A Framework for a Situated Design Agent: Gero and Fujii Revisited ⋆

John S. Gero

Krasnow Institute for Advanced Study and Volgenau School of Information Technology and Engineering, George Mason University, USA; Faculty of Information Technology, UTS, Australia

Gregory J. Smith

Faculty of Information Technology, UTS, Australia

Abstract

This paper describes situated designing as an interplay between interaction, constructive memory, the role of experience and situations. It revisits work of Gero and Fujii that introduced a framework for an approach to designing that was situated, with concepts being formed as a consequence of the situatedness of designing. Research done since that time has furthered our understanding of situatedness and constructive memory as they apply to computational models of designing. The model described by Gero and Fujii is re-interpreted in the light of such research. What situations are and how they affect designing is considered. Constructive memory is then illustrated via a metaphor using moiré interference patterns. Finally, an experiential model of situated agency inspired by Dewey and Bartlett is introduced.

1 Introduction

A designer interacts with their environment while designing. The environment contains data in the form of sketches and drawings, the client, other designers, and other things. These data are interpreted by designers as including the initial requirements and constraints about what is being designed, and of

⋆ This research is funded by the Australian Research Council, grant number DP0559885. Gregory Smith now works at the CSIRO Tasmanian ICT Centre, Australia.

Email address: john@johngero.com (John S. Gero).
URL: http://mason.cmu.edu/~jgero (John S. Gero).

the developing design. Designing is treated as an activity that produces the structure of an artifact capable of providing the behaviours and functions that are expected to be provided by the artifact. The result of designing is design descriptions, which explicitly depict the specification of the structure of an artifact. The consequences of designing are the behaviours of the artifact and the functions that the artifact has potential to provide. The structure of the artifact is presented in the form of a representation that is external to the designer.

In this work a designer is considered to be a design agent. The environment holds depictions of designs as external representations such as in drawings. These external representations are acted on by agents and become designs once they are interpreted by agents. The agents observe the environment and perform actions, so changing both the environment as well as the relation between the agent and that environment.

Gero and Fujii [1] presented a framework for concept formation in designing that was based on available cognitive and computational research. Their aim was to deal with these aspects of situatedness and emergence in designing. Since that time new insight has been gained into the theories underlying it, into the application of those theories to designing, and into methods of formalisation of those theories. This paper is the first of two in which we continue the work started by Gero and Fujii. This paper revisits the issues canvassed by Gero and Fujii by reconsidering what constitutes situated designing, what a constructive memory is for, and what role situations play in designing. The ideas presented here are drawn largely from cognitive science and artificial intelligence. However, we are not aiming at proposing a computational model explaining human cognitive processes. Rather, we aim to make a computational agent produce and represent what we call a design process.

2 Re-interpreting Gero and Fujii

> "What experience suggests about itself is a genuinely objective world which enters into the actions and sufferings of men and undergoes modifications through their responses", Dewey [2].

Experiences change us, changing how we later recall having had those past experiences. The authors of this paper have had many experiences since [1] was written, so it seems fitting that we begin with a re-interpretation of that work. Figure 1 and Table 1 show a simplified version of the diagram that was central to [1].

A sensor is a construct possessed by an agent that receives a stream of data

Table 1
Symbols used in Figure 1 and the associated Section 2 text

Symbol	Meaning
A	Action activators
C	Conceptors
D	Expectations of concepts
E	Effectors
H	Focused concepts
P	Perceptors
Q	Expectations of percepts
S	Sensors
V^e	Exogenous variables

from the environment. Perceptors, conceptors, (action) activators and effectors are also constructs. Sensors, perceptors, conceptors, activators and effectors may perturb each other; each particular perturbation is an effect by one construct on another. A sensation is a process of a sensor that constructs sense-data, where a process is the execution of a function by a construct. Perceptions, conceptions, actions and effections are also processes. Sense-data are notions of an agent. Percepts, concepts, acts and effects are also notions.

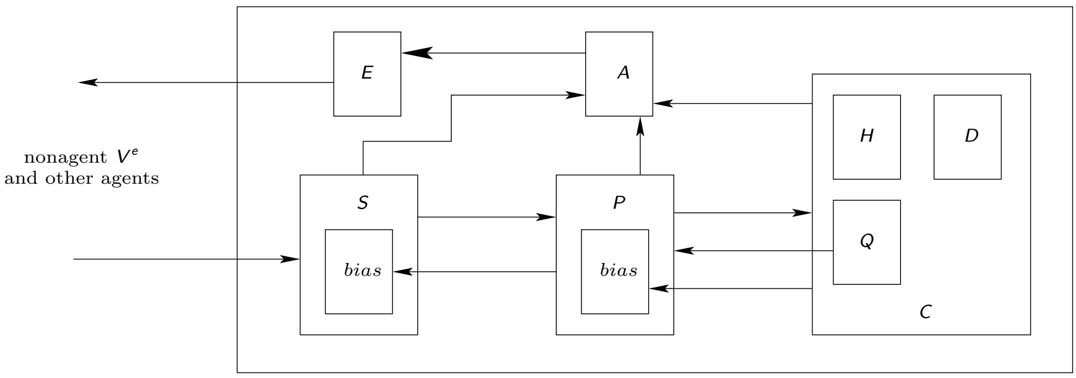


Fig. 1. A "framework for concept formation" revised from [1]. The largest box indicates one agent interacting with the rest of the environment. Smaller boxes indicate subsystems of the agent. Arrows indicate effects of subsystems on other subsystems. Effects within S, P, C, A and E are not shown. V^e are exogenous variables.

Agents interact with their environment via sensors and effectors. Sensors sense exogenous variables, and percepts abstract sense-data. Effectors change exogenous and autogenous variables, and acts abstract effect-data. An agent may possess multiple activators. A are activators of the agent that determine whether and when to effect actions, and actions are behaviours of activators.

There are three ways for sense-data to result in new effect-data, corresponding to three kinds of agent behaviour [3]. Reflexive behaviours are those processes in A that arise from sensor triggers. Reactive behaviours are those processes in A that arise from perceptor triggers. Reflexive and reactive behaviours are what are usually considered to be “situated action”: coordinated action without explicit deliberation [4], or what Agre [5] called improvised action. Such behaviours act *with* the environment rather than *on* it. But design agents need to deliberate if they are to achieve design goals. Such deliberations are reflective. Reflective behaviours are those processes in A that arise from conceptor triggers. They occur when an interpretation causes concepts to change, in turn causing new actions to be devised and effected.

V^e are exogenous variables. Activators use effectors to change the environment, to communicate with other agents. Actions that change their own exogenous variables involve changes both to V^e and to autogenous variables.

An agent may possess multiple sensors and multiple effectors. S are sensors of the agent, being those subsystems of agents that are effected by a system (agent or otherwise) not in the agent of that sensor. E are effectors of the agent, being those subsystems of agents that effect a system (agent or otherwise) not in the agent of that effector. A sensation is a process in a sensor that turns inputs into sense-data, and so is a behaviour of that sensor. Effections, similarly, are behaviours of effectors. Sense-data are employed to construct the agent’s own view of the environment.

An agent may possess multiple perceptors. P are perceptors of the agent and they transform sensory experiences in S into changes to percepts on the basis of H, Q and memories. Q are expectations of percepts and H are focused concepts that abstract the expected effects of actions. Percepts are not direct projections of sense-data but are acquired as a result of the interaction of sensory experiences and concepts.

Concepts abstract percepts so percepts are taken as the input by conception and perception is biased by conception. C are conceptors of the agent, and conceptions are behaviours of the conceptor. Percepts and concepts collectively constitute the agent’s interpretation of the environment. D are expectations of concepts, abstracting the interpretation of the environment.

Sensors, perceptors and conceptors can push or pull data. In a sensation push, for example, changes in exogenous or autogenous variables trigger sensory experiences. An example of a sensation push is a temperature sensor that works by a voltage being induced across thermocouple by the temperature of the environment. Another example is an agent registering a callback function with a world server, such that the callback is used by the server to notify the agent of changes to the environment. In a sensation pull, perception biases new sensory experiences such that some changes to exogenous or autogenous variables are filtered out, emphasised, or distorted.

3 Situated designing

These descriptions we would still largely support were it not for a concern that has grown as our ideas of situations and a constructive memories have firmed. Before addressing that concern, however, we need to say what it is that we want our situated design agents to do.

