Open, stable autonomic adaptation: a grand challenge (Extended abstract)

Simon Dobson and Paddy Nixon
Systems Research Group
School of Computer Science and Informatics, UCD Dublin IE
E-mail: simon.dobson@ucd.ie

Abstract

Treating autonomic systems as solving individual, isolated problems misses the opportunity to consider adaptation as an object of study in its own right. We propose a grand challenge of supporting adaptation in an open and stable manner, embracing the evolution of system goals and constraints over time without sacrificing our ability to describe, implement and analyse autonomic solutions. We suggest a possible approach using topology, and comment on the impact and problems of this novel approach.

1 Introduction

The motivations for autonomic systems research focus on the management of complexity and its consequent overall costs [5]. Large-scale systems face a diverse set of causes of complexity, including provisioning (matching resources to requirements), dynamism (of components, requirements and constraints), optimisation (prioritising the use of scarce resources) and openness (loose control of requirements and services). These apply equally to enterprise systems and to communications systems, and also to small-but-complex entities such as sensor networks and many embedded systems.

Addressing these problems generally involves building systems whose detailed behaviour adapts in some way in response to changes in the external environment in which the system operates. An enormous variety of techniques have been used across different problem domains [3]. Adaptation can be used to mask the impact of environmental changes from services and users, leading to the slightly paradoxical result of adaptive systems exhibiting *less* high-level behavioural variation than their non-adaptive peers.

This does, however, rather beg the question of what constitutes a legitimate adaptation: when does an *adapted* behaviour become a *wrong* or *unacceptable* one? The answer turns out to be rather more complicated than might be expected, and has implications for the specification, design, analysis and evolution of adaptive autonomic systems.

2 Open stability: the grand challenge

What does it *mean* to be an adaptive system? Clearly it does *not* mean that all behaviours can potentially be changed, since such a system could serve no useful purpose (other than, perhaps, its own perpetuation). This suggests that some aspects of the system's behaviour must remain constant (or at least change only slowly), whilst other aspects adapt in order to achieve this consistency.

One way to think about this is to define the set of acceptable externally-visible states of a system. This set defines the *behavioural envelope* of the system: the set of behaviours that we will accept as being "correct". The actual behaviour exhibited by the system may change, as long as it always remains within its behavioural envelope.

How does adaptation happen? Suppose the system is exhibiting a correct behaviour. A change in its environment may cause – or threaten to cause – its behaviour to move outside its envelope, triggering an adaptation that moves its behaviour to some point within the acceptable range. The later behaviour may be different to the earlier one whilst still being correct.

Applying adaptation in pursuit of stability raises a number of issues. For each adaptation tactic we deploy, we must ensure that it moves the system to an acceptable state from all states in which it might be applied. We must also ensure that each tactic is applied under the appropriate circumstances. We probably also want to perform the smallest adaptation consistent with achieving our aims.

Taken together, such an adaptive system is *stable*: a small change in its environment will cause its behaviour to change in such a way as to exhibit the smallest change consistent with remaining within its behavioural envelope.

For problems with well-defined and largely closed goals, this sort of adaptive behaviour is perfectly achievable – and has been demonstrated extensively within the literature. Perhaps the easiest examples involve power management, where the power drawn by enterprise servers is varied in response to changing load. (A good example is given by Kephart *et alia* [4].) Here the goal is to retain a given degree of throughput and/or latency in serving individual requests whilst minimising the power requirements. Solutions often involve using control theory to define how estimates of future service loads affect changes in power, and we can often prove the stability of the system under all sequences of stimuli.

Not all problem domains or adaptive tactics are so well-behaved, however. Adding new constraints to the system – allowing the available network bandwidth to vary, for example – would invalidate the control strategy that accounts only for latency and power, and adapting the strategy will often involve re-calculation *ab initio*. Systems using less mathematical techniques such as expert systems, policy-based management and the like may not have easily-proved stability properties, either in general or under extension. This makes openness the enemy of stability and predictability – but businesses often require both.

We could summarise this by saying that there is a difference between an autonomic *system* and an autonomic *solution* to a particular problem: the latter is bounded and determined, whilst the former must account for the evolution that most systems exhibit over time.

