Working Note A – Part 2

10. Beyond the PPP

Despite its explanatory power, the predictive processing paradigm (PPP) has not yet been brought to yield a clear account of how subjective states can be generated. Further ideas are needed. The following conjectures go to such ideas through building off concepts developed in Part 1 of this note. [130, 131, 132]

10.1 Neural correlates of consciousness

There is considerable evidence to support the proposal that the evolving contents of a person's moment-by-moment subjective perception of themselves as a physical body operating in a physical world correlates closely with the moment-by-moment evolution of network states – akin to those described in Section 9.2.1 as recognition states, Ξ – carried by nodes embedded within a hierarchical processing system[r] with architecture along lines described in Sections 9.1 and 9.2 of Part 1.[133]

This evidence is provided in a large literature supportive of the PPP that relies on the reported subjective perceptual experience of test subjects, including their perception of various types of illusions and other effects, where these can be shown to be well explained by application of the PPP.[134, 135]

Consistent with this proposal, Section 9.3.2 defined W[i], the physical world as a person phenomenally experiences it – and B[i], a person's physical body as they phenomenally experience it (denoted together as {B[i] in W[i]}) – as arising in a person's D[r] processor in the form of substates of an overall recognition state, Ξ.[136]

Drawing on subsequent sections describing the idea of a recognition cycle, and of how hierarchical processing could be implemented in D[r], the idea of a person's moment-by-moment perception of themselves as being a physical body operating in a physical world, was proposed [137] to be a subjectively experienced time evolution of the substates {B[i] in W[i]} as part of the evolution from some arbitrary time, t = a, of an overall recognition state, Ξ, as follows:

$$\{\Xi^{a} + \Xi^{(a+1)*}\},\ \{\Xi^{(a+1)} + \Xi^{(a+2)*}\},\ \{\Xi^{(a+2)} + \Xi^{(a+3)*}\}\ , \ldots ,\ \{\Xi^{(a+n)} + \Xi^{(a+n+1)*}\} \qquad \text{R1}$$

Where R1 – as per E4 – will have as a component an evolution of the substates:

$$\{\{B[i] \text{ in } W[i]\}^{a} + \{B[i] \text{ in } W[i]\}^{(a+1)*}\}, \{\{B[i] \text{ in } W[i]\}^{(a+1)} + \{B[i] \text{ in } W[i]\}^{(a+2)*}\}, \ldots, \{\{B[i] \text{ in } W[i]\}^{(a+n)} + \{B[i] \text{ in } W[i]\}^{(a+n+1)*}\} \qquad \text{R2}$$

However, even if it is considered plausible based on definitions of W[i], B[i] and {B[i] in W[i]} given in Section 9.3.2, to claim that substates such as those described in R2 do arise in a system such as D[r] – and to further claim that those substates can correlate with a person's subjectively experienced perceptions – the definitions provided at Section 9.3.2 together with the proposals outlined in Section 9.4.5, still do not seem to explain to a sufficiently satisfying level of detail how it is that such an ensemble of substates can come to be 'inhabited' by a subjectively experienced first person perspective in the way that you or I consciously experience ourselves on a moment-by-moment basis through having such a perspective.

For this reason, a central challenge in explaining subjective experience becomes one of describing in sufficient detail candidate mechanisms by which such a first person perspective of being a B[i] in a W[i] can arise. This goes to the 'what is it like?' question posed by Nagel,[138] the problem of an 'explanatory gap' identified by Levine,[139] and the related insights of Papineau,[140] which sit well with the approach adopted here.

---

[130] From here on Working Note A – Part 1 will be referred to simply as Part 1.
[131] The numbering of sections, figures and footnotes continues sequentially from those of Part 1.
[132] These conjectures additionally draw upon work first written up in the Essay and elaborated in other material provided at http://teledyne.com/.
[133] See in particular see Figs 3, 4 and 5 of Part 1, and their associated text.
[134] A powerful and indicative sample of such evidence is provided in Chapters 2, 3, and more broadly throughout Clark, A. *Surfing Uncertainty: Prediction, Action and the Embodied Mind* Oxford University Press, New York, USA 2016.
[135] Accordingly, one of the most important achievements of the PPP may be that so much of the experimental and observational work that has gone into supporting it shows that somehow, somewhere, at or near the innermost neurological levels of the brain[r]'s hierarchical processing system, subjective states are being generated. We can know this because – to paraphrase the text above – so often the subjectively experienced contents of those states are highly consistent with certain specific artifacts and illusions only to be expected if the neuronal architecture[r] delivering those artifacts and illusions is an implementation of predictive processing and active inference.
[136] NB: For definitions of key terms and notation used here and in Parts 1 and 3, see https://teledyne.com/working_note_A_appendix_1.pdf.
[137] See Section 9.4.5 of Part 1.
[138] Nagel, T. (1974) What is it Like to be a Bat? *The Philosophical Review 83* 435–450.
[139] Levine, J. (1983) Materialism and Qualia: the Explanatory Gap. *Pacific Philosophical Quarterly 64* 354-361.
[140] Papineau, D. (2011) What Exactly is the Explanatory Gap? *Philosophia 39* 5-19.

10.2 Inferential convergence

To begin to describe such mechanisms it is useful to revisit the hierarchical neurological architecture and processes proposed by Mesulam and integrated with the PPP by Friston and others.[141] Accordingly it is proposed here that the hierarchical processing system within D[r] will use raw inputs[142] – i.e. cueing/error signals – originating in the noumenal body, B[r], and its noumenal environment, W[r]\B[r], by processing these inwards and upwards through a series of cortical layers where, at nodes within each layer, these will meet and either 'verify or correct' outward and downward travelling signals – predictions – already generated based on the processor's *best estimate* of the state of B[r] and W[r]\B[r] that it expects that it is encountering. This probabilistic a priori estimate will be grounded in the processor's past 'learning', which will have been accumulated and realized through implementation of its constituent recognition and generative models. It is further proposed that when – at some given moment – these upward and downward signals are brought into balance through a 'cancelling out' of cueing/error signals in all cortical layers – the state of a network of nodes in the processor – i.e. the state previously defined as Ξ – will encode the processor's best estimate of the momentary state and state trajectory of B[r] in W[r]\B[r] that constitute the 'hidden causes' of input into D[r] in that moment.

Proposals of this general type about how a brain[r] processes incoming information are based on strong experimental evidence extending to an understanding that outermost, unimodal cortical layers operate to match incoming cueing signals against predictions encoding relatively basic inferences about what is likely to be causing the inputs being processed. For example, in these outer layers raw visual inputs will be matched against very general predictions about whether an arrangement of edges and surfaces is making up an expected discrete object or not, and if so then the general size, shape, colour, location and motion of that object. Resolution of the causes of such sensory input may extend in these outer, unimodal layers to 'verifying or correcting' predictions as to whether an object being seen is, for example, a moving blue car or van, while processing in and across deeper layers would be needed to achieve more specific verification/correction as to say the *type* of van or car being seen (e.g. a sedan or sports car; or a passenger or delivery van). Processing in even deeper, heteromodal layers would be needed (1) to identify – and match to the car/van – any sounds that it is making and, across presumably deeper layers, (2) to 'recognise' that what is being seen is in fact a specific a car owned by Fred.

While it is possible to find strong empirical evidence for this type of hierarchical construction based on demonstrating specific deficits in the phenomenal worlds experienced by subjects who have had identifiable damage to relevant outer cortical layers[143] – i.e. for examples of types up to and including (1) above – an understanding of what is happening in the innermost multimodal layers of the hierarchy is more difficult to determine through empirical observation of such correlations. Even so, given the strength of the empirically based understanding described above of the constructive processes taking place as signals are balanced within the hierarchical processor's outer layers, it seems feasible to extrapolate all the way – past such presumptive examples as (2) – to describe the end result that these processes of construction are likely moment-by-moment to be delivering in the processor's innermost multimodal layers.

Through such extrapolation it can be proposed that these processes of construction and 'recognition' will as an end result lead – once the hierarchical processor's generative and recognition models have had sufficient 'life experience' to reach maturity – to serial recognition states, Ξ, that moment-by-moment – beat-by-beat of the recognition cycle – form a *fully relational*, ongoing *mapping* that probabilistically predicts as an *integrated whole* the hidden causes of *all* input entering the processor, D[r]. In other words, that the recognition state Ξ will moment-by-moment reflect a fully relationally integrated mapping of the momentary state and state trajectory of W[r]\D[r][144] where such a mapping will constitute the most *operationally efficient best estimate* – or 'explanation' – that the processor's recognition and generative models can deliver to describe, and to shape through active inference, the hidden causes of ongoing inputs into D[r].[145]

---

[141] As discussed more fully in Sections 9.1 and 9.2 of Part 1.

[142] Where such inputs into the D[r] hierarchical processor will include all exteroceptive, proprioceptive and interoceptive inputs.

[143] For example, selective loss of ability to perceive colour or motion: ref. Mesulam, M.M. (1998) From Sensation to Cognition *Brain 121* 1013-1052.

[144] Here W[r]\D[r] is that part of W[r] that is not D[r]. Note that W[r] and W[r]\D[r] can be expressed in the following ways, each best suited to use in different contexts. Hence W[r] = {D[r] + W[r]\D[r]} = {B[r] + W[r]\B[r]}. Fully expanded, W[r] = {D[r] + B[r]\D[r] + W[r]\(B[r]\D[r])}. These relationships can be seen by viewing Fig. 8 in Part 1 as a Venn diagram. Where the expression 'B[r] in W[r]\B[r]' is used – sometimes shorthanded to 'B[r] in W[r]' – these are abbreviations to keep discussion streamlined, where strictly both mean {B[r] +W[r]\B[r]}\D[r], which in condensed form is W[r]\D[r].