We begin, as a thought experiment, considering a 17 year old student learning to drive for the first time with a driving instructor. For some reason the student has progressed through high school with the belief that "green" is the colour that everyone else would call red and that that "red" is the colour that everyone else would call green. Two types of beliefs are involved here. If the student is looking at a traffic signal then the first kind are percepts that turn visual sense-data into representations of what colour light is active at that time. The second type are concepts that map visual sense-data onto words "red", "green" and "orange" (and vice versa).

Assume that the student has passed a written or verbal test showing that they know their road rules, and are now going for their first drive with the instructor. The student knows that a "red" light means stop and a "green" light means go. Approaching a traffic signal for the first time, the instructor sees a "red" light and yells "stop", which confuses the student because they can clearly see a "green" light. Confusion all around.

The point of the story is to trigger thinking about about how to present agents and their environment in a way that describes what they do and what effect they have on each other, but without presuming any particular kind of implementation. In the case of the thought experiment the student and the instructor communicated and, until they actually approached a traffic light, believed that they understood each other. In agent terms, each agent receives messages from the other, interprets those messages and either sends messages in return or else acts directly on the environment. As long as messages received agree with expectations there is no reason for each agent to believe that there is any lack of common understanding. In actuality, though, the student and the instructor had individuated the world differently and had differing beliefs.

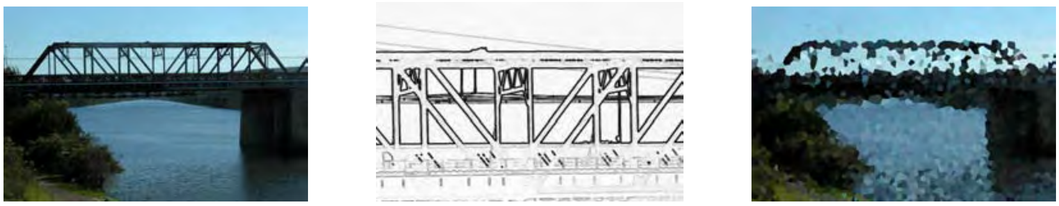
If we were using a possible worlds semantics for whatever descriptive language we chose then the semantic primitives would be sets of individual entities and sets of possible worlds. That is, we would say that there is an actual environment that is a certain way regardless of whether there are agents present to perceive it or not, and that environment is composed of entities that may or may not be accessible to agents. An alternative says that there can be no un-constructed natural facts and no independent world against which to determine the truth of a fact [6]. In this constructive view there is no isomorphism between an entity in the "real" environment and an internal model of an agent that perceives it [7]. The idea is of the perceiver of an entity in the environment as being like the captain of a submarine [8]: they have knowledge of the medium in which they are submerged but they cannot experience it directly.

That being the case, it is the agent that individuates what is in the environment. There should be no single objectively valid way of describing the set of entities that the environment consists of. An agent points a sensor at some location in an environment, the sensor receives a stream of data from the environment, and the agent interprets it. If only a part of an experience of an entity is actually what was sensed, memory of that experience at a later time will largely be a memory of an inference. If perception is constructed then a memory of it should also be, at least in part. Recollection involves a reconstruction of past experiences based on information in the current environment and on the way cognitive processing is currently accomplished – more like fantasising than looking up records [9].

If I strike a tuning fork, waves of air pressure result. These waves are not sounds. Sound requires a sensor (an ear or microphone) and perception, and interpreting sequences of sounds as music requires further reasoning. Recognising that there is a situation is like recognising that there is music, and so requires an agent. The agent affects entities in the system and new sense-data get interpreted, potentially resulting in changes to the situation. This is, to risk stretching another metaphor too far, a little like Heisenberg's uncertainty principle: the act of observation changes the system and so the situation. The Copenhagen interpretation of quantum mechanics says that an electron does not exist until something registers its existence:

> "No elementary phenomenon is a phenomenon until it is a registered (observed) phenomenon", Wheeler paraphrasing Niels Bohr [10].

This can be taken in two ways. The first takes the quote at face value: an electron *really does not exist* until there is something that registers that existence. It is a metaphor for a radical form of constructivism that begins at an agent's sensors and ends at its effectors. A less radical interpretation is to consider anything not empirically testable as being beyond *scientific* theory, and so the electron does not exist *from the viewpoint of quantum mechanics as a scientific theory* until the electron is observed. This is a metaphor for a less radical form of constructivism: the environment beyond an agent may well have some independent, objective existence but it is not accessible to the



(a) Box truss railway bridge over the Nepean River, Australia

(b) Visualisation of edge detection sense data

(c) Visualisation of colour sense data

Fig. 2. Examples of sensor data

agent except via sensors and effectors. We prefer the less radical metaphor.

Four important notions interplay in situated designing: the interaction between an agent and its environment, the influence of experiences of that agent on future patterns of interaction, the influence of constructive memory on how the agent behaves, and situations. The role of interaction was illustrated by the water flow and music analogies above. Such a role is well known in designing from Schön's [11] work on the "conversation" that occurs between a situated designer and the medium of the design. It is a dialectic view of designers experimenting with worlds that do not exist yet [12]. The role of experience is partly one of improving performance through learning, but that is not the only role. Memory plays a role that reflects how a situated agent has adapted to its environment. Constructive theories of perception and memory deal with how an individual agent interacts with its environment. At the lowest level, an agent receives sensor signals. These are uninterpreted and, as the world does not come pre-organised into categories, a process of assimilation constructs percepts out of these sensor signals according to the agents current goals and expectations. "A cognitive organism perceives (assimilates) only what it can fit into the structures it already has" (von Glasersfeld [7]).

Computationally this looks a little like perception as abduction. An agent receives a stream of input sense-data from the environment and sends a stream of output effect-data to the environment. The task of perception is to use existing knowledge to construct a set of descriptions consistent with existing knowledge. If we let Δ be the percept descriptions, Γ be received sense-data, and Λ be existing knowledge possessed by the agent, then what we are asking perception to do is find a Δ such that there is no $\Lambda \wedge \Delta$ for which Γ is false, or such that $\Lambda \wedge \Delta \models \Gamma$ [13]. This is abduction.

Percepts are pulled from a perceptor by changing biases and expectations in Λ and using these on existing Γ. Percepts are pushed through a perceptor by using the current Λ to update the current Δ on new Γ. Percepts can be noisy, ambiguous and even paradoxical, and all depend on sense-data so there are no objectively valid beliefs about the environment by agents.

Consider an agent looking at the view shown in Figure 2(a). If the agent has a sensor of visual edges (as shown in Figure 2(b)) then Γ will be a sequence of line intervals, Λ will be knowledge of the shapes things, and Δ will be a description that interprets these Γ. An agent with appropriate expectations (civil engineering knowledge Γ) should interpret Figure 2(b) as being a box truss; it is unlikely that the same interpretation would result from Figure 2(c). To get to a description of a truss would involve connecting together line intervals from Γ into polygons, aggregating these so that it was valid to infer Δ as being a kind of box truss. But for an agent to follow such reasoning requires as Λ quite a number of existing expectations. It is not the only valid interpretation. An agent with only Figure 2(b) sense-data and figure/ground expectations could validly decide that the view is a set of overlapping triangles with noise. An agent with only Figure 2(c) sense-data could employ histogram-based classification, or morphological operators for boundary detection, or any number of filters to arrive at other interpretations.

We posit that an agent individuates entities as being in the environment only when the situation is such that the agent ought to distinguish those entities from others. That is, an agent has a partial view of the environment and that view depends on expectations by the agent in the current situation. Figure 3(a) shows an example notion space by way of analogy. Figure 3(b) illustrates how the way that this is viewed could be changed by the situation via warping, filtering, and filling-in. These three enable the kind of constructive memory "schemata" that goes back at least as far as Bartlett [14]. It is Cherniak's [15] memory querying as reconstructing "what must have occurred" from a few fragments. The fragments are warped and filtered query inputs, and the current situation does the filling-in of what is missing from that warped and filtered input. This can mean that some notions are more quantitatively or qualitatively interesting, or are less quantitatively or qualitatively interesting, or that gaps can be filled where insufficient data are otherwise available.

