overview how the term “process” is used in SE context. We have also described how ambiguous the way of using the term of a “process” in SE discipline is compared to practical SE management. As there is ambiguity of the term of a “process” used both in a SE discipline community and in a practical SE management, we have introduced the theoretical backgrounds of a process-ontology and a topological model of the term of a “process” to be used in SE context where processes are understood as space-time events – either mental or physical.

As a result we think that a process-ontology and its topological model will provide an appropriate philosophical foundation to a SE discipline. They will give a common conceptual framework for the SE researches as well as for the SE practice; for example, it gives a possibility to compare their different SE process-models and concepts and to interpret the dependencies between two different SE models.

References

Davenport, T. H., 1993: *Process Innovation: Reengineering Work through Information Technology*. Boston: Harvard Business School Press.

Hocking, J. G. & Young, G. S., 1961: *Topology*. New York: Dover.

IEEE 1997: *IEEE 1074-1997 Standard for developing software life cycle processes.* Software Engineering Standards Committee of the IEEE Computer Society, New York.

IEEE 2004: *SEWOK Guide to the Software Engineering Body of Knowledge*, 2004 Version, IEEE Computer Society Professional Practices Committee, Los Alamitos, California.

ISO/IEC 1995: *ISO/IEC 12207 Information technology-Software life-cycle processes*. Geneva: ISO.

ISO/IEC 2003: *ISO/IEC 15504-2:2003 Information technology - Process assessment Part 2: Performing an assessment.* Geneva: ISO.

OMG 2005: *Software Process Engineering Metamodel*, v1.1. The Object Management Group, Inc. (OMG), www.omg.org.

Russell, B., 1927: *The Analysis of Matter*. London: Allen & Unwin.

Russell, B., 1948: *Human Knowledge*: *Its Scope and Limits*. London: Allen & Unwin.

SEI, 2002: *Capability Maturity Model Integration (CMMI), CMMI for Software Engineering (CMMI-SW, V1.1)*. Software Engineering Institute (SEI), Carnegie Mellon University, Pittsburgh.

Shaw, M. 1990: "Prospects for an Engineering Discipline of Software". *IEEE Software* **7(6)**, 15–24.

Weber, M., (ed.), 2004: *After Whitehead: Rescher on Process Metaphysics*. Frankfurt / London: Ontos Verlag.

Whitehead, A. N., 1929: *Process and Reality*: *An Essay in Cosmology*. New York: The Macmillan Co.