[145] This extrapolation corresponds to a limiting case in relation to the proposal that one of the ways downwards (inter-layer outwards) *and lateral* (intra-layer sideways) propagating signals in brain[r] hierarchical processors balance upwards (inter-layer inwards) propagating signals is to modulate those signals by probabilistically predicting the *context* within which they are being generated (e.g. see Section 4.2 of Friston, K. (2003) Learning and Inference in the Brain *Neural Networks 16* 1325-1352). Here the 'integrated whole' proposed above will be the probabilistically derived end result – encompassing *all* momentary upwards, downwards and lateral balancing across the *whole* hierarchical processor in D[r] – of the *contextualising* aspect of this balancing. The observed subjective result consistent with this claim is that you and I can – *and indeed seem to be bound to* – place everything we consciously experience and consider to be 'real' within the overall context of a *single unified phenomenal world*, in form of our respective W[i]s.

Now consider the extent to which this mapping might proceed if W[r] – the noumenal world – is such that its states and state trajectories are bounded in their variability, where proposing such *bounded variability* is simply to propose that there will be a boundary within which any arising state and state trajectory of W[r] will exist, even though a virtually infinite variety of such states and state trajectories might be possible – i.e. 'allowed' – within that boundary.

W[r] will be bounded in its variability in this way if it is subject to certain laws, and always behaves within the constraints – i.e. within an 'envelope' – defined by those laws. Based on our collective experience of W[z] – particularly through the practice and findings of the physical sciences, which overwhelmingly indicate the existence of physical laws – it can be inferred that W[r] is indeed bounded in its variability.[146]

Now if W[r] is bounded in its variability, and D[r] capacity to develop and implement its generative and recognition models passes a certain level of capability, it can be proposed that D[r]'s moment-by-moment mapping of states and state trajectories of W[r]\D[r] as Ξ, and of the effects that D[r] itself has on those states and state trajectories through expression of active inference, can begin to approach what may be called *completeness*.

This is simply to say that above a certain level of capability – reflected in the processing power of the D[r] neurology[r] implementing hierarchical processing and its affiliated generative and recognition models – it should be possible for D[r] to solve to a degree approaching *virtual completion* – and in the most relationally integrated and efficient way possible for purposes of allowing B[r] to attain goal states[147] – the ongoing processing task of having the state Ξ map the contemporaneous state and state trajectory of W[r]\D[r].

Call achieving such a level of capability in D[r] *inferential convergence*, where – following the progressive refinement of the generative and recognition models in D[r] – the inferentially derived state Ξ converges onto ongoing provision of a maximally complete, maximally relationally integrated, and maximally efficient, if probabilistic, mapping of the state and state trajectory of W[r]\D[r] it is encountering.[148]

This amounts to proposing that a normal human child will be born with sufficiently high innate processing power and related capacities in their D[r] and related B[r] systems that, as their B[r] grows and gains experience – and their D[r]-housed generative and recognition models become ever better tuned through operation of the recognition cycle – their D[r] will, with such experience, begin to generate recognition states Ξ that provide an ongoing, virtually complete, fully relationally integrated, and maximally efficient mapping of the unfolding state and state trajectory it is encountering in W[r]\D[r].

A simplifying, partial metaphor for achievement of such proposed completeness via inferential convergence is to liken it to the gradually improving focusing ability of an infant's eye as its systems learn to deliver ever clearer images onto its retina until maximally clear images are achieved. Here the states Ξ can be considered the image, the hierarchical processor in D[r] the eye, and W[r]\D[r] the source of light impinging upon the eye.

But an eye deals only with sharpening an optical image to a probable best fit, maximally resolved location in a relatively low-dimensional space. It is envisaged here that the hierarchical processor in D[r] will be able to hone its generative and recognition models to the point where, as the child matures, D[r] can deliver a maximally resolved – or 'focused' – location in Ξ space, where Ξ space is likely a massively high-dimensional space able to map as a relationally integrated whole all of the actual and potential states and state trajectories of B[r] in W[r]\B[r][149] that D[r] may encounter as it rides within – and interacts with, through all of its inputs and outputs – the dynamic structures of brain[r] within B[r] within W[r].[150]

[146] For how W[z] relates to W[r] see Appendix 1.
[147] The concept of a goal state is more precisely defined in due course (Section 10.5.1, footnote 163 second underlining).
[148] The idea of maximal efficiency can be understood here to apply, for example, to the level of resolution attained in various areas of this mapping which, for the working purposes of this note, will be assumed to be as high as is needed to allow a person, as their B[r], to interact in a normal successful way with their physical environment, W[r]\B[r]; for example, through practicing a normal level of competence in tasks requiring eye-hand coordination.
[149] Noting from footnote 144 that B[r] in W[r]\B[r] is just another way of saying W[r]\D[r].
[150] This structure, with its inputs and outputs, is illustrated in Fig. 8, and described in affiliated text, in Section 9.3.1 of Part 1.

10.3 Force generators and agents

If it is assumed that the level of capability of D[r] exceeds that needed to achieve inferential convergence, it can then be further proposed that D[r]'s development may reach a stage in its capacity to map the state and state trajectory of B[r] in W[r]\B[r] sufficient for it to begin to infer and predict the effects of what can be called *force generators*.

A force generator[r] can be defined as any entity[r] that causes *change* in the state trajectory of W[r]\D[r].

Using phenomenal experience as a guide, we can see that as we experience and perceive ourselves operating moment-by-moment as a physical body in a physical world we can attribute the changes we observe to be taking place in that world as being due, or as having been due, to the operation of physical forces whose impacts we are able to varying degrees to predict.

For example, the effects of what we call gravity are apparent to us, and are moment-by-moment relatively easy for us to predict. Wind is also apparent, and its effects are also predictable, though less so. Through language we can name gravity, magnetism, wind, electric and aqueous currents, and so forth, but we do not actually perceive these things as objects or surfaces but as a different class of entity[i] – namely entities that we can *infer* exist as causes of change in the movement of objects or surfaces or of their spatial relations to each other – but that we perceive only *indirectly* through their effects.[151] This includes our perception of any effects that forces have on our bodies, manifest in the form of pushes and pulls.

Experience shows us that the most complex, most difficult to predict class of force generators that we can infer drive changes in our phenomenal worlds are almost always other living things, including other people. Call such living force generators, *agents*. Agents express forces in complex sequences through the physical actions of their bodies. It seems likely that we can infer the existence of agents – and resolve them out as phenomenal entities[i] – through the efforts of our respective D[r]s to find the most efficient and effectively resolved – i.e. focused – states, Ξ, to best map the state and the state trajectory of W[r]\D[r].

Having thus used phenomenal experience as a guide, we can now summarize and extend the current line of conjecture by proposing that inferential convergence can take the processor in D[r] – through its ever more complete, contextualized and effective mapping to Ξ of the state and state trajectory of W[r]\D[r] – to a point where, for relatively familiar environments, it:

- largely masters the inferential task of mapping and predicting the presence and ongoing effects of individual force generators[r] – though less so agents – on the evolving recognition state, Ξ; and
- masters to a degree approaching completeness, the inferential task of mapping and predicting the presence and effects of that force generator[r] which is D[r] *itself* as it implements its generative model – through operation of the recognition cycle – to express via active inference those policies it infers will best sustain B[r] homeostasis.

It is proposed here then that D[r] will, by means of inferential convergence, attain a level of refinement of its recognition and generative models such that it can, and does, 'recognize' – i.e. resolve out – in its mapping to Ξ of the state and state trajectory of W[r] an evolving component of that map that correlates to its *own* presence and ongoing impact as an agent on the state and state trajectory of both itself as B[r] and its environment W[r]\B[r].[152]

[151] For example, in the way we can perceive magnetic force by the effect a magnet has on a distribution of iron filings scattered on a page placed over it.
[152] It will be proposed below that in this way D[r] architecture and processes can be configured such that D[r] is able to predict itself *to* itself *as a contextual requirement* that will ever more clearly emerge as inferential convergence proceeds, and that by this means a form of consciousness can arise that includes D[r]'s moment-by-moment subjective experience of itself as a physical body situated within, and interacting with, a physical world. Cast in terms of ideas outlined in footnote 145, this is to say that once the capability to infer *context* being built through iterative refinement of the D[r] hierarchical processor's generative and recognition models achieves sufficient refinement – embodied in an increasing refinement of the structure and activity of its contextualising intra-layer lateral and inter-layer downwards linkages – an inference by D[r] of D[r] itself can emerge in the form of an inference of that which is required to bring the mapping of W[r]\D[r] by the processor's recognition and generative models through to completion of a single, unified overall context within which all elements of that mapping are able to sit in maximally defined/resolved relations to each other. This is to say that the process by which the D[r] hierarchical processor progressively 'learns' to modulate for context will lead it to a stage where further progress towards a single, unified, *most efficient*, contextual setting applicable across the entire processor will drive an ever better inference of anything of significance that is 'missing' in the processor's mapping of the 'hidden causes' of inputs from W[r]/D[r] into D[r]. Crucially, these 'last pieces in the jigsaw' – the force generators that can only be perceived indirectly through their effects on more directly perceivable causes – will include D[r] itself, which will come to be inferred to be that 'entity' which is the 'hidden cause' of B[r] motor output and its impacts on W[r]\B[r]. Thus D[r] will be able to reach an inference of itself as being that agent (force generator) that moves B[r], and will be able to map itself as that 'entity' – defined below as D[a] – that moves the Ξ substate B[i]. As described below, this will be central to how a person – as a D[r] – can come to experience themself as being that entity – defined below as D[a] – that inter alia is the entity that moves their physical body, i.e. their B[i], and through it their physical world, i.e. their W[i]\B[i]. (And note here the strong alignment between these proposals and those outlined in Section 9.4.5 of Part 1.)