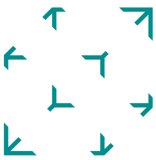
Figure 4(a) is a classic example from Kanizsa [16]. Persons viewing Figure 4(a) often interpretation it in the Gestalt continuation manner of Figure 4(b). Once that interpretation has been made the famous Necker cube interpretation can be made, depicted here as Figure 4(c). Figure 5(a) is Escher's Swans. This can be interpreted in a number of ways. Figures 5(b) and 5(c) are figure/ground interpretations where the focus of attention is on, respectively, white birds and dark birds. If the figure/ground shapes are ignored (treated as texture), a Möbius strip can be seen.

The view of perception as abductive inference on background knowledge and expectations is particularly applicable when the things being perceived are external representations by designers. A few more examples will illustrate. Why is Figure 6(a) a sketch of a church, and not just a set of scribbles? Why is Figure 6(b) a conceptual sketch of a house among trees, and not just a set

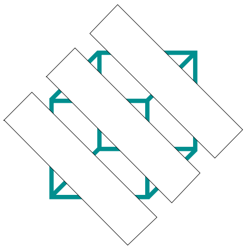
(a) Illustration of a space of notions as an abstract image of a space through which are distributed elements of varying types (although it is really of the Horsehead Nebula, from http://antwrp.gsfc.nasa.gov).

(b) A situation applied to Figure 3(a); in this example the space has been warped (as in this situation, some objects appear to be quantitatively or qualitatively different) and filtered (as in this situation, some objects are not quantitatively or qualitatively interesting)

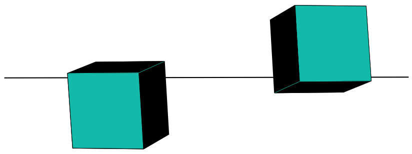
Fig. 3. Illustration of a space of notions



(a) Kanizsa-style occluded Necker cube

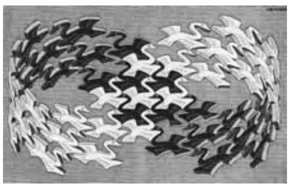


(b) Interpretation of Figure 4(a) as being an occluded thing



(c) Two different interpretations of the occluded thing entity

Fig. 4. Example of perception after Kanizsa [16]



(a) Swans, 1956 wood engraving by Escher

(b) Visualisation of a figure/ground interpretation with focus on white birds

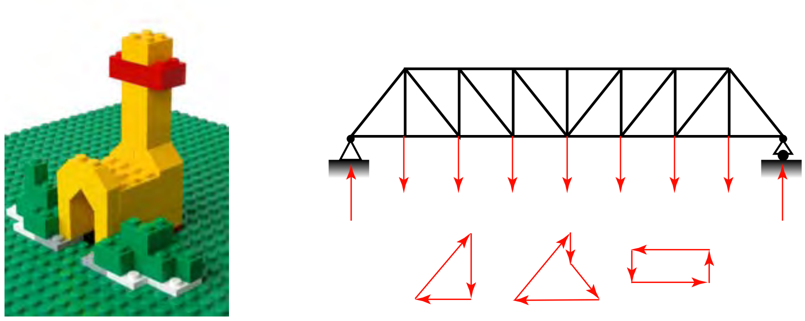
(c) Visualisation of a figure/ground interpretation with focus on dark birds

Fig. 5. Interpretation of Escher's Swans. Figures 5(b) and 5(c) are visualisations only: they are not intended to indicate what is represented as percepts or concepts. Figure 5(a) is reproduced from [17].

of scribbles? Is Figure 6(c) a sketch of a site layout or an interior layout? Maybe it really is just a set of scribbles? Why is Figure 7(a) a representation of a church, not just a pile of blocks? In each case our perception depends just as much on background knowledge and expectations as on sense-data. All four may be considered to be external representations, and all are clearly abstractions. For designers such abstractions maintain the interactive nature of external representations while still being conceptually flexible and "unrestrictive to thinking processes" [18] so as to encourage reinterpretation. External representations need not be so abstract as Figure 6, though. Figure 7(b) shows an engineering representation of Figure 2(a) generated for the purpose of determining truss member forces (the three vector diagrams graphically resolve reactions in the three left-most joints). While still obviously an abstraction of the thing depicted in Figure 2(a), Figure 7(b) has an obvious interpretation to anyone with the appropriate domain knowledge.

(a) Sketch (b) Sketch (c) Sketch

Fig. 6. Examples of design interpretation. Figure 6(a) is reproduced from [19]. Figures 6(b) and 6(c) are reproduced from [18].



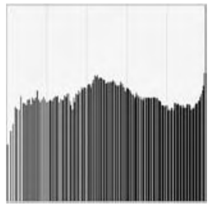
(a) Lego model (b) One representation of a bridge

Fig. 7. More examples of design interpretation

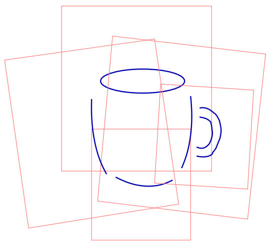
Figure 8(a) shows a cup that is a thing located in the environment. For the example assume that the agent is designing new cups. Figure 8(b) is of a sensor that does edge detection and Figure 8(c) is of a sensor that does colour detection. Perception will have expectations that are characteristic aspects of "cup-ness". The example in Figure 8(d) shows such expectations in the manner of Nelson and Selinger's [20] cubist metaphor. That is, that interpreting the category of things that are cups can be approached not as locating a single, homogeneous object but as a small number of key properties with local context that are assembled in a loose global context. The sense-data of 8(b) and 8(c) drive perception bottom-up but at the same time it relies on biases and expectations are such as shown in Figure 8(d). Figure 8(e) shows these same expectations being used in the reverse direction by actions to effect a sketch of a new cup, which if the same expectations hold should also be interpreted with “cup-ness”.

(a) A cup in view of a designer

(b) Edge detection sense-data from cup



(c) Colour histogram sense-data from cup



(d) Expectations of “cup-ness”

(e) Sketch of cup

Fig. 8. Perception of a cup . Figure 8(d) is after [20]

As an example of applying memory, consider a game. Adults reading this paper will find Figure 9 to be quite easy; most artificial agents will not. The four parts of Figure 9 are separated by enough whitespace that they can be interpreted as being four distinct entities. Given this interpretation, the “one of these things is not like the others” game needs a concept that fits three of the parts of Figure 9 but that does not fit the fourth.

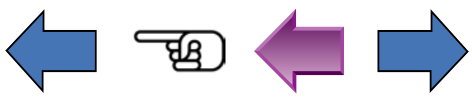
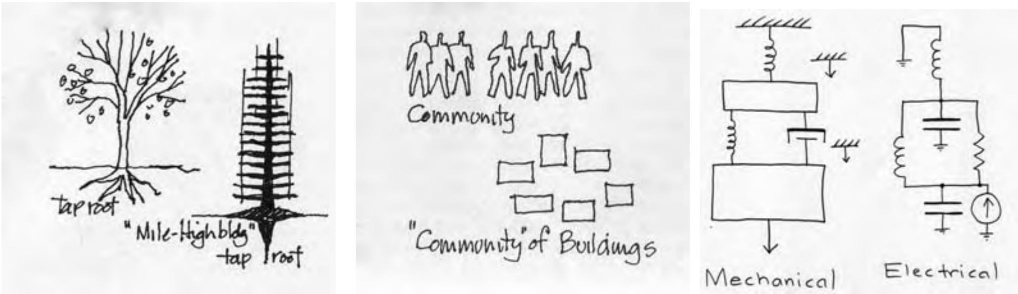


Fig. 9. Which of these things is not like the others? The official NIEHS [21] answer is number 4 (counting from the left) because it points to the right.

For Figure 9 the NIEHS answer partitions the relevant concepts into those interpretations with properties *pointing* and *left* versus those interpretations with properties *pointing* and *right*. This game is analogical, but the NIEHS answer considers only one possible partitioning the four. The validity of the



(a) Tree and building (b) People and buildings (c) Mechanical and electrical artefacts

Fig. 10. More analogies. Figures 10(a) and 10(b) are reproduced from [18].

answer depends firstly on the kinds of sensor of the agent that is playing the game, and secondly on the experiences (and hence memories) of that agent. An agent that has been focusing on edge detection perception of polygons may, in that situation, rationally find that *pointing* and $left$ are less important than the shapes that are bounded by the detected edges; that number 2 is the odd one out.

Analogies come in many forms. For example, complete the sequences ⟨*cow*, *moo*, *duck*, ?⟩ and ⟨*bicycle*, *ride*, *car*, ?⟩. Metaphors are analogies that require inference on what the properties may represent rather than a simpler analogical matching. The phrase[1] "love of money is the root of all evil" involves concepts reconstructed from "money", "root" and "evil", but having these constructs still has us a long way from understanding the metaphor. Understanding it requires additional inference.