We may therefore formulate our grand challenge for autonomic systems as being **the development of techniques for open, stable autonomic control**. This includes:

1. specifying **behavioural envelopes** for systems, constraining the extent of variation allowed;
2. defining **adaptation tactics** that can be used to adapt the system to keep it within its envelope;
3. deciding the **environmental triggers** that cause each tactic to be applied;
4. ensuring that the application of 3 and 2 satisfy 1 **under *all* combinations of circumstances**; and
5. allow the envelope, population of tactics and selection process to **vary over time**, subject to maintaining 4.

Points 1–4 ensure stability; point 5 ensures openness.

One might argue that this is an impossible dream rather than a grand challenge, but there are reasons to believe that it is at least partially achievable.

3 A topological view of adaptation

A behavioural envelope seems like a very abstract structure, but may be regarded as the sub-set of all the possible system states that we will deem "acceptable" in some sense. The notion of "acceptability" is defined externally to the system itself, and reflects the "purpose" or "goal" of the system in its wider environment. We might of course also model the system's environment as a set of states. Adaptation involves mapping each system state to another in response to a change in the environmental state.

Consider how this formulation works in practice. The environment evolves from one state to another as, for example, the number of simultaneous requests increases. The goal of autonomic adaptation is to cause a corresponding evolution in the system state that maintains the behaviour within the envelope.

We can formalise this rather easily. Define a system state space S defining all combinations of system variables: power consumption, average response time, and so forth. A subset $B \subset S$ represents the acceptable behavioural envelope. A similar environmental state space E represents all combinations of environmental variables we can measure. Adaptation consists of defining a mapping $f : S \times E \rightarrow S$ taking each system state to another for a given environmental state (figure 1).

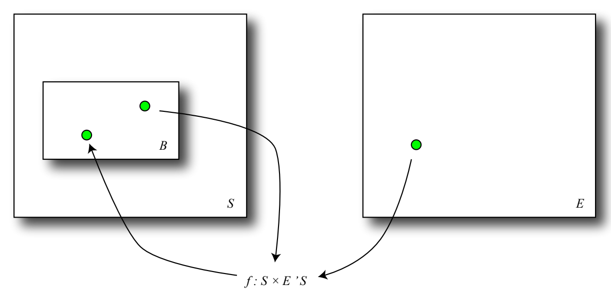

Figure 1. Adaptation

The behavioural envelope need not "expose" all the elements of the system's behavioural space. In our example above, we might consider that the visible portion of the envelope consists solely of the average response time to new requests for service, and elide the portion relating to power demand. We are then free to meet the visible constraints by *any* approach we deem appropriate, without this being exposed to users. It is this elision of information that allows for adaptation without affecting the user-visible quality of service.

We could define f piecemeal: define a set $F^* = \{f_i : S \times E \rightarrow S\}$ and a selection mapping $s : S \times E \rightarrow \mathcal{P}(F^*)$ that selects a set of possible adaptation tactics for each combination of system and environment. It need not be the case that all $f_i \in F^*$ will result in an in-envelope state for an arbitrary combination of system and environmental states: it need only be the case that $f_i(s, e) \in B$ whenever

$f_i \in s(s, e)$.

We might argue that f is defined too broadly, since it takes any arbitrary system state to another system state. The correctness of the adaptation implies at least that f is defined to always result in states in B given a state in B: $f(s, e) \in B$ for all $s \in B$. If we relax this so that $f(s, e) \in B$ for all $s \in S$ – we map *all* states, in-envelope or not, to in-envelope states – the result is the class of *self-stabilising system* studied by Dijkstra and others (see [6]).

Given that this is an autonomic system, we have a closed control loop: if $f(s, e) = s'$, our next environmental change from e to e' will result in an adaptation $f(s', e')$, and so forth. If we define the sequence of environmental changes as a function of time, $e(t)$, then f will define a system evolution from any given initial state. If we define f piecemeal as above, then if s is always a one-element set (only one tactic is available) the trajectory of f from any given initial state is entirely dependent on $e(t)$. If, however, s is potentially multi-valued, then the system may exhibit several different evolutions in response to given stimuli. The stability constraint forces all these traces to remain within B.

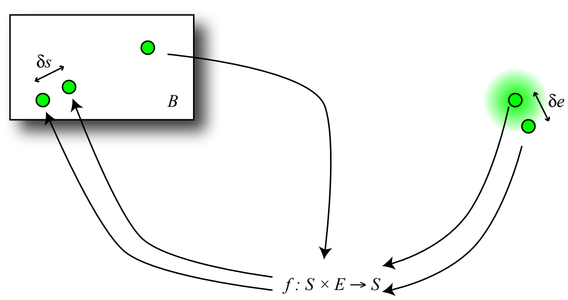

Figure 2. Adaptation in the presence of uncertainty

Of course if we have the option of multiple tactics and trajectories, we need some way in which to choose between them. One possibility is to endeavour to minimise the amount of change we observe for any given change in environment; another might be to minimise the *rate* of change, or some other function. Whichever approach we choose defines an adaptation *strategy* controlling how we apply the available adaptation tactics.

