# The Mechanical Homunculus

**Professor Stephen A. Gage**
The Bartlett School of Architecture
University College London
Wates House
22 Gordon Street
London
WC1 H 0QB
s.gage@ucl.ac.uk

**Chris Leung***
The Bartlett School of Graduate Studies
and Department of Computer Science
University College London
1-19 Torrington Place
London
WC1 E 6BT
christopher.leung@ucl.ac.uk

## Abstract

The use of mechanical environmental control devices that change according to changes in external environmental conditions is examined from a systems standpoint. Like the Watt's Governor these devices cannot be usefully understood when subdivided into homuncular elements linked through information flow, and any internal representation of their function is dynamical and depends on energy input. If they are to be treated as part of a wider system then communication must occur. Communication from the devices occurs when their behaviour is observed. Communication to the devices can best be effected by actively manipulating their energy input levels.

## Introduction

This paper has been written in order to clarify the way that some potentially very useful physical devices can be thought about in the context of a wider system. These devices mechanically change state as their operating environment changes and can be directly used to act on that environment or another physically adjacent environment. The devices are autonomous and require no power. They are usually regarded as being so basic that they are not worth thinking about in a sophisticated $21^{st}$ Century context.

The fact that they require no power makes them valuable and the fact that they are autonomous puts them into the same class of object as a person or an autonomous robot that can exist in an environment and act on it.

Probably the most important issue with regard to these devices is to understand where, how and if the concept of an internal model separated from physical operation is useful when we construct our own model of their behaviour. We can examine this question by looking at conceptualisations of the Watt's Governor. When we place these objects in a wider system it becomes important to know what additional physical modifications are required in order for them to modify their automatic operation. These modifications can allow communication between this type of device and entities that are often conceptualised as having a discrete internal model (people) or entities where this conceptualisation is embodied in a discrete digital control device.

*Sponsored by Haque Design + Research Limited, London, http://www.haque.co.uk

## 1 Passive environmental control

Some of the most reliable mechanisms that are used in environmental modification are mechanical devices. Many of these devices are passive and do not require anything to operate apart from an ambient energy input. As a result they are of considerable interest to Architects and Engineers who seek to reduce the overall energy consumption of buildings by using this type of device to modify the environment in buildings. Examples include traditional bi-metallic thermostats, humidity sensing ventilator openers and wax piston activated window openers.

The traditional thermostat is a bi-metallic switch with a variable set point. The active bi-metallic element is exposed to a room's ambient air temperature. When the room air falls below the set point temperature the bi-metallic element switches on an air-heating device. When the room air returns to a temperature at or above the set point the air-heating device is switched off. The thermostat forms part of a simple feedback system that includes the room air, the room air heating device and (usually) a human being who varies the set-point temperature on the thermostat. The thermostat can be placed anywhere within the environment that it forms a part of. Other similar systems incorporate devices that must be placed on the boundary between one environment and another. For example, an automatic greenhouse ventilator sits on the boundary of a greenhouse. A wax piston on the inside of the ventilator (i.e. within the controlled environment) opens the vent when the air in the greenhouse is above a set temperature and closes it when the air in the greenhouse is below a set-point temperature.

The design strategy for any such device is to look for a physical transform that can be driven by a change in the controlled environment and can in turn directly control a modifier of the environment.

Figure 1: A sculptural object-environment passive response shown in a "stowed" state

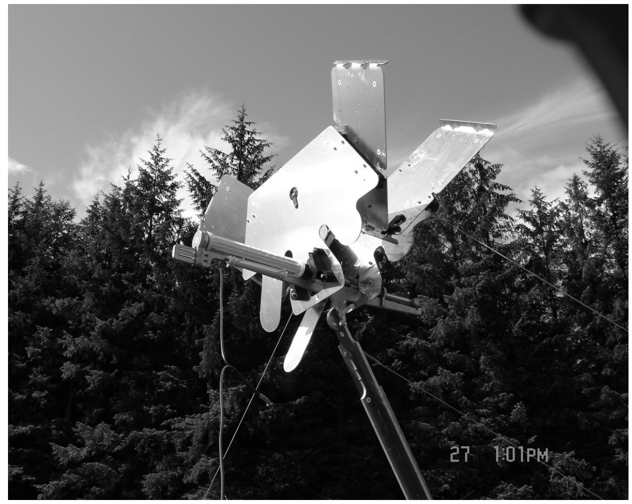


Figure 2: A sculptural object-environment passive response shown in a "deployed" state

## 2 Another class of device

There are other potentially useful reactive mechanical devices that operate in a loose and less coupled manner to modify an environment. These devices exist in one environment and modify the behaviour of another environment.