Consider extending the game into a more advanced version where additional inference is required: some small amount of reasoning about ways that separate images may be interpreted as being similar. In Figure 10(a) the annotation indicates one way that the two depicted entities may be similar: that the tap root of the tree is similar to the support of the building. This could be in terms of structure (descends underground below the entity) or behaviour (supports the entity to stop it from falling over). Looking at the left hand sketch in Figure 10(a) results in a perceptual interpretation, and hence concepts, of what a tap root structure and behaviour are. A query on this concept results in properties appropriate to that sketched tap root. But some of these properties apply analogically to buildings also, so we have an analogy from one perceptual interpretation to another. Also, having found the analogy may cause others to be recollected such as between the tree trunk and the structures around the elevator shaft.

Figure 10(b) shows an analogy from a concept (that of community) to drawing actions that cluster buildings and hence to expectations for perceptual inter-

[1] The phrase is from I Timothy 6:10, King James Bible

pretation. Figure 10(c) visualises a conceptual analogy from one representation of a mechanical artefact (as a mass-spring-damper system) to an equivalent electrical representation (because the behaviours of the two have the same differential equations and their components have analogous behaviours).

Analogy is important to designing [22] and to creativity [23], and one approach to creative problem solving is to deliberately use analogy as a route to lateral thinking. But constructive memory means that this need not always be deliberate. False recognition is where a person claims that a novel word or event is familiar, and an intrusion is the production of novel information [24]. That is, false recognition is a false memory on the perception side of the agent and intrusion is a false memory on the action side. But to creative designers these effects may not be disadvantageous. Perhaps some part of lateral thinking is the reconstruction of memories "incorrectly"?

4 The constructed environment

Sensors have structure, and that structure influences what is sensed. If a robot senses temperature via an analog thermocouple and a A/D converter, the physical construction of the thermocouple and the resolution and sampling rate of the A/D converter will influence what stream of bits arrive. Data from a sonar sensor or an audio microphone will at any time have temporal correlations to immediately prior sensed data. Data from a vision sensor will at any time have both temporal and spatial correlations to other sensed data. If the brain of the student from Section 3 sends a signal to a muscle to move their arm (an effector signal) and if that arm moved then causality in the environment would likely result in the instructor sensing that movement. The sensed data cannot be completely random or there could be no causality in the environment, no structure in the sensor, and an agent would not be able to interpret such data other than as noise.

So some sensed data must correlate with some other prior sensed data. But that does not imply that what is sensed is entity structure in an FBS [2] sense. There may be correlations within data pushed into a sensor but that does not mean that the sensor is just reading off properties of structured entities from within the agent's view. It is agents that are situated and it is agents that individuate the environment. "Reality" to an agent is a subjective construction rather than an objective thing acquired by sensors. An environment could be described in a Gero-and-Fujii manner as a space of not-necessarily-unique and not-necessarily-disjoint exogenous variables. These variables would have some kind of pseudo-structure but that would not be structure as understood by

[2] See [25] for the FBS design ontology.

a design agent inspecting a designed artifact. Structure as is understood by the agent is an interpretation constructed from sense-data and expectations. Behaviours of artifacts are interpretations of changes to interpreted structure.

Put another way, we can certainly say that the environment contains data in some form. We as observers of our environment could apply information-theoretic techniques to model the information of the environment, and the existence of correlations between environment data is what allows us to view that information as anything other than noise. But the information itself is only an artifact of our modelling - it isn't in the environment.

An observation commonly attributed to Einstein [3] is that insanity is doing the same thing in the same way but expecting a different outcome. Well, maybe: it depends on what "the same" means. An agent only needs partial awareness of the environment to act on it. There may actually be important environmental differences that the agent is not aware of. Secondly, it is not only the agent's interpretation of the environment that is constructed. Memories are also constructed. We can understand this by contrasting it to the usual problem in machine learning. Usually there is some underlying target function that the learner is trying to learn a model of. A reinforcement learner learns a Markov decision process model from environmental interaction while undertaking a particular task. So what gets learned is a model of the particular task undertaken in a particular environment and context.

But there is a problem when we are concerned with constructing memories in the manner suggested. The problem isn't whether the learning algorithm is online or offline, it isn't whether concepts are learned in a supervised or unsupervised fashion, and it isn't whether the learning algorithm is incremental or non-incremental. Markov decision process models have been adapted to new tasks but that isn't our problem either. Rather, it is something like the reverse. The problem with modelling learning in this way is having to learn a concept that will be forever preserved, with the agent adapting it to fit a new task, which in turn will be forever preserved. What happens if the agent learns to represent the intension of a concept X in terms of an existing concept or property A, and at some later time its understanding of A changes? Surely that implies that the concept X also changes? Another story will illustrate. One of the authors of this paper once worked on large military and civil computer systems such as flight simulators. The customer of one project had an unusual requirement. Since they were making milestone payments, the project had to backup and document work sufficiently that at any time the system (software and hardware) could be restored to be as it was at the time of a desired

[3] http://www.quotationspage.com

previous milestone [4]. This was very difficult because the more that the system incrementally developed, the less likely it was that an old test rig or algorithm or piece of hardware would work precisely as it had previously. It wasn't that what an algorithm, for example, did previously was wrong. It is just that that the algorithm was operating under different circumstances previously to now and so when it was called, and what arguments it was being called with, had changed.

What we want of memory in a situated agent is for experiences to be guided in familiar ways. This temporally twofold:

- Projecting active experiences into the future:
  "experience in its vital form is experiental, an effort to change the given; it is characterized by projection, by reaching forward into the unknown", Dewey [2].
- Recognising having previously had a similar experience:
  "reference of its present image to some past reality. In memory we recognize its presence; i.e., we know that it has been a previous element of our experience", Dewey [26].

How that projection or recognition occurs depends on what the agent now knows. We want to allow it to continue to learn from experiences throughout its life. Memories of the agent should not be static constructs that are to be preserved verbatim forever. Rather, subsequent experiences of the agent should structure recollections of past ones.

5 A Metaphor for a Constructive Memory

Section 2 contain descriptions that we would still largely support were it not for our concern of readers from a range of backgrounds fixing on differing understandings of these, and so interpreting those descriptions, to the authors. To this end we sketch here an alternative, metaphorical view of reasoning by a situated agent. This metaphorical view is intended to encourage thinking about the problem differently. We don't intend for it to be implemented as described. With a different mindset established, Section 6 will briefly introduce a new model that will be detailed in [27].

To begin the metaphor, look at Figure 11. It is a photograph of two identical mesh strainers nested one on top of the other. A moiré interference pattern appears if the strainers are slightly out of alignment. The perceived moiré pat-

[4] The objective was not to actually do this; the project just had to be able to demonstrate it could be done if required.

tern depends on the mesh geometries, the relative positions and orientations of the strainers, photographic details such as aperture, printing details (try comparing this viewed electronically on-screen to printed), and perception by the viewer.

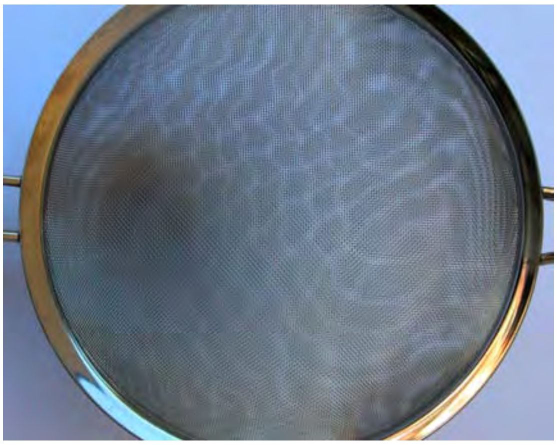
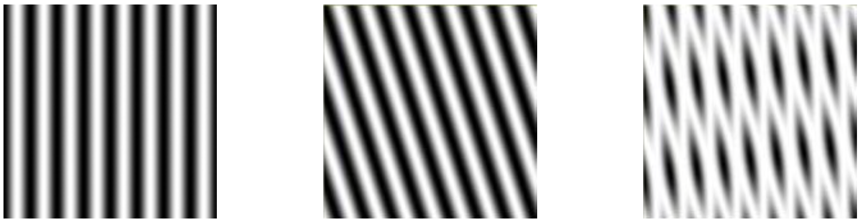


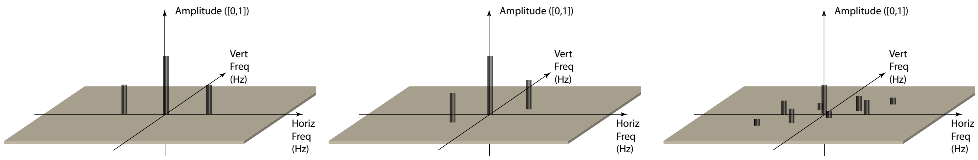
Fig. 11. Photograph of nested strainers.

