Panel Report: "Grand Challenges of Network and Service Composition"

Giuseppe Valetto[1] and Fabrice Saffre[2]

[1] Telecom Italia Lab
`Giuseppe.Valetto@tilab.com`
[2] British Telecom
`fabrice.saffre@bt.com`

Abstract. This brief report intends to summarize some of the things we learned during WAC2004 sessions "Network Composition" and "Negotiation and Deployment", and that were highlighted during the subsequent panel discussion. We try to focus especially on the following aspects: traits and trends of convergence emerging from the rather diverse findings presented in those sessions; controversial or divergent opinions on some of those findings; open issues that should be addressed by the autonomic communication community and how to tackle them; and major research directions that seem likely to emerge and shape a significant part of the autonomic communication landscape.

1 Findings and Commonalities

We notice how a certain set of issues surfaced repeatedly in both sessions. Fig. 1 attempts to draw some relationships among those issues, starting from the idea of *Adaptation* as a generic term for the self-regulating operation of an autonomic system. Several of the works presented deal with *Composition* as an important form of adaptation and a first-class autonomic primitive. *Policies* and *Negotiation* are widely regarded as useful means to strategize about and achieve adaptation through composition, in a top-down vs. bottom-up fashion: policies can dictate the terms of composition (e.g. *when, what and how*), while negotiation can be used to spontaneously reconcile competing and/or conflicting policies, and converge towards stable system configurations. Finally, *Semantic Knowledge* can be used to obtain an explicit and abstract representation of autonomic adaptation, including composition. That representation remains formal and hence can be distilled into policies, but at the same time provides the ability to understand discuss, communicate and review the autonomic behavior of a system.

Another evident common element is that wireless communication provides major motivation and an obvious test bench for the investigation of autonomic paradigms: LANs, PANs, sensor networks, as well as ubiquitous, ambient and ad hoc networking, are among the wireless and mobile contexts that look like natural catalysts for autonomic capabilities. It seems also evident that all of them require some kind of context-awareness (e.g. location-, user-, service-awareness) embedded in the autonomic communication facilities, possibly as part of the knowledge base.

Other agreed-upon, significant enablers of autonomic communication are transparent addressing, seamless handover, strong decentralization of all adaptation mechanisms, and a regard for issues like resource, security and trust management as first-class elements.

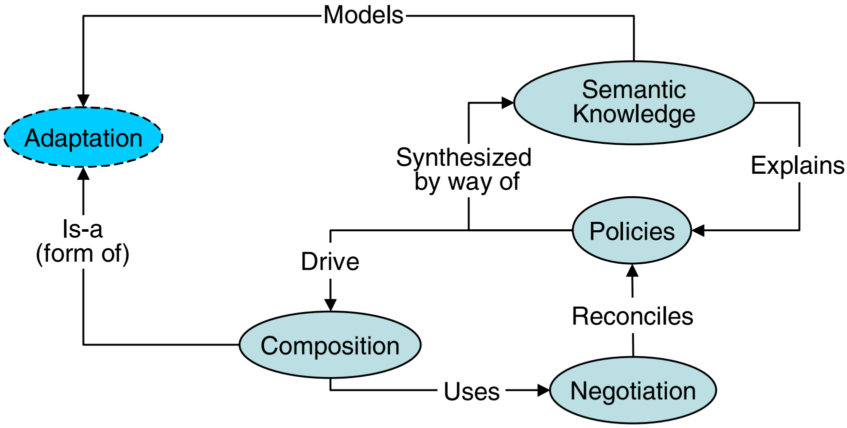

Fig. 1. Common issues and their relationships

2 Divergences

Along many of the presentations and discussions, the issue of *emergent behavior* was often either explicitly mentioned, or alluded to. However, we have recorded diverse opinions on its actual relevance to the problems in autonomic communication. On the one hand, it seems clear that it has a clear appeal, because of its resonance with the biologic metaphor at the very basis of all things autonomic, and its promise to enable complex, numerous and strongly autonomous elements that co-exist in the same environment to act in concordance and with a common sense of purpose. On the other hand, there is a feeling that the biologic metaphor should not be over-stretched, and a concern that such an approach – as well as many evolutionary or "cognitive" approaches - could break once applied to communication infrastructures and services, under their extremely demanding timeliness and predictability requirements.

In a similar fashion, different opinions exist on whether autonomic features should originate from explicit vs. implicit provisions (e.g., dedicated protocols that account for and codify features such as robustness, flexibility and fault-tolerance vs. spontaneous or stigmergetic interactions among communication elements that converge towards a mutually sustainable and satisfactorily functional configuration).