An example of such a device is a sunshade that deploys to shade part of a building when the sun falls on it. The heat energy in sunlight is used to drive the sunshade. Such a device can be designed so that a local response to changing conditions takes place. It is useful to consider the solar collector unit as being situated within a crude model, so that the shade assembly is a physical homunculus of the part of the building that is being shaded. Recent work has explored the potential of these devices for both practical and aesthetic effect, an Architect's residency in the Kielder forest by Sixteen*(makers) experimented with a dynamic sculpture shown in figure 2 in a deployed state due to high ambient temperatures and in figure 1 in a stowed state due to lower ambient temperatures [Ayres, 2006]. The underlining working principle of these devices is stimulus-response; there is no feedback from the environment being controlled back to the device. Their behaviour is not teleological in Weiner's sense [Gage, 2007a].

## 3 The 'Watt's Governor'

These devices are essentially alien in a $21^{st}$ Century systems environment. They present us with three significant and difficult challenges:

1. They work without requiring the use of a formal programming representation for their own control. How can we think constructively and inventively about them and similar devices that are yet to be invented?
2. Is it possible for an array of such devices to operate collectively, for example to anticipate events?
3. How can we envisage the interaction of people with these devices?

When we consider a building surrounded by an array of autonomous devices that respond to climatic conditions we are instinctively drawn to the image of a building surrounded by trees. Trees can protect from the sun in summer, and if they are deciduous allow sun to penetrate in the winter. However, outside of fables there is little evidence that trees can talk to each other or respond to people.

We describe above how the devices that we make are driven to change physical configuration as a result of changes in external ambient conditions. They are analogue climatic measuring devices that represent an aspect of the changing climate in a useful way.

The issue of representation is central to the answer of our three questions especially because questions 2 and 3 suggest that some form of intelligence might exist in the assembly of climate, devices, building and occupants.

We can find no direct academic reference to crude stimulus-response analogue devices that help us; however, there are many references to the Watt's Governor. The Watt's Governor exists as a central metaphor in cybernetics and control theory. It is a close-coupled mechanical device that forms part of a feedback system and is therefore more sophisticated than the stimulus-response mechanisms that are the subject of this paper. In [Gelder, 1995] Tim van Gelder uses the metaphor of the Watt's Governor to illustrate both the essential nature of it's mechanical connectivity and the fact that it's "information" only exists when it is in a *dynamical state*. Tim van Gelder compares a mechanical Watt's Governor with a hypothetical computerised steam Governor carrying out the following operations:

1. Measure the speed of the flywheel
2. Compare the actual speed against the desired speed
3. If there is no discrepancy, return to Step 1 Oth-

erwise,

**a** Measure the current steam pressure;

**b** Calculate the desired alteration in steam pressure

**c** Calculate the necessary throttle valve adjustment

4. Make the throttle valve adjustment
5. Return to Step 1
[Gelder, 1995]

The Watt's Governor cannot be understood in isolation. The Governor sits in a dynamic system that also includes the engine, a boiler, a pressure release valve and the person who stokes the boiler. On inspection the computerised steam Governor can be broken down into the following devices:

- A machine to give rotary motion using steam
- A source of steam
- A measuring device that gives current steam pressure (output digital)
- A measuring device that gives rotation speed (output digital)
- A memory device to give desired speed (input & output digital, includes a clock)
- A throttle on the steam line (input digital)
- A computational device to calculate throttle valve adjustments (input & output digital, includes a clock)

The use of devices that use digital electronic communication protocols means that these devices can be physically remote from one another and that multiple substitutions are possible.

It is very different when mechanical linkages directly amplify the effect of shaft speed and use the effect to control the speed.

It is difficult to see how a mechanical Watt's Governor could anticipate events, though relatively easy to see how the computerised Watt's Governor could access other, more sophisticated digitally encoded information. Two papers stand out: "Anticipation in Cybernetic Systems: A Case against Mindless Anti-Representation" [Schomaker, 2004] and "Dynamics and Decomposition: Are They Compatible?" [Bechtel, 1997]. Both papers are centred on problems of cognition and both extend into neuroscience and robotics; both use the Watt's Governor as a key reference. Both authors quote van Gelder [1995][1999a] and both attempt to refute him and also Van Orden [Orden and Papp, 1997] by suggesting that the Governor is either an abstraction or a linked set of abstractions. Bechtel's description of the mechanical Watt's Governor consists of a deconstructed set of parts, each passing "information" to its neighbours. Bechtel believes that the mechanical and digital components can be considered to be equivalent. He describes these components as homunculi, little agents responsible for one operation in the overall activity of the system.