This kind of interference is also evident in other domains. The variations in air pressure caused by a tuning fork, and the corresponding electrical signal in a microphone or human ear, are time-varying sinusoidal signals (assuming for this example a perfect tuning fork and microphone). If a second source is added nearby to the first, the resulting composite is not simply a pair of separate signals. A new signal is introduced by their composition. For a pair of sinusoidal signals, it is also a sinusoid with a frequency equal to the difference between the original two. Musicians would recognise the composite sound as containing the "beats" that are used when tuning.

Examples like this satisfy the Superposition Theorem because the result is a linear combination of the components, but many examples are nonlinear. Nonlinear systems provide for the possibility of emergence of behaviours from interactions between the components. Figure 12 shows an example by Amidror [28] of an image or light source over which filters are placed. Figures 12(a) and 12(b) show two filters used individually and Figure 12(c) shows both used together. Moiré effects are evident; the reason for their appearance is as follows. Any periodic (or finite aperiodic) sequence has an associated spectrum that can be obtained via a Fourier transform. The sequence may be a time-varying signal, a spatial distribution such as an image (a sequence periodic in 2D, requiring a 2D Fourier transform), or a vector (such as in the convolution-based memories of [29]). If the system is linear and additive then the composite spectrum is the addition of the component spectra. But Figure 12(c) is not additive. Rather, the component transmissions are multiplicative [28]. The example uses cosine filters so as to visually simplify their spectra. The spectrum of a cosine wave is is a pair of impulses for the *cos* term and one impulse for the DC offset. Multiplication in the sequence (time/space) domain is equivalent to convolution in the frequency domain, so if $F_1(u,v)$ and $F_2(u,v)$ are the spectra of Figures 12(d) and 12(e) respectively then $F_3(u,v) = F_1(u,v) \circledast F_2(u,v)$ where $\circledast$ denotes convolution. The resulting $F_3(u,v)$ spectrum of Figure 12(f) contains $3 * 3 = 9$ impulses, not 5. This nonlinear composition introduces new artefact signals and it is these that we perceive as the moiré effects. It is a case of the whole being greater than the sum of the parts - not only does the result contain the original parts but it also contains additional parts that arise from their interaction.



(a) Vertical filter (b) Angled filter (c) Composed filters



(d) 2D spectrum of vertical filter (e) 2D spectrum of angled filter (f) Convolved 2D spectrum

Fig. 12. Example from [28]. Figure 12(a) shows a monochrome filter with transmittance $(\frac{1}{2}cos(2\pi fx) + \frac{1}{2})$, where x is horizontal distance, y is vertical distance and f is band frequency. A transmittance of 1 means full transmission of light (white) and 0 means no transmission (black). Figure 12(b) is Figure 12(a) rotated slightly, or $(\frac{1}{2}cos(2\pi f(xcos(\theta) + ysin(\theta)) + \frac{1}{2})$, and Figure 12(e) is its spectrum. Figure 12(c) shows the moiré effects that are produced when the two filters are overlayed, and Figure 12(f) is its spectrum.

An agent interacting with such signals will sample them at some rate. The sampling rate must be not less than twice that of the highest frequency component. Failure to sample at a sufficiently high rate results in the loss of part of the signal, a phenomenon usually called "aliasing" but which is also interference. Think of a digital photograph or scan of a repeating pattern like a fence. Aliasing may result if the resolution is too low relative to the size of the repeating element. This is why moiré patterns sometimes appear in photographs and scans; it is because of too low a sampling rate (resolution). Think of a spot painted on a rotating wheel, an agent's sensor that strobes a light at some rate, and at each strobe measures the position of the spot. The complexity of the resulting sense-data stream depends both on the behaviour (angular velocity) of the wheel and on the behaviour (strobe rate) of the sensor. It is not an intrinsic property of the physical wheel: "it emerges from the interaction of system state dynamics and measurement as established by an observer" [30]. We could label these phenomena as "modulation" (as in AM radio transmission) or "interference" (as in aliasing) depending on whether we desired them or not.

We can use these ideas to metaphorically describe how a constructive memory works for a situated design agent. Consider an agent doing routine designing and interacting with an external representation of a design. The agent will have existing memories learned from previous experiences, illustrated by Figure 13. These memories are patterns in a space of notions. In Figure 13 the memory is shown as the pattern, and the scope of the space is indicated by the circular boundary.

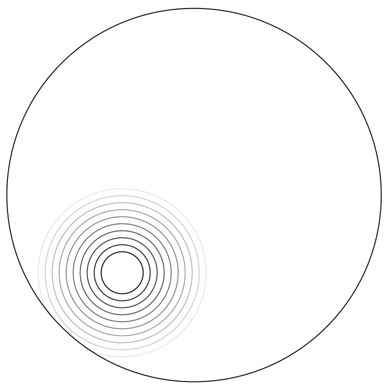
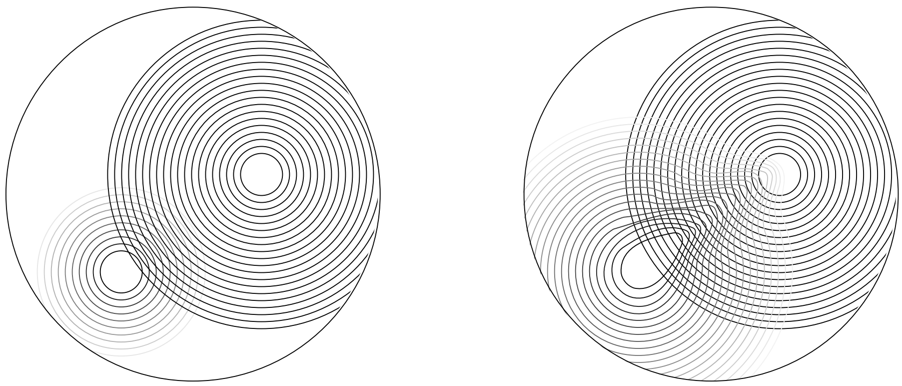


Fig. 13. Illustration of a memory

The pattern looks vaguely like a stone dropped into a pond, right? An experience is something occurring in an agent and so is a process. It is like ripples spreading and decaying in the pond. But the ripples need a medium, and that medium is the pond. For experiences the medium is memory. If I say "I am experiencing fear again" I mean two things. firstly, there is a process going on now and the way I describe to you what that process feels like is as "fear". Secondly, I have experienced fear before and I am using in the present something drawn from an experience that occurred in the past. so "experience" is used in two ways: for an ongoing process now, and for what happens when we use the past in the present. In Section 6 we unify these by denoting both as an experience $\mathfrak{e}_i^j$ of agent $\mathfrak{a}_i$ over temporal extent (t_1, t_2). If (t_1, t_2) includes now then it refers to an ongoing process. If (t_1, t_2) is in the past it refers to how that past experience affects experiences now.

What we don't mean by this is that all experiences get stored verbatim, case-based reasoning style, for later retrieval. What gets stored between the past and now are changes to memories, with multiple experiences possibly affecting multiple memories and vice versa, from which the past experience is reconstructed now. An experience isn't just called up - a current experience needs to construct it. It needs to be made useful now. Something from memory isn't an experience unless it gets used. Memories are how the agent's constructs adapt to experiences, so constructs for perception and the like, and so on each have state. Different states are like different patterns, so absorbing new experiences is like learning to recognise new patterns. Constructs can contain other constructs, so memories can be associated with others that are more abstract, less abstract, or at a similar level of abstraction. Examples include perceptor constructs as a set of pattern recognisers for perception, conceptor constructs as semantic net associations of concepts or as recognisers of serial order patterns, and action activator constructs as an action behaviour network.