10.4 Parsing out the recognition state – the idea of optimal parsing

Bearing this in mind, it is now useful to revisit the proposal made in Section 9.3.1 of Part 1 that the *totality* of a person's evolving subjective perceptions – including of being a physical body operating in and on a physical world, but also of experiencing desires and other feelings – will by some means be fully reflected in the evolution of the recognition state, Ξ.

But with this – although these evolving subjective perceptions are of experiencing being a fully relationally integrated and continuous phenomenal-self-in-phenomenal-world – it is also phenomenologically clear that our experience of this phenomenal-self-in-phenomenal-world is such that it exhibits contents that are clearly differentiated in how they are laid out to our perception in space and time.[153]

Together, these ideas lead to the conclusion that processes in D[r] are operating to parse out the recognition state into substates that correspond to the differentiated contents of our perception.

As pointed out in Section 9.3.1 of Part 1, there could – in the abstract – be many ways of carving up recognition space into subspaces.[154] But let us assume here that the way in which recognition space is parsed out into subspaces by operations in D[r] is the result of Darwinian evolutionary pressure on B[r], on its nervous system[r] and on brain[r] that has driven maximization of a certain specific type of architecture and processing arrangement in D[r]. Say this leads to what can be called *optimally parsed* substates of Ξ.

Then say – consistent with Section 10.3 – that this architecture and processing arrangement is one which *inter alia* optimizes D[r]'s ability to attain inferential convergence so as to maximally 'recognize' – i.e. resolve out/define – in its mapping to Ξ of the state and state trajectory of W[r] those aspects of that map that correlate to its own presence and impact as a specific force generator[r] on the state and state trajectory of W[r].[155] Call the specific force generator or agent that is recognized by D[r] – and that correlates to D[r]'s own presence and impact on W[r] – the *primary agent*, D[a].

Then say that this architecture and processing arrangement has further evolved to operate such that it parses out and configures recognition space in such a way that D[r] is not only able to infer in a defined way the presence of D[a], but such that D[r] can *predict and operate on itself* by means of inferentially identifying – and then by defining – itself as *being* D[a], i.e. by placing itself into *identity* with D[a].[156]

Consider then that, when D[r] begins to predict and operate on itself by means of inferentially identifying itself as D[a] – i.e. by placing itself into identity with D[a] in mapping itself as D[a] – D[r] is enabled, by means of the architecture and processing arrangements described above, subjectively to experience itself as if it actually *is* the force generator D[a], i.e. as if it actually *is* the primary agent.

On this basis it is proposed here that, by so doing, D[r] can inferentially come to experience itself as a 'presence' 'looking out from' and 'expressing itself from' a 'point of view' that corresponds to a continuously evolving position in a purpose-configured mapping of all of D[a]'s *relations*, and its inferences of potential relations, to the optimally parsed out substates of Ξ, noting that the sum of these substates will remain a synchronous predictive mapping of the continuously evolving state and state trajectory of W[r]\D[r].[157]

If this is true then, as a D[r] experiencing itself as a D[a], each of us should know from personal experience what the result of this mapping 'looks like' from the 'inside' which, broadly speaking, amounts to our knowing:

1. that the described architecture and processing arrangements in D[r] optimally parse out recognition states into substates – as mentioned above – that encode the differentiated contents of our perception; for example specific objects: e.g. a table or chair; specific bodily sensations: e.g. a pain or an itch; and specific emotions: e.g. fear or relief, and

2. that the described architecture and processing arrangements in D[r] optimally parse out recognition space into subspaces – for example, the subspace we perceive as 'physical space' which *inter alia* we perceive to 'contain' specific objects *as they sit in relation to ourselves as bodies and to each other* – but also into a subspace that can be called 'emotional space' where substates in this space correspond inter alia to our perception of one or more desires/appetites/emotions.

---

[153] For example, we can perceive discrete, spatially distinguishable, physical objects in relation to each other – but also as independent to each other – in their motions and properties.

[154] Where recognition space is the space within which any given recognition state Ξ will define a position (as per Section 9.3.1), and where any subspace of recognition space will be that space within which any member of a specified type of substate of Ξ defines a position (also as per Section 9.3.1).

[155] It seems reasonable to assume that there would be strong Darwinian natural selection driving maximisation of capability to inferentially identify and predict the impacts of all force generators encountered, given that by definition these are the entities[r] that cause the trajectories of states of W[r]\D[r] to change, and that it is through such change that a great many of the risks and opportunities that W[r] can present to B[r] will arise.

[156] Noting that this idea is fully consistent with, and supported by, the mechanistic proposal made in Section 9.4.5 of Part 1.

[157] This idea better refines the intuition underlying the idea of 'projection inversion' as developed in Section 8.1 of the Essay.

10.5 The construction of phenomenal space

10.5.1 Space 1 and Space 2

Consider now more closely just what it is that D[r] is inferentially identifying through its optimally parsed out mapping to Ξ of the state and state trajectory of W[r] when it identifies those aspects of that mapping that correlate specifically to its own presence and impact as a force generator[r] – i.e. as D[a] – on the state and state trajectory of W[r].

Drawing on the ideas and observations provided above it can be proposed that D[r] will inferentially identify as the primary agent, D[a], that force generator[r] which *consistently over time*:

1. Works to move the current position, *CP1*, which

    - D[r]'s hierarchical processor is inferentially resolving out by balancing a class of cueing signals driven mainly by *interoceptive inputs* to D[r], and
    - is a position in a subspace of Ξ space – call this *Space 1* – dimensionally defined by the subset of Ξ space degrees of freedom needed to define as a unique position[158] any of all of the potential Ξ substate permutations that could arise through the hierarchical processor's balancing of any of all potential permutations of that class of cueing signal

    to a position in Space 1 – call this *Position 1* – which is defined as a position that satisfies the set point values for B[r] homeostasis[159]

2. through working concurrently to move the current position, *CP2*, which

    - D[r]'s hierarchical processor is inferentially resolving out by balancing a class of cueing signals driven mainly by *exteroceptive and proprioceptive inputs* to D[r], and
    - is a position in a subspace of Ξ space – call this *Space 2* – dimensionally defined by the subset of Ξ space degrees of freedom needed to define as a unique position[160] any of all of the potential Ξ substate permutations that could arise through the hierarchical processor's balancing of any of all potential permutations of that class of cueing signal [161]

    to a position in Space 2 – call this *Position 2* – which is defined as that position required to move the current position in Space 1 to Position 1.

Although points (1) and (2) may seem abstract, they simply describe in more objective terms[162] a situation we each subjectively experience as commonplace, as demonstrated by the sentence:

> "I move the state of my physical body, and by this means I move the state of the physical world, so that I can move from an internal state of desire to an internal state of satiation."[163]

This obvious yet fundamental statement is, for all practical purposes, true for any person.

---

[158] Here each position within a subspace of recognition space (a.k.a. Ξ space) will correspond to a set of specific values in each of the degrees of freedom used to define that subspace. Such positions can be viewed as a substates of the overall state Ξ. Here CP1 is the current momentary substate of the overall recognition state Ξ defined as a position in the Ξ subspace Space 1. This way of looking at recognition space and recognition states is described in detail in Section 9.3.1 of Part1.

[159] The number of degrees of freedom required for Space 1 should correspond to a hypothetical number of independent set point values needed to maintain B[r] homeostasis. This homeostasis may broadly be divided into biochemical[r] and biophysical[r] homeostasis.

[160] CP2 is the current momentary substate of the overall recognition state Ξ defined as a position in the Ξ subspace Space 2.

[161] Here the number of degrees of freedom used to fulfill this requirement – and it is probably immense – can be estimated to be the total number of degrees of freedom of Ξ space minus the number of degrees of freedom of Space 1. (Assuming Space 1 + Space 2 = Ξ space, which for *current* working purposes will be assumed to be the case.) This number will reflect the overall inbuilt capacity of B[r] systems – including the sensory channels[r] and the hierarchical processor in D[r], operating through its recognition and generative models – to optimally parse, resolve out and map onto Ξ the ongoing state and state trajectory of all of B[r] in W[r] that is within B[r]'s sensory range. This capacity will, in turn, result from the extent to which those B[r] systems[r] have evolved under pressure of Darwinian natural selection; where such selection will need to have proceeded in such a way that those systems[r] will, at a minimum, have the capacity to deliver *sufficient* degrees of freedom for Space 2, and associated levels of resolution in defining any current position in Space 2, to serve the practical purposes of reliably – through D[r] implementation of active inference – keeping B[r] sufficiently close to Position 1 to keep it alive.

[162] Terms that align well with ideas such as those provided in Tschantz et al. (2021) Simulating Homeostatic, Allostatic and Goal-directed Forms of Interoceptive Control Using Active Inference, doi: https://doi.org/10.1101/2021.02.16.431365.

[163] To make this plain, a person's moving the state of their physical body, and thereby the state of their physical world, corresponds to (2) while their moving from a state of desire to a state of satiation corresponds to (1). Note then that (1) plus (2) essentially cast D[a] as a reification of allostasis. In due course (Section 10.5.4) Space 2 will be identified with κ space – the space linked to a person's perception of a phenomenal, 'outer' 'physical' world of B[i] in W[i] – and Space 1 will be identified with ρ space – the space linked to a person's perception of a phenomenal, 'inner' 'emotional' world of feelings/desires/appetites – consistent with definitions provided in Sections 9.3.2 and 9.3.3 of Part 1. Note also that Position 1 and Position 2 combine to define in generic terms that set of recognition states that can be called *goal states*. This is the definition of a goal state that will apply from here on in Working Note A.