The homunculus metaphor of cognition derives from the homunculus of Alchemy, a "false human being". The alchemical homunculus is made from bones, hair and semen buried in dung for 20 days before turning into a foetus. Cognitive scientists have used this metaphor to describe a virtual man existing in the brain and some have then deconstructed the man to become a whole population of mechanistic operators.

## 4 Envisaging mechanical environmental modifiers

Mechanical environmental modifiers are close-coupled devices. A holistic conception of a particular device must incorporate the role of the device as an environmental modifier, the mechanical driver of the device, the energy source and the logic that translates the available energy and drive into an appropriate modification. If the homuncular analogy is of value it stops at the scale of the device. Any further subdivision is of no value and is the equivalent of attributing independent functionality to the revolting components of the Alchemical original in their primary non-functional state.

## 5 Digital links: Why make them?

The hypothetical computerised steam Governor is similar in its operating principles to comparable combustion engines with digital controls. The physical separation of sensing, actuation and computation enables duplication; complex mechanical linkages that are subject to wear and distortion are eliminated.

These reasons might be good enough to abandon the mechanical Watt's Governor in the $21^{st}$ Century. Assume nevertheless that the mechanical Governor remains much more effective than the computerised version; there could still be compelling reasons to abandon it. The Governor might want to "talk" to other Governors in a wider system, "talk" with people remote from it, modify its behaviour in response to stimulus and "remember" previous behaviour. These are attributes of digital systems that are hard (but not impossible) to replicate in mechanical devices.

## 6 Internal representation

This leads to a philosophical question; What are the needs of the wider system of the engine, boiler, stoker and Governor to have for an internal representation of itself? The main argument in [Schomaker, 2004] is directed against anti-representationalism:

> "The work of Brooks [Brooks, 1999] in robotics, Braitenbergs [Braitenberg, 1984] revival of the ideas of Grey Walter [1963][1950] on emergent complexity in behaviour and the work of ecological psychologists Gibson [Gibson, 1979] and Kelso [Kelso, 1995] have provided its basis. A new belief has emerged under the name of anti-representationalism, which is strongly opposed to the notion of representation in cognition. Partly this development is spurred

by the very fruitful insight that some forms of behavioural complexity can be brought about by simply mechanisms at a low systematic level, taking into account embodiedness of the organism and its embeddidness in an (ecological) context. Although we do not really know what representation are, and although it is probably is good thing to be sceptical of the constructed representations in some toy models of cognitive science and traditional artificial intelligence, there is a fundamental problem with extreme anti-representationalism. In the context of the current paper, to represent the external world means to present this external world again internally within a cognitive system, in an adopted form its physical projections which is however sufficiently informative for the organism to allow for selection, preparation and control of behaviours, which are conducive to survival in its ecological Niche. It should be noted for the sake of argument that contrary to [Gelder, 1999b] it is assumed here that a given dynamical-system parameterization (e.g. an instance of Watt's Governor) actually does represent aspects of its environment, while functioning. The difference between a biological control system and an engineered negative-feedback device such as Watt's Governor is that the parameters of the biological control system are autonomously tuned during organism-world interaction. No matter what type of dynamic system is being implemented, the essence is that the parametric details are being tuned by a complex neural apparatus itself."
From [Schomaker, 2004]

This is quoted at length because his penultimate statement is quite extraordinary in the context of over forty years worth of attempts to engineer electro-mechanical devices that possess the latter attributes.

This search was part of the classic AI endeavour of the latter half of the $20^{th}$ Century that attempted the task using the symbolic representations advocated by Bechtel. The work of Brooks [Brooks, 1999] offered an alternative distributed model of layers of control in a subsumption architecture, a kind of distributed representation. This approach is of immediate interest to designers who wish to incorporate autonomous devices into buildings because it allows us to recognise the essentially decoupled nature of buildings. Buildings are decoupled from immediate functionality [Gage, 2007b] and building users are physically decoupled from buildings and each other. The continuous reconstruction of a building as a complete system must recognise the fluidity of these interactions and go beyond the work of Brooks into the field of constructing interactive systems where autonomous agents can coexist with each other in what is, none the less, a goal directed context.