The ripples in the pond don't only depend on the stone, though. They also depend on the structure and behaviour of the pond bed and bank, obstacles in the pond, the wind, and the water. Likewise, the situation at the time that a memory is absorbed influences how it gets absorbed and how it is recollected. This is shown in Figure 14(a) as an interference pattern because the current situation influences reconstruction. Because the situation will be different at the time of recollecting to the time of initially remembering, the same experience may later result in a different recollection, both because the memory itself may have changed and because the situation will be different. It is a little like pulling on a spider web: what comes out isn't just that piece of the web but lots of other interconnected threads. Recalling an experience itself results in changes to memory because expectations change, because different parts of the memory are active after the query than before, and because the memory space adapts with the experience. This is illustrated by Figure 14(b).

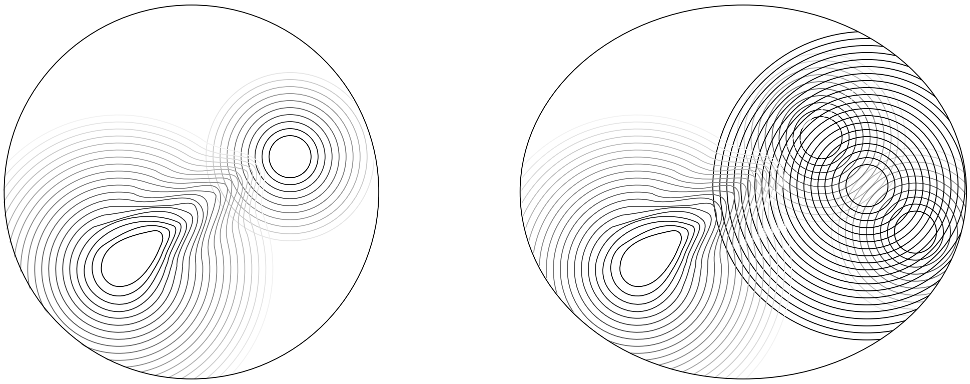


(a) The current situation as an influence on recollection

(b) Memory adapted to the situation

Fig. 14. Adapting to the situation

Our agent now gets a new interactive experience by activating an action and receiving in response new sense-data to be interpreted. Figure 15(a) shows the memory after absorbing another experience. Each experience has a region of the memory space where it dominates. There are also regions where they influence each other.



(a) The memory after a new experience (b) Memory space after creative designing

Fig. 15. Further adaption

Figure 15(a) is from an experience obtained during routine designing. Suppose that the agent undertakes a creative designing task. This entails extending the design space by introducing new properties and by changing the subspaces described by some existing properties. Consequently, if the agent learns something new then bounds of the space of relevant constructs will enlarge, and new memories will result. Figure 15(b) shows one such enlarged space.

The space may have changed due to there being new properties, or because memories were expressed in terms of prior properties which have since been re-interpreted. We can compare this to a machine learning task like learning a decision tree. In both cases, consider what should happen if that which the agent understands a notion to be changes at a later time? Does the change invalidate the decision tree in strict machine learning terms? Probably, yes. But we would say that how the agent recollects the knowledge described by the decision tree changes with new experiences and with the situation, and that is desirable that this be the case. For example, as perception is abductive, sense-data received now may be therefore interpreted differently than previously and different actions may be selected in similar situations than would have been the case previously. If this dependence of reconstructions in the current situation and on later experiences afterwards leads to an agent making a prediction that it finds to be invalid, then the agent is free to later adapt the classifier. But "invalid" is situated and subjective on the part of the agent. The classifier would only need to change if the agent detected errors with respect to its own expectations of the effects of actions. Some prior experiences will thereafter be recollected differently, and a more recent experience is more likely to be recollected than an old one due to changes in memory spaces and more recent situations and experiences adapting particular memory subspaces.

So in Figure 15(b) the pattern in the top right has moved due to the spaces of some prior properties being now understood differently. The pattern in the

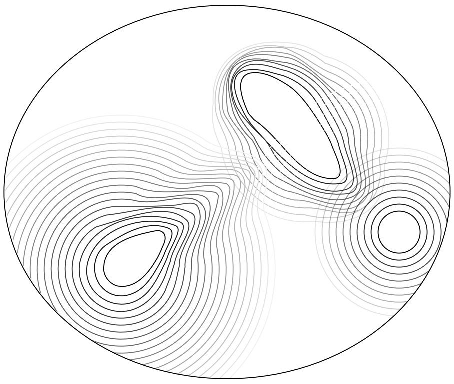


Fig. 16. Adapted memory of the agent

lower right is the new memory that caused the change to the space. Overlayed is the situation at the current time. It is again shown as an interference pattern because the current situation influences how the memory is constructed and absorbed. Figure 16 shows the resulting memory patterns.

6 Overview of an experiential model

Our current view of constructive memory was inspired by two phrases. The first is Clancey [32] paraphrasing Dewey:

> "Sequences of acts are composed such that subsequent experiences categorise and hence give meaning to what was experienced before".

The second is by Bartlett [14]:

> "Remembering is not the re-excitation of innumerable fixed, lifeless and fragmentary traces. It is an imaginative reconstruction, or construction, built out of the relation of our attitude towards a whole active mass of organised past reactions or experience, and to a little outstanding detail which commonly appears in image or in language form."

This view of situated designing starts with experiences. An experience is not of a disembodied agent. It is to do with interaction of the agent with an environment ξ. An experience is also not something static; it is dynamic and is of an agent coupled to its environment. So we say that there are agent-entities (agents) denoted by $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_N$ and thing-entities (non-agent things) denoted by $\mathfrak{t}_1, \mathfrak{t}_2, \ldots, \mathfrak{t}_M$, where N is the number of agents and M the number of things. An agent $\mathfrak{a}_i$ is composed of an agent-thing $\mathfrak{t}_i$ (where this $\mathfrak{t}_i$ is a part of ξ and is what in humans is the body) and some number of construct-entities (constructs) $\mathfrak{c}_i^1, \mathfrak{c}_i^2, \ldots, \mathfrak{c}_i^N$ (that in humans is the nervous system), where N is the number of constructs of $\mathfrak{a}_i$.

A coupling between $(\xi - \mathfrak{a}_i)$ and an agent-thing $\mathfrak{t}_i$ is an e-experience (an exogenously generated experience) of $\mathfrak{a}_i$. An example is robot navigation experiences involving sonar sensual experiences and motion effectual experiences. A coupling between the agent-thing $\mathfrak{t}_i$ and agent constructs $\{\mathfrak{c}_i^j\}$, or within $\{\mathfrak{c}_i^j\}$, is an a-experience (an autogenously generated experience) of $\mathfrak{a}_i$. An example is a human moving their arm, involving sensual experiences of proprioception and motor effectual experiences. Further, e-experiences and a-experiences may perturb each other directly or indirectly. An e-experience perturbs an a-experience if the agent interprets that e-experience. An e-experience fails to perturb an a-experience if the agent ignores it. An a-experience perturbs an e-experience when the agent acts on its environment – when that agent perturbs other agents or things.

We denote experiences of agent $\mathfrak{a}_i$ as $\mathfrak{e}_i^1, \mathfrak{e}_i^2, \ldots, \mathfrak{e}_i^N$, where N is the number of experiences of $\mathfrak{a}_i$. If an e-experience is able to perturb an a-experience and vice versa, an agent must be able to have multiple concurrent experiences . An e-experience involves entities perturbing each other but where one of the entities is $\mathfrak{t}_i$ of $\mathfrak{a}_i$ and the other is either another agent $\mathfrak{a}_j$, $j \neq i$, or a thing $\mathfrak{t}_m$. An a-experience also involves entities perturbing each other, but where both are part of $\mathfrak{a}_i$. Experiences and notions describe agents at a logical level. They are supervenient or emergent on constructs, which describe agents at a physical level. The "memory" is the set of constructs with their states, and these support experiences. The future of experience $\mathfrak{e}_i^x$ at some time t will depend on

- $\mathfrak{e}_i^x$ up until t – we also call this the trajectory of $\mathfrak{e}_i^x$ to t,
- What perturbations of $\mathfrak{e}_i^x$ there have been up until time t – we call this the history $hist(\mathfrak{e}_i^x)(-\infty, t) \sqsubseteq \mathfrak{e}_i^x$ until t, and
- The situation at t.