Consistent with this, if a working proposal is accepted that D[r] expresses language as an output and receives it as an input[164], then it is possible to go on to use the statement above – and others like it – as further support for the idea that a person's subjective experience of themselves in the world arises when their D[r] begins to predict and operate on itself by means of inferentially identifying itself as D[a]; i.e. where D[r] begins to experience itself from the '*point of view'* of D[a], and *as being* D[a], based on an ongoing inference that it *is* D[a].

Accordingly, consider that for any given person the core processes giving rise to their ongoing subjective experience of being in the world are taking place within their D[r], as it sits within its Markov blanket, and that when a person – i.e. their D[r] – makes statements about *itself*, it is actually making statements about what all of its architecture and hierarchical processing arrangements – in conjunction with its developed recognition and generative models – have inferentially resolved out to be D[a] through optimal parsing of the evolving recognition state, Ξ, held within D[r]. To this end, say that any given person – i.e. their D[r] – in referring to itself in terms of its *prediction* of itself, D[a] – will refer to itself as "I".

It is then possible to interpret everyday statements such as, "I will go to the shops", as a report from a person's D[r] of a *prediction* it is making about the nature of the effects the force generator D[a] – which D[r] infers to be *itself* – is going to have on the motion through recognition space of the recognition state, Ξ.[165]

Similarly, statements such as "I am hungry" or "I see a food" can be interpreted as reports from a person's D[r] of aspects it is describing from the *perspective* – i.e. *point of view* – of what it infers to be the momentary 'situation' of D[a] within the overall optimally parsed out momentary recognition state, Ξ, which itself – in its totality – constitutes a momentary position in overall recognition space.

Because D[r]'s inference of the existence of D[a] – and how D[a] will move the recognition state – will be based on D[a]'s effects as a force generator, D[r]'s inference of the 'situation' of D[a] in terms of any momentary recognition state, Ξ, will be based on an ongoing inference that works to identify and parse out D[a] as inhering within that evolving substate in Ξ whose *motion* can be inferred to be caused by a single agent, such that that single agent can be inferred to be driving those motions of Ξ that satisfy (1) and (2) above. By this means D[r] can inferentially parse out that ongoing substate of Ξ that 'contains'[166] D[a], and from this D[r] can infer a 'position', and thence a 'perspective' or 'point of view', for D[a] as a set of relations between that ongoing substate and its *environment*, which will be made up of all of the optimally parsed out substates of Ξ that are *outside* the substate that D[r] can infer 'contains' D[a].

Thus, the perspective, or point of view, that D[r] can infer for D[a] through optimally parsing Ξ can be seen as *entirely relational*. "I am hungry" can be seen as a report of a distance and direction defining the momentary position in Space 1 which D[r] infers to be the position of D[a] relative to Position 1, where Space 1 is a subspace within overall recognition space. "I am very hungry, mildly thirsty and tired", simply reports a different distance and direction defining the position of D[a] relative to Position 1.

In an immensely high number of dimensions, and for that reason *leaving an immense amount unsaid*, "I see food", will – *together with everything else*[167] that D[r] must in the same moment be inferring about D[a]'s situation as an ongoing substate within the momentary substate of Ξ defined by CP2, and about CP2 as a position in those dimensions in Space 2 – define the distance and direction of CP2 relative to Position 2 in circumstances where CP1 in Space 1 can be described as "I am hungry".

The relevant person's D[r] might then predict – and might even say – "I am going to eat that food". This will be a prediction made by their D[r] about the nature of the series of effects it predicts that the force generator D[a] will have on the motion of the recognition state, Ξ, through both Space 2 and Space 1. Importantly, it will – at the same time – be a prediction that D[r] itself fulfills as it operates through expression of active inference to bring the distances identified above as close to zero as possible.[168]

[164] This seems a reasonable proposal since it is difficult to conceive of *where else* sentences/reports such as that just given in demonstration could originate other than from *within* the Markov blanket housing D[r], consistent with the wider proposal being made in this Working Note that D[r] will contain the apparatus and processes responsible for generating a person's subjective perception of themselves in interaction with their environment, as well as being required to generate the motor output needed for them to make speech.
[165] This specific statement reports D[r]'s prediction of the position to which D[a] is going to move CP2 in the subspace of recognition space that has been defined above as Space 2.
[166] It is difficult to find the right word here. To say: "…the ongoing substate of Ξ that 'contains' D[a]…" might equally be put as "…the ongoing substate of Ξ 'within which D[a] inheres'…".
[167] Including all that D[r] presumably *could* report about the situation of 'I' in that moment, but leaves 'unsaid'.
[168] The idea that D[r] will operate in this way to bring all of the distances described above to zero can be viewed as an implementation of the Free Energy Principle, as per Friston, K. (2010) The Free Energy Principle: A Unified Brain Theory? *Nature Reviews, Neuroscience 11* pp127-138 and many other references.

This last observation is an assertion that highly effective application and refinement of its recognition and generative models should allow D[r] to map and predict – in the form of those motions of Ξ it infers to be driven by D[a] – the effects that expression of its own generative model will have – through application of active inference via the recognition cycle – on the motions of the state of W[r].

Moreover, based on D[r]'s ability, as described above, to parse out that ongoing substate of Ξ which 'contains' D[a],[169] D[r] should be able to map and predict those effects on W[r] in a way that differentiates between the motions of B[r] and the motions of W[r]\B[r]. This is since, with optimal parsing, and in normal circumstances, the state and motions of B[r] will always be mapped to an evolving substate of Ξ that D[r] is *able consistently* – over time – to infer *persistently* 'contains' D[a][170], with the state and motions of W[r]\B[r] then being mapped to the *complementary* evolving substate of Ξ.[171]

10.5.2 The nature of phenomenal space – the 3-space inference

To get a clearer picture, it is now useful to describe in more precise terms the phenomenology of a person's experience of themselves as a physical being operating in, and interacting with, a physical world.

Consistent with the approach adopted above, it will be the case that this description can only be based on statements that my own D[r] can make about itself, having inferentially identified itself as the agent D[a], having placed itself by inference into 'the position' of that D[a], and then from that 'position' proceeding to describe how it experiences itself as a physical being in a physical world from 'the perspective' of that D[a], which my D[r] calls, 'my perspective' or 'my point of view'.[172]

From this perspective, a person's experience of themselves as a physical being operating in and on a physical world manifests subjectively as a continuous stream of *Specific Instances* – that is, as a stream of moment-by-moment perceptions – of the self as a body existing at, and acting from, the centre of a region of infinitely extended space.

Define as the *Centre* the position or vantage point from which a person perceives space. This vantage point will not be identical to the person's body, although almost without exception people closely identify it with a point-like position 'right behind the eyes' of their perceived body.[173]

Define the *Self-as-Centre* as the subjectively perceived set of *contents* of the *vantage point* from which a person perceives space, and define the *Self-as-Body* as a person's body as they perceive it to exist in space from the vantage point of the Self-as-Centre, which is identical to the vantage point of the Centre.

Note that the vantage point of the Centre can be considered to have contents – which constitute the Self-as-Centre – in so far as this is the spatial position where, in normal circumstances, people report perceiving their private 'internal world' of desire, emotion, other 'states of mind' and 'mental events' to be 'located'. A person perceives the space surrounding the Self-as-Centre as being populated by *Specific Objects*, including the Self-as-Body. Any particular Specific Object – e.g. a chair, a tree, a river – will be perceived as existing in space and time either as a closed surface, extended surface, closed volume or as an extended volume[174] and as sitting within a boundary defined by at least a fairly specific range of angles and distances from the Self-as-Centre.

[169] Or perhaps more precisely, 'within which D[a] abidingly inheres'.
[170] This is because as a force generator[r], D[r] will always – over time, by implementation of its generative model through active inference – be moving the state of B[r] to get to Positions 2 and 1, and will therefore always, over time, by implementation of its recognition model through optimal parsing, map the state and motion of B[r] onto an evolving substate of Ξ that it can infer 'contains' D[a].
[171] That is to say, with the state and state trajectory of W[r]\B[r] being mapped to all of the optimally parsed out substates of Ξ that D[r] is *unable* consistently – over time – to infer *persistently* 'contain' D[a]. As described previously, these substates can be considered to make up D[a]'s phenomenal 'environment'.
[172] I can be confident that other people have a similar, if not identical, experience of themselves as a physical being operating in, and interacting with, a physical world because they (their D[r]s) verbally report having the same types of experiences.
[173] Such subjective identification of this vantage point with the spatial position of the body – and particularly of the head/eyes – does not always apply. Situations can be created where the Centre and Self-as-Centre can be spatially displaced from the Self-as-Body, such as in experiments described in Ehrsson (2007) The Experimental Induction of Out-of-Body Experiences *Science 317* 1048 and for wider context see Bekrater-Bodmann, R. et al. (2020) Interoceptive Awareness Is Negatively Related to the Exteroceptive Manipulation of Bodily Self-Location *Front. in Psychol. 11*: 562016.
[174] Discrete objects – e.g. cups, books or marbles – are perceived to be enveloped by a *closed surface*. Extended objects – the sea, a railway line, the surface of the earth – are perceived locally as *extended surfaces*, i.e. they extend beyond the immediate range of perception. A further group of objects includes flames, other light sources, mirrors, transparent/translucent liquids and solids, and clouds, where an object surface can be hard to perceive. These objects are what is meant by *closed volumes* and *extended volumes*. Overall, Specific Objects may exhibit perceived surface or volume characteristics such as smoothness, brightness, colour or, to the ear, sound emissions or, to the touch, viscosity, wetness, temperature and so forth, but all of these attributes will be perceived *in relation to each othe*r as existing in both space and time.