An early figure in this is Gordon Pask with his "Colloquay of Mobiles" at the Institute of Contemporary Arts in London, an exhibition that took place in 1968. Pask's mobiles varied their behaviour in response to each other and to human intervention [Pask, 1967]. In this the mobiles were designed to deal with changing environmental conditions. Consider Pask's description of the goals of his mobiles:

1. The goals of the several mobiles should be partially incompatible so that the mobiles compete with one another.

2. Some of the goals should be incompatible of attainment by any one mobile on its own. In order to achieve such a goal, at least a pair of mobiles must co-operate and in order to co-operate, they must communicate with one another.

3. The main goals of a mobile should be decomposable into sub-goals so that any mobile contains an hierarchical organisation.

4. Co-operative interaction must involve main goals and sub-goals so that there are several levels of communication in the system.

5. The pursuit of the lowest level sub-goals should be carried out by autonomously acting programs embedded in each mobile. Whereas selection of these programs depends upon communication mediated feedback, their execution does not. This is one way (incidentally, a biological important way) of decoupling the mobiles and maintaining their individually integrity.
[Pask, 1967]

Pask's mobiles were driven by an array of analogue components coupled with electro-magnetic counters. These devices were more sophisticated then a Watt's Governor and much more complex than a stimulus-response device.

Pask's construction of learning [Pask, 1975] and Glanville's construction of memory [Glanville, 1976] are both based on circular iteration and are dynamical systems in van Gelder's terms. Schomaker's description of representation is close to Glanville's description of memory; he accepts that the mechanical Watt's Governor can represent aspects of its environment as it spins. It is not too much of a jump to suggest that the loose system of dynamical interactive objects in Pask's mobiles are both the actuality and the representation of the environment in which they occur and the immediate (but not long term) memory of the past states of this environment. Pask enthusiastically sought to find modes of parallel analogue computing because he was of the view that only this technology could reflect the interactions that he saw in the world. The linear state transformations in silicon chip computers are pale shadows of this but they are extraordinarily powerful. They can be used to mimic dynamical memory and learning and "to automatically tune parametric details during organism-world interactions".

## 7 Information flow techniques

If we wish to use stimulus-response mechanisms for environmental modification that are directly powered by ambient energy in a system that must learn to modify its behaviour either globally or locally then we must provide "access windows" in these mechanisms that permit two-way information flow and associated computation. Information from the device is simple in that a substantial array of possible sensors can be used. Information flow to the device is more complex because the device is not "switched on to activate". The only way to change the state of the device is to mimic changes in ambient conditions (see below).

## 8 Local and global behaviour

We return to the example of a facade of solar shutters each individually driven by a solar collector: the facade treated in this way will simply respond to local conditions. If a tree shades a part of the facade at a particular time of day the shades in that part will remain open for that part of the day. A facade that consists of an array of these passively driven homunculi will in the first instance display only local behaviour. However, if there is a finite deployment time it may be useful not to experience sunlight on an unshaded window while the shade absorbs the necessary solar energy to deploy. Now imagine that the shades can "talk" to each other; knowledge that a shade is in sunlight can be passed to a neighbour and a vector of change can be locally established. A shade can begin to close before a shadow leaves it. Alternatively, the whole facade can be observed by a camera looking for the same vector information and appropriate shades can be driven to deploy. Just like the spinning Watt's Governor the state of the shutter is given by the energy state that the solar collector is in; neither are switches in the usual sense. In order to change the state of the shutter some method must be found to replicate the behaviour of the sun. This will cost energy. If the shutters are being deployed for the convenience of users who may or may not be behind them then a deployment that costs energy should only be made if the shutter "knows" that there is someone in or approaching the space it is servicing. If this is not the case then an anticipatory deployment is unnecessary. User shutter deployment can also take place through direct user intervention using the same energy input route.

## 9 Validation

The internal representation load in this array is not heavy. However a considerable representation problem does occur when we look to test an environment of this nature before it is built. We must not only model the behaviour of the components in different environmental conditions, we must also model the behaviour of the local climate and the behaviour of people on a moment-to-moment basis. The latter are very difficult. It may well be better to use very approximate techniques for this and construct sufficient "slack" in the physical hardware to allow for running modification during service.