3 Open Issues and Next Steps

A number of suggestions and questions have been raised, trying to indicate how to focus the undergoing work and discussion of the autonomic communication community towards a set of incremental goals that can help map this largely uncharted territory.

A major issue seem to be trying to define the boundaries of the autonomic communication domain: it seems obvious that it comprises self-* issues within a network infrastructure, as well as at the juncture of different networks; it is also evident that it extends to a degree up to the level of the services carried by those networks, but what kind of services are going to be affected the most – and in what application domains - is not equally well understood.

A related issue is the categorization of the techniques that are relevant for autonomic communication: in Session 2A and 2B – and throughout the Workshop in general – a wide spectrum of techniques, ranging from control theory (with its hard mathematical foundations), to bio-inspired techniques (for instance, emergent behavior), and from distributed negotiation algorithms to knowledge-based reasoning have been presented and debated. It must be better understood what solutions are suitable for what problems, in a multi-dimensional space that includes qualitative parameters, such as problem type, attainable scale, level of impact, as well as quantitative properties, such as timeliness, predictability, reliability.

Another open issue of great interest of the community is the level of *transparency* that autonomic communication facilities shall strive for. Transparency is in itself a multi-faceted concept, since it encompasses technical factors (e.g., non-intrusiveness in affected systems) as well as human-observable factors (e.g., the ability to understand, analyze and influence autonomic behavior on the part of technicians, as well as seamless and “hidden” operation from the point of view of end users).

An operational suggestion to address some of the questions above is to work together to propose and formalize a sort of “autonomic communication reference problem” (or problem set), against which proposed approaches should be evaluated, independently of their application domain. The reference problem might be characterized as a “check-list” of observable and demonstrable autonomic features, which an autonomic communication system should strive to address, thus enabling a form of comparison with respect to other solutions.

4 Research Directions

Research issues in autonomic communication are likely to revolve around achieving two complementary objectives: identifying suitable design principles and testing implementation/deployment strategies. As for the first, many sources of inspiration are available from the natural world including, but not limited to, biological systems. As for the second, a number of well-established domains in computer science (machine learning, formal methods etc.) offer reliable tools and a substantial body of knowledge to start with. This is not the place to review these, however, so we chose to sketch a picture of what we believe will be a fundamental common denominator to all future research in the field of autonomic systems.

Traditional engineering starts by specifying desirable system-wide characteristics and then designs/selects individual components under the assumption that the whole is only the sum of its parts. In extremely large distributed systems, this “top-down” approach is under severe strain to deliver viable solutions, which is a major rationale for autonomic communication. Yet this new paradigm raises issues of its own, mostly due to the apparition of complexity (in the restrictive scientific sense of emergent global properties within large ensembles of interacting units). However, despite being aware of this difficulty, many technologists seem reluctant to cross the cultural barrier between a proven and immensely successful paradigm (inherited from the industrial revolution) and the newer science of complexity, which is less well understood by engineers.

Complexity science provides powerful methods for dealing with probabilistic predictability and describing in a rigorous and useful way systems comprised of individually unpredictable elements. Over the last three decades, it has been extensively demonstrated that variability in the individual response of its constituents does not necessarily translate into the frequency distribution of a system's states exhibiting a similar amount of 'noise'. On the contrary, the huge number of interactions and the presence of intricate feedback loops often mean that the system as a whole can only exist in a limited number of configurations, despite the largely random behavior of individual units. The science of complexity mainly consists of identifying these configurations, determining their probability of occurrence, and understanding/characterizing transitions between them and trajectories leading to them (e.g. bifurcation).

The sheer size of a large network comprised of many thousands of components means that the state of a large distributed computing environment will virtually always be the result of an unforeseeable combination of many events, and so can be described best probabilistically. While there may be an increased recognition of this situation, there is a poor awareness of the methods capable of dealing with it. The heterogeneity of the underlying infrastructure (in terms of purpose, capability, and ownership) precludes a centrally imposed set of rules defining the function and privileges of every participant. Instead, we must find ways to engineer autonomic principles, like self-configuration, into individual elements and their interactions, so as to allow them to deal with unexpected situations, requests, combinations of events, etc.

Complex systems theory and modelling can and must help us understand which macroscopic behaviour is more or less likely to emerge from the many interactions between heterogeneous devices. The real challenge is not to cope with microscopic unpredictability - the conceptual tools required to handle its macroscopic effects are readily available. The difficulty resides in identifying and weighting the factors involved, so that the purpose of fine-tuning the local rules is not defeated by the presence of 'hidden variables' capable of pushing the entire system into an unexpected/undesirable state.