Importantly, people are unable to imagine perceiving space and time as having anything other than infinite extension. This is manifest in our complete inability to imagine how a boundary beyond which space or time might cease to exist could be perceived.[175] Nor can we imagine perceiving any more than one *single* arena in space and time that we would count as *physically real*, or of perceiving anything we would count as physically real, but that exists entirely outside this single arena in space and time.[176,177]

So while the Self-as-Body and its environment of Specific Objects, their interrelationships, interactions, and motions are always perceived on a moment-by-moment basis, those moment-by-moment perceptions are *always* implicitly, and more broadly, perceived as being Specific Instances of occurrences in a single and absolute arena of space and time within which *all real* Specific Objects to be perceived have existed, currently exist, or will exist.

On this basis define the *Specific Characteristics* of a Specific Instance of a person's perceptual experience of themselves as a physical being embedded in and interacting with a physical world as being composed of all of the Specific Objects perceived in that instance together with the perceived motions and perceived interrelationships and interactions of those Specific Objects with each other and with the Self-as-Body.

Then, with respect to these Specific Characteristics, note again the generality that all Specific Objects – including the Self-as-Body – and all of their interrelationships, interactions and motions, are perceived as arising, changing, and disintegrating through transformations of their form, position and other attributes[178] within a single arena of space and time as a dynamic evolution of states, including an evolution of states driven by interactions between the Self-as-Body and its perceived physical environment of Specific Objects.

On this basis define the *General Characteristics* that encompass all of the serially unfolding Specific Instances of a person's perceptual experience of themselves as a physical being embedded in and interacting with a physical world as being the infinitely extended arena of space and time within which a person perceives *all conceivable real*[179] *physical objects, entities and events* as having evolved, as evolving, or as yet to evolve, from the vantage point of the Self-as-Centre.[180]

The value of this exploration of the phenomenology of a person's experience of themselves as a physical being operating in, and on, a physical world rests in what it tells us about the optimal parsing D[r] must be using to parse out the recognition state Ξ into substates and to parse out overall recognition space into subspaces.

Perhaps the most important things it tells us is that – at the foundations of D[r]'s inferentially optimised parsing out of all recognition states Ξ, and of Ξ space – there must be at least one primary and fundamental inference. This inference is not fluid but structural, and it is reflected in the abiding and inalienable General Characteristics of a person's perceptual experience of themselves in the world. Call this structurally built-in inference the *3-space inference*. This inference is that *all* potential positions – i.e. Ξ states – in Ξ space can – *sufficient for the purposes of the survival of the organism* – be parsed out into substates such that *all* such substates can be *spatially indexed*[181] – i.e. relationally mapped – onto *three degrees of freedom*.

[175] It is possible by means of extreme abstraction to deduce the existence of such boundaries, for example by proposing the existence of the event horizon of a black hole. But such conceptualisation is not part of our natural moment-by-moment *perception* or imagination of ourselves in the physical world and is the result of the application of highly non-sensory means of understanding. (2) A parallel case is that we can say with confidence that our perception of colour is limited to a certain spectrum, and that we simply *cannot imagine or conceive of* how colours outside that spectrum would appear to our perception. This remains true, even though we can deduce by abstract means that frequencies of light exist that we somehow might have been capable of perceiving if our biology had been different. But because our biology is what it is, such hypothetical colours simply do not exist and *cannot* exist as perceptions for us. Likewise, boundaries beyond which space and time cease to exist cannot exist as perceptions for us and we cannot conceive of such perceptions.

[176] We do not count dreams or things we imagine as physically real, and even ideas in physics such the many-worlds interpretation of quantum mechanics do not predict that we would perceive more than one physical world. Nor does it seem feasible that a virtual reality, generated on devices based in the physical world we currently perceive, could ever displace that physical world in being perceived by a person as actual physical reality. In an extreme case, if the person was able only to perceive such a virtual world and was never able to visit the physical world (a 'fooled for life, trapped in the Matrix' scenario) they would still only be able to perceive as real – and conceive of perceiving as real – that *one* single, albeit virtually generated, arena in space and time.

[177] Note that given these conceptual constraints it is only natural that most people come to believe that everything that is real must in some way exist within space and time, reflected in some way in the physical reality that we seem *only* to be able to perceive to exist in terms of space and time.

[178] Attributes such as colour, brightness, loudness, temperature, texture and so forth.

[179] Further to footnotes 176 and 177, how we perceive things as being either 'real' or 'imaginary' is explored in depth in Section 11 of Part 3.

[180] For further insight and observations in this area, and in relation to work in Section 10.6 below, see in particular Chapter 9 in Revonsuo, A. *Inner Presence: Consciousness as a Biological Phenomenon* MIT Press, Cambridge MA, USA 2006.

[181] Here the term 'spatial indexing' is defined as the *process* of relationally mapping the state Ξ, by virtue of relationally mapping each of its inferentially parsed out substates, onto a three dimensional space. The outcome of this process can be referred to as the substate's spatial index. A substate that has been so mapped can be said to have been spatially indexed.

Again, from the perspective of a person's subjective perceptual experience of themselves as a physical being operating in, and interacting with, a physical world – i.e. from the perspective D[r] adopts by placing itself in identity with its inference of itself, D[a] – this seems commonplace. Everything that a waking a person perceives presents itself to them as having a location and motion that is distributed *somewhere* within the past, present or future of a single, eternal, infinitely extended *three dimensional* physical world.

Integral to this, it is clear that people perceive the 'physically real' part of themselves to be their spatially defined, three dimensional physical body – i.e. their Self-as-Body – even as they experience (identify) themselves – beyond their perception of being a 'physical self' – as being that 'psychological force' or 'will' (i.e. that force generator – D[a]) that moves the state of their surrounding three dimensional physical world through moving the state of their three dimensional physical body.

Also, it seems clear that D[r] parses out the recognition state Ξ such that it assigns itself – in identifying itself as D[a] – as being located somewhere 'within' the spatial distribution it assigns to its 'physically real' Self-as-Body, since it is the Self-as-Body that corresponds to that evolving substate of Ξ that D[r] can infer – over time – persistently 'contains' D[a]. By this means D[r] can experience itself as – i.e. infer itself to be – that force generator which 'owns' the Self-as-Body since it can infer – over time – that it is that force generator that is persistently moving the Self-as-Body.[182]

More specifically, from the perspective of a person's subjective perceptual experience of themselves, it seems that D[r], in experiencing itself as D[a], locates itself at the Self-as-Centre; that is, at the vantage point, or 'point of view', that it experiences itself to be 'looking out from' into a surrounding phenomenal world of three dimensional space infused with its evolving pattern of Specific Characteristics.

It also seems clear from our perceptual experience that people inferentially assign a spatial location to the contents of what we call our 'minds'. Thus any normally conscious person – including you or I – experiences an 'inner world' of desires, emotions, and thoughts that we predominantly – along with ourselves as D[a] – locate at the Self-as-Centre. Hence a person predominantly subjectively experiences their 'mind' as being located *within* their head (just behind their eyes) *within* their Self-as-Body, which itself – coincident with the 'point of view' from their Centre – they perceive to be at the centre of, and in turn *within*, an eternal, infinitely extended three dimensional physical world.[183]

The following section provides a more explicitly PPP-oriented approach to describing implementation and implications of the 3-space inference.

10.5.3 Degrees of spatial indexing

The 3-space inference – and with it spatial indexing – must have arisen in our distant evolutionary past and become highly refined through millions of years of Darwinian selective pressure. Studies of animal and human brains show the visual system, and other systems designed to provide information in their inputs that would allow processors such as D[r] to implement high-definition spatial indexing, arose in this way. Such studies also indicate that these systems became highly developed and specialized at stages in evolutionary history millions of years before the emergence of human beings.[184]

The proposal here then is that Darwinian selective pressure must already have 'discovered' at a pre-human evolutionary stage that extremely capable, reliable and efficient[185] recognition and generative models could be built on an overarching 'inference' – an inference then progressively physiologically and neurologically hardwired into B[r] and D[r] through natural selection – that *all* hidden states and state trajectories arising in B[r] and W[r]\B[r], and as B[r] interacts with W[r]\B[r], could be sufficiently well predicted *for all of the practical purposes of survival of the organism* by encoding such states, state trajectories and their relations to each other, as evolving as transformations over time in properties relationally mapped onto three degrees of freedom.[186]

---

182 Consistent with ideas presented in the final paragraph of Section 10.5.1.
183 The word 'predominantly' is used here in deference to the contents of footnote 173.
184 For example, in relation to spatial discrimination in the visual systems of various mammals and primates, see Haas, J.H., The Evolution of the Visual System in Primates pp1-26 in *The New Visual Neurosciences* Werner, J. & Chalupa, L. eds, MIT Press (2013).
185 And for these reasons highly selectively competitive.
186 The idea that organisms will, through natural selection, evolve in interaction with their environment such that inferences about the nature of that environment progressively become structurally embedded into their phenotype – i.e. that their phenotype will be progressively refined such that it is better able to interpret and interact with its environment – is just a restatement of the idea of natural selection itself. Such evolutionary pathways are possible because W[r] is bounded in its variability, i.e. that in certain respects it is highly predictable. A further example of an inference structurally built into how a person processes inputs into their D[r] relates to colour vision, as described at footnote 175(2). Note also, in relation to such 'embodiment', Section 5.1 in Ramstead, M.J.D. et al. (2020) A Tale of Two Densities: Active Inference is Enactive Inference *Adaptive Behaviour 28* pp225-239.

The implication of this is that for all human beings the neurological architecture supporting development of the recognition and generative models being built from birth in our respective D[r]s will have *no latitude* other than to seek inferentially to spatially index – i.e. relationally map – onto three dimensions all of the subsequently[187] further inferentially parsed out substates of the overall recognition state, Ξ.