## 10 Conclusion

This gives us an overall picture of an environment being modified by homunculi being placed outside it which in some ways resembles the subsumption architecture proposed by Brooks. The homunculi operate as physical models of the environmental conditions that need to be modified and by operating they modify them. The homunculae operate autonomously and without any additional formal self-representation. They can be made accessible both on their output side and on their input side. Local communication allows for object-object and user-object interaction. Global observation by a vision system allows for a system observer-object interaction. It is possible in this context that anticipatory behaviour can be constructed through multiple routes each incorporating their own mode of representation and a protocol of communication that crosses all routes.

At the Bartlett, we are currently developing similar devices with more complex functionalities, our view is that these will be of great value in the built environment. It is also possible that the conceptualisation that we offer here may have wider application in the systems community. It shows a way of uniting the local bottom-up world of the anti-representationalists with the global top-down world of the representationalists in a de-coupled dynamical system based on communication.


## Acknowledgements

Contributions to this paper result from the Architect's residency by Sixteen*(makers) in the Kielder forest, Northumberland, United Kindgom during since 2003. Contributions to this paper are also from research work carried out as part of the EngD (Engineering Doctorate) programme in VEIV (Virtual Environments Imaging and Visualisation) have been supported by a grant from the EPSRC (Engineering and Physical Sciences Research Council) in the UK.



## References


[Ayres, 2006] Phil Ayres. Constructing the specific. In Kaas Oosterhuis and L. Feiress, editors, *GameSetandMatch II*, pages 314–321, Rotterdam, 2006. Episode Publishers.

[Bechtel, 1997] William Bechtel. Dynamics and decomposition: Are they compatible? In *Proc Australian Cognitive Science Society*, 1997.

[Braitenberg, 1984] Valentino Braitenberg. *Vehicles: experiments in synthetic psychology.* MIT Press, Cambridge, MA, USA, 1984.

[Brooks, 1999] Rodney A. Brooks. *Intelligence without representation in Cambrian Intelligence : the early history of the new AI.* MIT Press, Cambridge, Massachusetts, USA, 1999. (originally in AI Journal, Vol. 47 pp 139-159).

[Gage, 2007a] Stephen A. Gage. The boat/helmsman. In *Technoetic Arts*, volume 5, pages 15–24, 2007.

[Gage, 2007b] Stephen A. Gage. Constructing the user. In *Systems Research and Behavioural Science*, volume 24, pages 313–322, 2007.

[Gelder, 1995] Tim van Gelder. What might cognition be, if not computation? In *The Journal of Philosophy*, volume 92, pages 345–381, 1995.

[Gelder, 1999a] Tim van Gelder. *Mind and Cognition*. Blackwell, 1999.

[Gelder, 1999b] Tim van Gelder. Revisting the dynamical hypothesis. *Preprint*, 99, 1999. University of Melbourne, Department of philosophy.

[Gibson, 1979] J. J. Gibson. *An ecological approach to visual perception*. Hougton Miffin, Boston, USA, 1979.

[Glanville, 1976] Ranulph Glanville. What is memory that it can remember what it is? In R Trappl, editor, *Recent Progress in Cybernetic and Systems Research*, volume 7 of *Third European meeting on cybernetics and systems research*, Washington D.C., USA, April 1976. Hemisphere Press.

[Grey, 1950] Walter W. Grey. An imitation of life. *Scientific american*, pages 42–45, May 1950.

[Grey, 1963] Walter W. Grey. *The living brain*. W.W. Norton & Company, 1963.

[Kelso, 1995] J. J. Scott Kelso. *Dynamic patterns: the self-organisation of brain and behaviours*. MIT Press, Boston, 1995.

[Orden and Papp, 1997] Guy C. Van Orden and Kenneth R. Papp. Functional neuroimages fail to discover pieces of mind in parts of the brain. In *Philosophy of science*, pages 85–94. University of Chicago, USA, 1997.

[Pask, 1967] Gordon Pask. *Cybernetics, art and ideas*, chapter A comment, a case history and a plan, pages 76–99. Studio vista, London, 1967. (originally in "Cybernetic Serendipity", Rapp & Carroll).

[Pask, 1975] Gordon A. S. Pask. *Conversation, cognition and learning : a cybernetic theory and methodology*. Elsevier, Amsterdam, 1st edition, May 1975.

[Schomaker, 2004] L. Schomaker. Anticipation in cybernectic systems: A case against mindless anti-representationalism. In *IEEE International Conference on Systems, Man and Cybernetics*, pages 2037–2045. IEEE, 2004.