An experience can be a part of another experience. We write $\mathfrak{e}^x \sqsubseteq \mathfrak{e}^y$ to mean "experience $\mathfrak{e}^x$ is a part of experience $\mathfrak{e}^y$", with "part of" taken to be Seibt's [33] process mereology part of relation. We write $\mathfrak{e}^x \bigcirc \mathfrak{e}^y$ to mean "experience $\mathfrak{e}^x$ overlaps experience $\mathfrak{e}^y$", with "overlap" taken to be Seibt's process mereology overlap relation [5] . Given these we say that when one experience perturbs another, or $\mathfrak{e}^x \rhd \mathfrak{e}^y$, there will be some overlap between them.

An important idea from the Bartlett quote is that recollections of earlier experiences may vary. Projections into the past and future are not against

[5] Actually Seibt wouldn't use overlap here as she deals with "free" processes (like the abstract process of running) which become concrete when they "interfere" with spacetime. Free process interfering is more ontology than we need here so we ignore the distinction.

recordings of experiences, they are against reconstructions of experiences. Let situations of $\mathfrak{a}_i$ be $\mathfrak{s}_i^1, \mathfrak{s}_i^2, \ldots \in \mathfrak{S}_i$ and let $recall(\mathfrak{s}_i^j, \mathfrak{e}_i^x, \mathfrak{e}_i^y, t)$ mean that recalled experience $\mathfrak{e}_i^y$ is familiar at time t to $\mathfrak{e}_i^x$ in situation $\mathfrak{s}_i^j$. That is, $\mathfrak{e}_i^y$ from some time $t' < t$ is similar in some way to $\mathfrak{e}_i^x$ at the current time t, given situation is $\mathfrak{s}_i^j$. Now $\mathfrak{e}_i^x$ and $\mathfrak{e}_i^y$ have different temporal extents (projecting into the future, recognising similar pasts) so let $aproj_i(\mathfrak{s}_i^j, t, \mathfrak{e}_i^k)$ be $\mathfrak{e}_i^k$ projected with the same time scale but with t located at time zero.

$$aproj_i : \mathfrak{S}_i \rightarrow \mathbb{T} \rightarrow \mathfrak{E}_i \rightarrow \mathfrak{E}_i$$

We use "projected" to mean predicted or devised, not simply as a mathematical transformation. The projection $aproj$ is affected by the age of the recalled experience: the older the original experience is, the less likely it is that it should be recalled. It is also projected with respect to the current situation.

We need to be able to judge the similarity of two projected experiences. [34] measure process equivalence is by finding those firing sequences of a Petri net model M_1 that also appear in a Petri net model M_2. In that light we define similarity of two experiences as that part of one experience that is recollected from another experience

$$similarity(\mathfrak{s}_i^j, t, \mathfrak{e}^{hx}, \mathfrak{e}^{hy}) = \frac{|\mathfrak{e}^{hx} * \mathfrak{e}^{hy}|}{|\mathfrak{e}^{hx}|}$$

where $\mathfrak{e}^{hx} = hist(aproj(\mathfrak{s}_i^j, t, \mathfrak{e}_i^x))(0)$ and $\mathfrak{e}^{hy} = hist(aproj(\mathfrak{s}_i^j, t', \mathfrak{e}_i^y))(0)$ are projected histories of the experiences, $|\mathfrak{e}^{hx}|$ is some quantitative measure of the extent of $\mathfrak{e}^{hx}$. The $*$ is the mereology product relation of Seibt [33], where the product of $\mathfrak{e}^x$ and $\mathfrak{e}^y$ is all overlapping parts of $\mathfrak{e}^x$ and $\mathfrak{e}^y$.

$$\mathfrak{e}^z = \mathfrak{e}^x * \mathfrak{e}^y \iff \forall \mathfrak{e}^w \bullet (\mathfrak{e}^w \sqsubseteq \mathfrak{e}^x \wedge \mathfrak{e}^w \sqsubseteq \mathfrak{e}^y \Rightarrow \mathfrak{e}^w \sqsubseteq \mathfrak{e}^z)$$

Let $recall(\mathfrak{s}_i^j, \mathfrak{e}_i^x, \mathfrak{e}_i^y, t)$ be true if $\mathfrak{e}_i^x$ at the current time t is similar to $\mathfrak{e}_i^y$ at some prior time and there are no other more similar experiences that can be recalled.

$$recall(\mathfrak{s}_i^j, \mathfrak{e}_i^x, \mathfrak{e}_i^y, t) \Rightarrow \underset{\mathfrak{e}_i^y, t'}{\operatorname{argmax}} \left(\frac{|\mathfrak{e}^{hx} * \mathfrak{e}^{hy}|}{|\mathfrak{e}^{hx}|} \right)$$

with $\mathfrak{e}^{hx}$ and $\mathfrak{e}^{hy}$ as above. Notice that similarity is with respect to the current situation and the definition of $|\mathfrak{e}|$. Recall isn't looking up experiences in storage somewhere; it is reconstructing past experiences in the current situation. If a function of memory is to guide an experience in familiar ways, but the agent has adapted to other experiences since that familiar experience was active, what the agent remembers at the later time may not be what was originally the case. One reason is that the agent adapts to subsequent experiences, hence recollections of what was experienced before may be different.

$aproj$ isn't simply timeshifting but is reconstructing at $t \in \mathbb{T}$ according to $\mathfrak{s}_i^j, \ldots \in \mathfrak{S}_i$. Another reason is that memories of experiences are dynamic and interlinked, so a recollection is a reconstruction rather than a lookup: imperfect recollection accompanied by filling-in of what is missing, and the spider web. A third reason is that the agent adapts to subsequent experiences, hence some recollections of what was experienced before may be interpreted differently. Finally, experiences are subject to the current situation: $aproj_i$ depends on $\mathfrak{s}_i^j$.

Dewey characterised experiences in terms of continuity and interaction. For continuity something persists and for interaction something changes. So experiences have a temporal aspect, but they cannot be solely temporal. Suppose we want to look closely at an experience and see what this "something" of experiences is, so we fixate on an experience at a particular time. This fixation is a function from experiences onto some subspace that we denote as $\mathfrak{N}_i$. This subspace is in the space of experiences, and each contains "somethings" that have meaning to agent $\mathfrak{a}_i$. So we call the subspace $\mathfrak{N}_i$ the space of notions of the agent. We use the word "notion" to maintain independence from any particular kind of agent representation. For convenience we write $\mathfrak{e}_i^x(t), t \in \mathbb{T}$, for that particular subspace of $\mathfrak{N}_i$ that was what $\mathfrak{e}_i^x$ was like at time $t \in \mathbb{T}$. The concepts $\mathcal{C}_i$, percepts $\mathcal{P}_i$, acts $\mathcal{A}_i$, sense-data $\mathcal{S}_i$ and effect-data $\mathcal{E}_i$ of an agent are all subspaces of $\mathfrak{N}_i$.

Maher and Gero [3] described three kinds of agent reasoning: $reflexive$, $reactive$ and $reflective$ but here we start with two more kinds that we shall call $senseless$ and $effectless$, corresponding to the Rosenschein and Kaelbling [35] "pure action" and "pure perception" agents. For entity $\mathfrak{o}$,

$$\begin{aligned}
&senseless(\mathfrak{o}) \iff \exists i \bullet \mathfrak{a}_i = \mathfrak{o} \wedge \mathcal{S}_i = \emptyset \\
&effectless(\mathfrak{o}) \iff \exists i \bullet \mathfrak{a}_i = \mathfrak{o} \wedge \mathcal{E}_i = \emptyset \\
&interactive(\mathfrak{o}) \iff (\exists i \bullet \mathfrak{a}_i = \mathfrak{o}) \wedge \neg senseless(\mathfrak{o}) \wedge \neg effectless(\mathfrak{o})
\end{aligned}$$

For $\mathfrak{o}$ to be an agent it must be $interactive$ and either (i) have at least one agent experience perturbed by at least one external process, or (ii) have at least one external process perturbed by at least one agent process. The conditions in the disjunction are satisfied when the agent is $reflexive$, $reactive$ or $reflective$.

$$\begin{aligned}
isagent(\mathfrak{o}) \Rightarrow\ & interactive(\mathfrak{o}) \wedge \\
& (reflexive(\mathfrak{o}) \vee reactive(\mathfrak{o}) \vee reflective(\mathfrak{o}))
\end{aligned}$$

$reflexive$ agents are $interactive$ with at least one effection perturbed by at least one sensation. Effection and sensation are of effectors and sensors respectively, so the condition holds when the experiences of an effector without any sensors are different to the experiences with at least one sensor present. That is, when there is at least one sensor that perturbs at least one effector.