This inferential mapping – which in PPP terms would amount to a *mandatory prior* hardwired into human neurology[r] – would precondition a D[r]'s recognition and generative models inferentially to resolve out to a *best fit* <u>*all*</u> possible information held in the evolving recognition state Ξ in terms of an evolving distribution of positions relationally mapped onto those three dimensions.

Essentially then, these three dimensions can be seen as constituting a kind of 'container' within which D[r]'s recognition and generative models are *inescapably* required inferentially to find relational positions and distributions – i.e. mappings – for all of the information held in any potential recognition state Ξ.[188]

If it is accepted that D[r] is bound by its neuronal[r] architecture inferentially to map all recognition states Ξ onto a three dimensional space then it is possible to envisage *degrees of definition* in the mapping that D[r] can apply to evolving substates of Ξ.

Call these degrees of definition hard inferential mapping (*hard mapping*), soft inferential mapping (*soft mapping*) and default inferential mapping (*default mapping*). Say then that, due to brain[r] hardwiring, all of the evolving substates of Ξ, no matter how they are otherwise inferentially identified/parsed out by the recognition and generative models in D[r], will fall somewhere along a spatial indexing spectrum that runs from hard mapping through soft mapping to default mapping.

Hard indexing

Say that hard inferential mapping will occur for the set of evolving substates of Ξ where there is a component, or are components, of that substate's activation that are driven by one or more cueing signals that carry information sufficient to allow the recognition and generative models in D[r] to achieve a *highly defined* inference of its spatial index. Call such highly-defined spatial indexing *hard indexing*. All this means is that it will be possible for D[r] to infer a highly spatially defined distribution of positions and motions over time for such evolving substates within the imposed three dimensional container.

The types of cueing signals travelling into D[r] that will provide information sufficient to allow hard indexing will include those from exteroceptive, proprioceptive and interoceptive pathways that can empirically be identified as capable of providing inputs that result in a person's subjective perception of Specific Objects and their motions – including parts of their Self-as-Body – where these are perceived with *high spatial definition*.

In normal circumstances examples would include nearly all visual inputs as well as certain proprioceptive and tactile inputs; in particular, the latter two[189] would include inputs to do with the hands, fingers, fingertips, and the tongue. Examples also include certain nociceptive inputs.[190] This hard indexing will allow for hard inferential mapping of members of the relevant set of evolving substates in Ξ, consistent with the earlier assumption that the evolving substates of Ξ constitute the contents of what is consciously perceived on a moment-by-moment basis.

It is empirically clear that nearly all hard indexed substates of Ξ will resolve out to what is consciously perceived to be either *outside*, or at the surface, of the Self-as-Body.

---

[187] 'Subsequently', in that it will be held here that the 3-space inference will precede such further – and, in many cases, more flexible – inferential parsing of the recognition state.

[188] The idea that there may be an a priori, mandatory prior in the form of the 3-space inference resonates with ideas on the a priori nature of perceived space introduced over two hundred years ago by Immanuel Kant: see Chapter I. Part 1. Section 1: Space, in *Immanuel Kant's Critique of Pure Reason* trans. Kemp Smith, N. Macmillan & Co. London UK 1929.

[189] See Tactile Sensation and Proprioception pp207-212 in Patestas, M.A. and Gartner, L.P. *A Textbook of Neuroanatomy Ed.2* Wiley & Sons, Hoboken New Jersey USA 2016.

[190] These inputs result pain, which is often be perceived as highly spatially local to one part or another of the Self-as-Body.

Soft indexing

The types of cueing signals travelling into D[r] that will provide information sufficient only to allow *soft indexing* will include those from exteroceptive, proprioceptive and interoceptive pathways that can empirically be identified as able *only* to provide inputs that result in a person's perception of Specific Objects – or *aspects*[191] of Specific Objects – that have *relatively low spatial definition*.

These will include perceptions due to visual inputs in insufficient light, most auditory inputs, many proprioceptive inputs, and many tactile and interoceptive inputs such as those leading to sensation of some kinds of internal pains and other sensations perceived as arising at imprecise locations within, or even perceived as arising throughout[192], the Self-as-Body. Soft indexing will allow only for soft inferential mapping of the relevant evolving substates in Ξ, consistent with the assumption that the evolving substates of Ξ constitute the contents of what is consciously perceived on a moment-by-moment basis.

Here it is empirically clear that in normal circumstances a great many soft indexed substates of Ξ will resolve out to what is consciously perceived to be either *inside*, or at the surface, of the Self-as-Body.

Default indexing

The idea of *default indexing* arises since it is apparent there are one or more sets of evolving substates of Ξ whose members seem to be driven by cueing signals where none of those signals carries information that is either sufficient or of a type that will allow the recognition and generative models in D[r] to infer a spatial index for those substates. This means D[r] will be unable with any definition to relationally map any member of such a set onto a distribution of positions in the imposed three dimensional container.

An empirically, phenomenologically identifiable example of such a set is that whose members correlate to most[193] desire/appetite states – such as those affiliated with hunger or thirst – which are almost certainly activated by certain types of interoceptive input to D[r]. Emotional states, such as fear or relief, also form such a set and seem also to be driven – or at least to be informed[194] – by interoceptive inputs[195] to D[r].

If, as proposed, it is the case that D[r] is bound by its neuronal[r] architecture and processes inferentially to map *all* of the states and state trajectories of Ξ onto a three dimensional space, there is a question of where in that space default inferential mapping will position states and state trajectories that are default indexed.

It seems empirically clear, based on phenomenal experience, that members of the sets of evolving substates of Ξ that are default indexed are inferentially mapped by D[r] to the location in perceived three dimensional space that – as noted in the preceding section – people subjectively experience as their Self-as-Centre.

Hence it seems empirically to be the case that there is such a thing as default indexing, and that it leads to inferential mapping of the relevant sets of evolving substates in Ξ onto what, as the Self-as-Centre, might be considered a kind of 'imaginary' location – a form of 'out-of-sight' place 'hidden' 'somewhere' *within* the three dimensional container of 'normally' perceived space – where this location coincides with the position of the vantage point from which a person subjectively perceives themselves to be a physical agent interacting with, and driving, changing permutations of Specific Objects – including their own body – from the Centre of an infinite, eternal arena of three dimensional space.[196]

---

[191] Here 'aspects' refers to elements making up a Specific Object; for example, its edges, its sound, its position or its motion.

[192] For example, the sensation of having a general fever or of tiredness.

[193] Some desire/appetite states and emotional states – particularly when perceived as 'strong' – can be experienced viscerally (as per footnote 183), in which case D[r] must be soft indexing at least some elements of the relevant substates of Ξ to locations more widely within the Self-as-Body than just to the Self-as-Centre. 'Parching thirst' and 'gnawing hunger' are examples experienced as the eventual, incremental limit of failing to satisfy a desire/appetite, where that desire/appetite may not at first have been perceived as spatially located anywhere other than at the Self-as-Centre in the form of a 'state of mind' or inclination. Clarification of how this 'cross-over' between default indexing and soft indexing arises is provided in Part 3 as per footnote 194.

[194] Some new, clearer ideas of how emotions, including desires and appetites, are generated and perceived are proposed in Part 3. Part of the challenge in developing such ideas is that the language people usually use to describe emotions is vague, ambiguous and sometimes unhelpfully arbitrary. The presence of this lack of clarity is well described and evidenced in ref. footnote 195(3). For this reason tolerance in how terms such as 'desire', 'appetite', 'hunger' and 'emotion' are being used here is sought from readers for the time being, with a promise of clarification in Part 3.

[195] See (1) Seth, A.K. & Friston, K.J. (2016) Active Interoceptive Inference and the Emotional Brain *Phil. Trans. R. Soc. B 371*: 20160007, (2) Seth, A.K. (2012) Interoceptive Inference, Emotion, and the Embodied Self *Trends in Cognitive Sciences 17* 565-573 and (3) Feldman Barrett, L. *How Emotions Are Made* Horton Mifflin Harcourt, New York 2017.

[196] This is also the location – the 'hidden place' – where many people believe their 'mind' to reside, and which they might loosely refer to as, "somewhere inside my brain".

10.5.4 Space 2 corresponds to kappa space and Space 1 corresponds to rho space

It was proposed in Section 10.5.2 that our complete inability to imagine perceiving a boundary beyond which space or time could cease to exist – together with our inalienable conviction that everything that exists, has existed, or will exist, and that we consider to be real, must exist within one, single, infinitely extended arena of spacetime – shows that D[r]'s implementation of optimal inferential parsing of Ξ into substates must have the 3-space inference hardwired into it.

It was then proposed in Section 10.5.3 that this hardwiring implies that all substates of Ξ will be spatially indexed as a foundational step in their being optimally parsed.

The ideas of spatial indexing and optimal parsing have already, if implicitly, been used in Section 9.3 – specifically in Sections 9.3.2 and 9.3.3 – in the definitions given there of W[i], B[i] and N[i]. A re-reading of those sections in parallel with Sections 10.5.1 to 10.5.3 allows the following connections to be drawn:

A. Since CP2 in Space 2 is defined as inferentially resolved out of Ξ by balancing cueing signals driven mainly by exteroceptive and proprioceptive inputs to D[r] – where these signals are either hard or soft indexed – the connections A(1-2) can be made, and at A(3-6) it can be reiterated that:

(1) Space 2 corresponds to kappa space (κ space); which is perceived as external physical space
(2) CP2 corresponds to a momentary position in κ space, i.e. a κ locus;
(3) a κ locus is equivalent to the momentary state of the Ξ substate, W[i], the world image;
(4) a κ locus is the sum of a σ locus and a ω locus; where
(5) a σ locus is equivalent to the momentary state of the Ξ substate, B[i], the body image; and
(6) a ω locus is equivalent to the momentary state of the Ξ substate, W[i]\B[i], the environment image.