If a system is sensor-driven then the environmental state $e(t)$ at any time is not known precisely, but only within the tolerances of the various sensors. This is something that is well-understood in pervasive systems [7] but which has received less prominence for autonomic systems. We can characterise the relationship between stability and uncertainty as constraining the divergence in the mappings: observations only known within a range $e + \delta e$ must result in a correspondingly small behavioural changes $s + \delta s$ (figure 2). If this is not the case, then a small error in the knowledge of the environment can cause arbitrary changes in behaviour, which is clearly undesirable.

Openness may be defined as a transformation between spaces. Adding a new environmental constraint to a system, for example, is modeled by adding an additional dimension to the environmental state space; similarly a new behaviour can be modeled by expanding the system state space. To achieve open adaptation we must be able to lift adaptive mappings into the new spaces. We may see similar shapes to the lifted mappings, or there may be more dramatic changes: the point is that the process of evolution is well-defined (at least in principle) and admits some well-founded mathematical techniques for its analysis and solution.

3.1 Making it happen

We can continue these observations, but for the purposes of the grand challenge it is sufficient to observe that this formulation is essentially one of topology: we define a system's "shape," match its dynamic evolution on this "shape" to the "shape" of its environment subject to stability constraints, and allow the "shapes" to evolve in a controlled way in support or openness. Although this is a novel area for computer scientists, it is well-accepted and -understood in other branches of science and mathematics. Treating an adaptive system as a problem in dynamics allows us to deploy the techniques of dynamical systems that have proved themselves to be extremely powerful [1, 2].

The impact of achieving such a challenge would be immense, although the technical challenges are equally large. Such models allow closed-form analysis and prediction, allowing us to state the properties of adaptive systems with confidence. As well as the scientific benefits, such clarity has a business benefit in allowing potential purchasers to see clear added value. It further provides a sound foundation for reasoning about adaptive systems and their evolution over time in response to additional constraints and behaviours.

Technically, turning the challenge into a programme of research and development is as much a problem of mindset as of technology. Once can only analyse adaptive tactics that are stated in analytic form, meaning we must formulate the effects of (for example) control-theoretic and AI-based strategies in a common framework. This Whether this is possible or desirable remains to be seen, but early work is promising. Perhaps the most important first steps are to develop a number of exemplar systems to validate – or, indeed, invalidate – the approach and provide guidance as to the application of the techniques to realistically complex systems.

References

[1] S. Dobson. An adaptive systems perspective on network calculus, with applications to autonomic control. *International Journal of Autonomous and Adaptive Communications Systems*, 1(1), 2008. To appear.

[2] S. Dobson, E. Bailey, S. Knox, R. Shannon, and A. Quigley. A first approach to the closed-form specification and analysis of an autonomic control system. In *Proceedings of the 12th IEEE International Conference on Engineering Complex Computer Systems*. 2007.

[3] S. Dobson, S. Denazis, A. Fernández, D. Gaïti, E. Gelenbe, F. Massacci, P. Nixon, F. Saffre, N. Schmidt, and F. Zambonelli. A survey of autonomic communications. *ACM Transactions on Autonomous and Adaptive Systems*, 1(2):223–259, December 2006.

[4] J. Kephart, H. Chan, R. Das, D. Levine, G. Tesauro, F. Rawson, and C. Lefurgy. Coordinating multiple autonomic managers to achieve specified power-performance tradeoffs. In *Proceedings of the 4th IEEE Internatioanl Conference on Autonomic Computing*. 2007.

[5] J. Kephart and D. Chess. The vision of autonomic computing. *IEEE Computer*, 36(1):41–52, January 2003.

[6] M. Schneider. Self-stabilization. *ACM Computing Surveys*, 25(1):45–67, March 1993.

[7] J. Ye, S. McKeever, L. Coyle, S. Neely, and S. Dobson. Resolving uncertainty in context integration and abstraction. In D. Cotroneo and J. McCann, editors, *Procedings of the International Conference on Pervasive Services*. 2008. To appear.