$$\begin{aligned}
reflexive(\mathfrak{o}) \iff\ & interactive(\mathfrak{o}) \wedge \\
& \exists i, \mathfrak{e}_i^x, \mathfrak{e}_i^y, t \bullet \mathfrak{a}_i = \mathfrak{o} \wedge \\
& \qquad \mathfrak{e}_i^x(t) \subseteq \mathcal{E}_i \wedge \mathfrak{e}_i^y(t) \subseteq \mathcal{S}_i \wedge \\
& \qquad \mathfrak{e}_i^x \rhd \mathfrak{e}_i^y
\end{aligned}$$

reactive agents are *interactive* with at least one one action perturbed by at least one perception. The reason that *reactive* is not described as "$reflexive$ where ..." is that an agent may be both reflexive and reactive, or it may be reactive but not reflexive. Similarly, a reflective agent may or may not be reactive and it may or may not be reflexive. Why is an agent with percepts but not acts called reflexive? Because it is what drives the effectors that matters. Why have percepts in a reflexive agent? To determine the current situation.

$$\begin{aligned}
reactive(\mathfrak{o}) \iff\ & interactive(\mathfrak{o}) \wedge \\
& \exists i, \mathfrak{e}_i^x, \mathfrak{e}_i^y, t \bullet \mathfrak{a}_i = \mathfrak{o} \wedge \\
& \qquad \mathcal{A}_i \neq \emptyset \wedge \\
& \qquad \mathfrak{e}_i^x(t) \subseteq \mathcal{A}_i \wedge \mathfrak{e}_i^y(t) \subseteq \mathcal{P}_i \wedge \\
& \qquad \mathfrak{e}_i^x \rhd \mathfrak{e}_i^y
\end{aligned}$$

reflective agents are *interactive* with at least one action perturbed by at least one conception.

$$\begin{aligned}
reflective(\mathfrak{o}) \iff\ & interactive(\mathfrak{o}) \wedge \\
& \exists i, \mathfrak{e}_i^x, \mathfrak{e}_i^y, t \bullet \mathfrak{a}_i = \mathfrak{o} \wedge \\
& \qquad \mathcal{A}_i \neq \emptyset \wedge \mathcal{C}_i \neq \emptyset \wedge \\
& \qquad \mathfrak{e}_i^x(t) \subseteq \mathcal{A}_i \wedge \mathfrak{e}_i^y(t) \subseteq \mathcal{C}_i \wedge \\
& \qquad \mathfrak{e}_i^x \rhd \mathfrak{e}_i^y
\end{aligned}$$

Reflective construction of particular notions may eventually lead to reactive reasoning over similar notions if that refection recurs sufficiently. This is non-associative adaption to continuity of experiences, and it is often called habituation. Reactive reasoning may similarly lead to new reflexive reasoning.

7 Conclusions

In this paper we began the process of documenting knowledge collected from many years of research into agents and situated designing. We revisited the issues canvassed in Gero and Fujii [1] by reconsidering the notion of what constitutes situated designing. This has brought further insight into the interplay between situated interaction, constructive memory, the role of experience, and our understanding and description of what a constructed environment is. We continue this work in [27], where we consider the kind of memory an agent of the kind described should have. It describes what is required of a constructive memory system that could be used to build what woul be the basis of ongoing efforts to build situated design agents and intelligent design tools.

References

[1] J. S. Gero, H. Fujii, A computational framework for concept formation for a situated design agent, Knowledge-based Systems 13 (6) (2000) 361–368.

[2] J. Dewey, The need for a recovery of philosophy, in: Creative Intelligence: Essays in the Pragmatic Attitude, Henry Holt and Company, 1917, pp. 3–69, reprinted in *John Dewey: The Middle Works, 1899-1924*, Vol. 10: 1916-1917, Southern Illinois University, 1985, pp. 3–49.

[3] M. L. Maher, J. S. Gero, Agent models of 3D virtual worlds, in: G. Proctor (Ed.), ACADIA 2002: Thresholds, Pamona, CA, ACADIA, 2002, pp. 127–138.

[4] W. J. Clancey, 'Situated' means coordinating without deliberation, mcDonnell Foundation Conference on 'The Science of Cognition' (1992).

[5] P. E. Agre, Computation and Human Experience, Cambridge University Press, Cambridge, UK, 1997.

[6] A. Kukla, Social Constructivism and the Philosophy of Science, Routledge, 2000.

[7] E. von Glasersfeld, Radical Constructivism: A Way of Knowing and Learning, The Falmer Press, 1995.

[8] I. E. Gordon, Theories of Visual Perception, John Wiley and Sons, 1997.

[9] R. K. Guenther, Human Cognition, Prentive-Hall, 1998.

[10] J. A. Wheeler, At Home in the Universe, American Institute of Physics, 1994.

[11] D. A. Schön, D. Wiggins, Kinds of seeing and their functions in designing, Design Studies 13 (2) (1992) 135–156.

[12] G. J. Smith, J. S. Gero, What does an artificial design agent mean by being 'situated'?, Design Studies 26 (2005) 535–561.

[13] M. Shanahan, Perception as abduction: Turning sensor data into meaningful representation, Cognitive Science 29 (2005) 103–134.

[14] F. C. Bartlett, Remembering: A Study in Experimental and Social Psychology, Cambridge University Press, 1932.

[15] C. Cherniak, Minimal Rationality, MIT Press, 1986.

[16] G. Kanizsa, Organization in Vision: Essays on Gestalt Perception, Praeger, 1979.

[17] M. C. Escher, The Graphic Work of M. C. Escher, Oldebourne, 1967.

[18] P. Laseau, Graphic Thinking for Architects and Designers, 2nd Edition,

John Wiley and Sons, 2001.

[19] N. Crowe, P. Laseau, Visual Notes for Architects and Designers, John Wiley and Sons, 1984.

[20] R. C. Nelson, A. Selinger, A cubist approach to object recognition, in: International Conference on Computer Vision (ICCV98), Bombay, India, 1998, pp. 614–621.

[21] NIEHS, What's wrong with these pictures?, US National Institute of Environmental Health Sciences, `http://www.niehs.nih.gov/kids/wrong.htm` (2005).

[22] L. Qian, J. S. Gero, Function-behaviour-structure and their roles in analogy-based design, Artificial Intelligence in Engineering Design, Analysis and Manufacture 10 (1996) 289–312.

[23] M. Runco, S. Pritzker (Eds.), Encyclopedia of Creativity, Academic Press, 1999.

[24] A. Riegler, Constructive memory, Kybernetes 34 (2005) 89–104.

[25] J. S. Gero, Design prototypes: A knowledge representation schema for design, AI Magazine 11 (4) (1990) 26–36.

[26] J. Dewey, Psychology, Harpers & Brothers, 1887, reprinted in *John Dewey: The Early Works, 1882-1898*, Vol. 2: 1887, Southern Illinois University, 1969, pp. 3–21.

[27] J. S. Gero, G. J. Smith, A framework for a situated design agent: Constructive memory, to submit (2007).

[28] I. Amidror, The Theory of the Moiré Phenomenon, Kluwer Academic, 2000.

[29] T. A. Plate, Analogy retrieval and processing with distributed vector representations, Expert Systems: The Journal of Knowledge Engineering 17 (1) (2000) 29–40.

[30] J. F. Kolen, J. B. Pollack, The observers' paradox: Apparent computational complexity in physical systems, Journal of Experimental and Theoretical Artificial Intelligence 7 (1995) 253–277.

[31] M. Bunge, Ontology II: A World of Systems, Vol. 4 of Treatise on Basic Philosophy, D. Reidel Publishing Company, 1979.

[32] W. J. Clancey, Situated Cognition, Cambridge University Press, 1997.

[33] J. Seibt, Free process theory: Towards a typology of occurrings, Axiomathes 14 (2004) 23–55.

[34] W. M. P. van der Aalst, A. K. A. de Medeiros, A. J. M. M. Weijters, Process equivalence: Comparing two process models based on observed behavior, in: S. Dustdar, J. L. Faideiro, A. Sheth (Eds.), International Conference on Business Process Management (BPM 2006), Vol. 4102 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 129–144.

[35] S. J. Rosenschein, L. P. Kaelbling, A situated view of representation and control, Artificial Intelligence 73 (1995) 149–173.