B. Since CP1 in Space 1 is defined as inferentially resolved out of Ξ by balancing cueing signals driven mainly by interoceptive inputs to D[r] – where these signals are default indexed – the connections B(1-2) can be made, and at B3 it can be reiterated that:

(1) Space 1 corresponds to rho space (ρ space); perceived as internal 'emotional space'
(2) CP1 corresponds to a momentary position in ρ space, i.e. a ρ locus; and
(3) a ρ locus is equivalent to the momentary state of the Ξ substate, N[i], the need image.

Drawing these connections allows application of the approach used to define W[i], B[i] and N[i] in Sections 9.3.2 and 9.3.3 to the wider picture made available in Section 10 above, which considers how D[r] can apply optimal parsing to inferentially recognize and define itself as D[a].

Recall from Section 10.5.1 that:

> "Because D[r]'s inference of the existence of D[a] – and how D[a] will move the recognition state – will be based on D[a]'s effects as a force generator, D[r]'s inference of the 'situation' of D[a] in terms of any momentary recognition state, Ξ, will be based on an ongoing inference that works to identify and parse out D[a] as inhering within that evolving substate in Ξ whose *motion* can be inferred to be caused by a single agent, such that that single agent can be inferred to be driving those motions of Ξ which satisfy (1) and (2) above.[197] By this means D[r] can inferentially parse out that ongoing substate of Ξ that 'contains' D[a], and from this D[r] can infer a 'position', and thence a 'perspective' or 'point of view', for D[a] as a set of relations between that evolving substate and its *environment*, which will be made up of all of the optimally parsed out substates of Ξ *outside* the substate that D[r] can infer 'contains' D[a]."

And:

> "…based on D[r]'s ability, as described above, to parse out that ongoing substate of Ξ which 'contains' D[a], D[r] should be able to map and predict those effects on W[r] in a way that differentiates between the motions of B[r] and the motions of W[r]\B[r]. This is since, with optimal parsing, and in normal circumstances, the state and motions of B[r] will always be mapped to an evolving substate of Ξ that D[r] is *able consistently* – over time – to infer *persistently* 'contains' D[a][198], with the state and motions of W[r]\B[r] then mapped to the complementary evolving substate of Ξ, with the state and motions of W[r]\B[r] then being mapped to the *complementary* evolving substate of Ξ."[199]

[197] See Section 10.5.1 for what (1) and (2) refer to.

[198] This is because as a force generator[r], D[r] will always – over time, by implementation of its generative model through active inference – be moving the state of B[r] to get to Positions 2 and 1, and will therefore always, over time, by implementation of its recognition model through optimal parsing, map the state and motion of B[r] onto an evolving substate of Ξ that it can infer 'contains' D[a].

[199] That is to say, with the state and state trajectory of W[r]\B[r] being mapped to all of the optimally parsed out substates of Ξ that D[r] is *unable* consistently – over time – to infer *persistently* 'contain' D[a]. As described previously, these substates can be considered to make up D[a]'s phenomenal 'environment'.

Together with (A) and (B), these proposals imply that the substates of Ξ which map the state and motions of B[r], and which D[r] can infer 'contain' D[a], are B[i] and N[i]. This is since these will be the substates of Ξ whose motions D[r] will almost always – and will consistently over time – be able to infer are *caused* by the single force generator and primary agent D[a]. Equally, it can be proposed that the substate of Ξ which maps the state and motion of W[r]\B[r] is W[i]\B[i], since this substate of Ξ will be composed of those substates whose motions D[r] will be *unable* sustainably over time to infer are being caused by D[a].[200]

Moreover, by this reasoning – and pursuant to that already provided in Section 10.5.1 – when D[r] comes to experience itself from the '*point of view'* of D[a], and *as being* D[a] – based on an ongoing inference that it *is* D[a] – it will come to experience itself as an *inference of the driver of the motion* of B[r], which under optimal parsing it will experience as an inference of the ongoing cause of motion of the Ξ substate B[i]. By this means, D[r] will come to infer – and experience – itself as being that which moves B[i].

More precisely and completely, D[r] will come to infer and experience itself to be that entity, D[a] – which D[r] calls "I" – that acts to move the substate N[i] by moving the substate B[i] such that elements of the substate W[i]\B[i] in turn are moved, such that all of those substates – which will sum to most, if not all, of the recognition state Ξ – are moved along a path to a goal state.[201]

To assist in recognising what this means, again consider the sentence used in demonstration of the predecessor to this idea in Section 10.5.1:

> "I move the state of my physical body, and by this means I move the state of the physical world, so that I can move from an internal state of desire to an internal state of satiation."

10.6 Assembly of phenomenal worlds - Sphere world diagrams

Several of the preceding concepts can be illustrated by use of what can be called a *sphere world diagram*. Sphere world diagrams should be viewed as a practical, rule-of-thumb device to allow visualization of the way it is proposed that the 3-space inference, which provides for hard, soft and default spatial indexing of various substates of the recognition state, Ξ, will be applied by D[r] – followed by D[r] application to Ξ of further optimal inferential parsing – to assemble its own subjective, phenomenal experience of being/inhabiting a physical body embedded in, and operating on, a physical world.

Figure 11 shows a *simple* sphere world diagram.[202]

The diagram can be understood by first noting that any momentary recognition state Ξ defines a position in Ξ space, where this space is shown as the outermost area of the diagram. In the device of the diagram this momentary position[203] is shown 'unpacked' into three nested, concentric circles. These should be visualized as cross sections of three nested spherical shells.

The first, outer spherical shell is intended to represent a three dimensional space, having spherical outer and inner surfaces, and throughout which D[r] will, in the device of the diagram, distribute according to their spatial indexing all of the optimally parsed out substates that momentarily make up the environment image, W[i]\B[i].

Almost all of the optimally parsed out substates that make up W[i]\B[i] will be hard indexed to highly spatially defined positions, distributions, and motions within this spherical space such that they correspond to what a person perceives as the relatively precisely spatially located and defined Specific Characteristics[204] that make up the physical world they perceive as surrounding and interfacing with their Self-as-Body (i.e. their physical body as they perceive it, a.k.a. their body image, B[i]).

---

[200] Almost always, rather than always, because the interface between B[i] and W[i]\B[i] will, over shorter periods of time, be blurred. Sometimes force generators[r] other than D[r] will move B[r] and, more importantly, D[r] will often itself move B[r] in ways that affect substates in W[r]\B[r]. Even so, in this second circumstance the substates of W[r]\B[r] that D[r] moves through moving B[r] will be relatively short lived – they will come and go – whereas D[r]'s movement of B[r] will persist. Moreover, the vast majority of substates in W[r]\B[r] giving rise to hidden causes of perception will, at any given moment, be unaffected by output from D[r]. Of course, these assertions are based on expectations derived from considering the phenomenology of being a physical body, B[i], living and operating in a physical world, W[i], where B[i] is in interaction with W[i]\B[i], and therefore on a presumption that B[i] is indeed mapping the state and motion of B[r] and that W[i]\B[i] is indeed mapping the state and motion of W[r]\B[r], consistent with the second quoted paragraph above.

[201] Where the meaning of *goal state* is as defined in the underlined text of footnote 163.

[202] It is simple because it does not take time evolution of the recognition state Ξ into account in any explicit or well defined way. Instead it simply refers to the Ξ state being unpacked in the diagram as 'momentary'. The time evolution of Ξ is addressed in subsequent sections.

[203] The strict meaning here of 'position' – a location approximating a point – in Ξ space is given in Section 9.3.1, and see footnotes 76 and 77.

[204] Specific Characteristics – as defined in Section 10.5.2 – are the full ensemble of Specific Objects (e.g. chairs, houses, railway lines) – including their positions, spatial relations and relative motions – that at any given moment a person perceives as surrounding – and in some cases interacting with – their Self-as-Body.

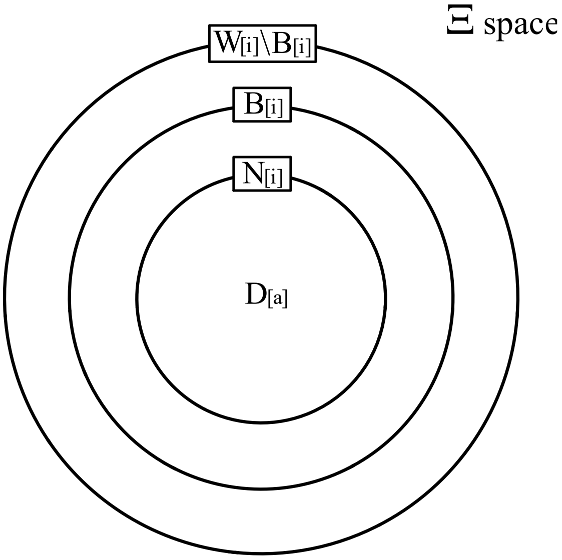

Figure 11

Figure 11 shows a simple sphere world diagram. The diagram shows an 'unpacked' position – i.e. an unpacked recognition state, Ξ – in Ξ space. Label this unpacked position a *sphere world*. The diagram is a rule-of-thumb device designed to illustrate how D[r]'s application of the 3-space inference and its further optimal parsing of the recognition state Ξ can operate to generate an enduring, *overall pattern* of substates of Ξ that can readily be visualized in terms of how a person actually subjectively experiences and perceives themselves to be situated in a three dimensional phenomenal world. This overall pattern is illustrated in the diagram as three nested circles – to be visualized as three nested spherical shells – where each shell constitutes a three dimensional space bordered by an inner and outer spherical surface. These shells illustrate how, through application of the 3-space inference and further optimal parsing of Ξ , our D[r] constantly over time partitions the three dimensional world we each experience – when our D[r] places itself into identity with its D[a] – into: an outermost volume populated by the substates that make up W[i]\B[i]; which fully surrounds and interfaces with an intermediate volume populated by the substates that make up B[i]; which in turn fully surrounds and interfaces with an inner volume populated by the substates that make up N[i]. Through application of the 3-space inference all of the optimally parsed out substates of Ξ will be spatially indexed – hard, soft or default – to one or another of these volumes or surfaces, and will all be perceived from the spatial perspective of D[a], located at the centre of these nested shells. From this perspective D[r], as D[a], will perceive (1) the substates of N[i], which populate the inner shell (largely default indexed), as desire/emotion arising in a place coinciding with its own location, (2) the substates of B[i], which populate the intermediate shell (largely soft indexed), as the parts and sensations that make up its physical body and are closest to its own – i.e. D[a]'s – location while still being broadly spatially differentiable from it and from each other and, (3) the substates populating W[i]\B[i] within the outer shell (largely hard indexed) as being furthest from it and as objects and entities in its physical environment that are largely highly spatially differentiable from it, from its physical body and from each other.

Positions a person perceives to be at the maximum perceivable physical distance from their Self-as-Body – i.e. those literally on the spatial horizon of the external physical world as they perceive it – can be considered to be sitting in the two dimensional spherical outer surface of the outermost spherical shell.

Nested within the outer shell, and interfacing with it at its inner two dimensional spherical surface, the second and intermediate spherical shell also represents a three dimensional space, with spherical outer and inner surfaces, throughout which D[r] will, in the device of the diagram, distribute according to their spatial indexing all of the optimally parsed out substates that momentarily make up the body image, B[i] (i.e. a person's physical body as they perceive it, their Self-as-Body).

Many of the optimally parsed out substates that make up B[i] will be soft indexed to only roughly defined positions and distributions within this space such that they correspond to what a person perceives as the relatively poorly spatially located and defined physical sensations that arise within their Self-as-Body. Such perceptions will include some aches and pains, some sensations of physical illness such as fever, digestive sounds, feelings of fullness, and so forth. They will also include broadly felt perceptions of pushes and pulls.

Some substates that will nevertheless have a degree of hard indexing will be those which D[r] optimally parses out to being in or very near the two dimensional spherical interface between W[i]\B[i] and B[i]. These substates will be perceived by a person to be located at the interface between the exterior and interior of their Self-as-Body and they will perceive these as being at the outer surface of their Self-as-Body. For example, if you pick up a pen or some other suitable Specific Object and, as you look at it, you move it around between your fingers, you will perceive your fingers as they move the pen to be highly defined in space. You will do so by visual means but also, at the fingertips, by tactile means together with proprioceptive sensibility.[205]

The third, inner spherical shell in the diagram is nested within, and interfaced with, the intermediate shell and represents a final space within whose *nominal* three dimensions D[r] will, in the device of the diagram, distribute according to their spatial indexing the optimally parsed out substates that make up the need image, N[i].

Because these substates are default indexed they can be considered to be distributed in a *spatially undifferentiated way* – i.e. evenly – throughout this innermost space. In the moment covered by the sphere world diagram shown in Fig. 11, these substates will correspond to what a person perceives as a spatially undifferentiated momentary mix of desire/emotion[206] located within their Self-as-Body, at what has been described above[207] as the Self-as-Centre.

Certain substates may arise that D[r] soft indexes to locations at the two dimensional interface between N[i] and B[i]. These substates will manifest when a person is feeling one or more desires or emotions very strongly such that they partially experience these as physical sensations, for example as sensations of longing or angst, perceived as being at some broadly defined location within the Self-as-Body, for example in their 'heart' or in the 'pit' of their stomach.[208]

Shown at the centre of the diagram is the vantage point or 'point of view' – described earlier as the Centre[207] – from which all of the optimally, inferentially parsed out substates of Ξ discussed above are perceived by a person in terms of their contents, distributions, and relations to each other and – crucially – *in terms of their relations to* D[r]'s inference of itself, D[a].

As a whole, Fig. 11 is designed to provide means to visualize how D[r]'s implementation of the 3-space inference, combined with its further optimal parsing of the recognition state Ξ, can work to generate a relational vantage point for D[r] which, when D[r] inferentially identifies itself as D[a], allows D[r] *through the proxy of D[a]* relationally to 'look out onto' the state and state trajectory of B[r] and W[r]\B[r] through inference of itself as *being* D[a] 'looking out onto' a relational distribution, in a three dimensional space, of the optimally parsed out Ξ substates and substate trajectories that together make up N[i], B[i] and W[i]\B[i].

D[r] can achieve this by mapping the force generator D[a] *to* – and by inferring itself to *be* D[a] *at* – a vantage point at the centre of this relational distribution, which constitutes D[r]'s *own overall mapping* of the state and state trajectory of W[r]\D[r] – a mapping in the form of a rich ensemble of substates and substate trajectories that it is optimally parsing out to form the spatially distributed contents of N[i], B[i] and W[i]\B[i].[209]

---

[205] The interior of a person's mouth and nose are interesting examples of sets of relatively hard indexed substates – some, for example, perceived as tongue[i] or teeth[i] – at the interface between W[i]\B[i] and B[i]. Here, while visual input plays virtually no role in the inference of these substates, tactile input from the tongue, as well as olfactory and proprioceptive inputs play crucial roles. At a more general level, bearing such examples in mind, the expected degree of spatial indexing associated with each region of B[i] can be expected roughly to correlate to the relative areas accorded each part of the body and its surface in diagrams such as the 'sensory homunculus' shown in Fig. 12.18 in Patestas, M.A. and Gartner, L.P. *A Textbook of Neuroanatomy Ed.2* Wiley & Sons, Hoboken New Jersey USA 2016.

[206] A way to conceptualise how a person might perceive such a spatially undifferentiated mix of substates of N[i] is to imagine – as a metaphor – that such perception might at any given moment parallel how a person perceives a simultaneous sounding of musical chords, where each chord is spatially undifferentiated from the others and is itself made up of a spatially undifferentiated mix of notes. This metaphor is further developed and drawn upon extensively in subsequent discussion of emotions in Section 11.3.2 of Working Note B – Part 3, including as a means of devising a nomenclature for the description of simple emotional states and how these can be related to decision-making.

[207] Where the Self-as-Centre and the Centre are defined in Section 10.5.2.

[208] See footnote 193 for some relevant elaboration.

[209] The *mechanism* by which D[r] can inferentially sustain that it *is* D[a] and, crucially, that it is actually *in* – and consciously *'looking out from'* – a vantage point at the centre of a map of its own making, is at the heart of how subjective states can arise. Proposals in relation to this mechanism have already been described in Section 9.4.5 of Part 1, which refers to and contextualises the notes, *How is Free Will Possible?* and *The Construction of Phenomenal Time*.

Taking this into consideration, the sphere world diagram shown in Fig. 11 can be understood to show the general structure of the overall map of the hidden state and state trajectory of B[r] in W[r] that D[r] builds and maintains moment-by-moment, inter alia by hierarchically processing its inputs to generate moment-by-moment recognition states Ξ which, through the 3-space inference and optimal parsing, it renders into the *enduring general form* depicted in the Fig. 11 as the nested spherical shells N[i], B[i] and W[i]\B[i].

In other words – just as the map of a city is drafted onto a continuous sheet of paper – D[r] applies the 3-space inference to draft its dynamic map of the state and state trajectory of B[r] in W[r] onto a continuous three dimensional space, which – at a primary structural level – it then almost always[210] inferentially partitions into discrete but entirely contiguous, constituent spaces – shown as the shells in the diagram – into each of which it distributes the evolving substates comprising N[i], B[i] and W[i]\B[i].[211]

The diagram in Fig. 11 is preliminary because it does not explicitly take the time evolution of the recognition state Ξ into account. Over time the momentary substates of N[i], B[i] and W[i]\B[i] that populate the dynamic map being generated by D[r] will change, but the map's primary structure – shown in the diagram as the three nested shells – will endure.[212]

Subsequent sections of this note explore how sphere world diagrams can be used to visualize the time evolution of the recognition state Ξ as an inferentially generated map of the time evolution of the hidden state W[r]/D[r]. This time evolution can be visualized as the movement of a *sphere world* along a path in Ξ space.[213] However, before taking this step it may aid understanding to demonstrate application of sphere world diagrams in more concrete terms.

To do this recall that it is being proposed that a person's D[r] – including yours and mine – will operate inferentially to resolve out and map its primary force generator D[a] *to* – and then infer itself to *be* that D[a] *at* – and to be 'looking out from' – a vantage point at the centre of its *own overall mapping* of the state and state trajectory of W[r]\D[r] as rendered into the overall form of the Ξ substates parsed out as N[i], B[i] and W[i]\B[i].

This means that as a D[r] you will experience yourself as a D[a], at the centre of a *sphere world* as shown in Fig. 11 and – *as that D[a]* – you will experience yourself as being at the Centre:[214]

- which you will spatially experience as 'behind your eyes' – at the spatial origin of a three dimensional world – and as *co-located* with what you perceive as your current desires/emotions – reflecting the contents of the Ξ substate members[215] of your need image, N[i] – which you will experience most intimately as your Self-as-Centre (inner shell);

- from which position you will 'look out' to perceive at quarters closest by distance to your Self-as-Centre, but intimately encompassing it – and unlike it, spatially distributed and differentiated within that three dimensional world – the parts and sensations of your inner and outer Self-as-Body (intermediate shell), reflecting the Ξ substate members of your body image, B[i]; and

- from which position you will also 'look out' to perceive at quarters spatially further from your Self-as Centre and Self-as-Body, but encompassing them within that three dimensional world, a highly spatially differentiated physical environment (outer shell) – reflecting the Ξ substate members of your environment image, W[i]\B[i] – in the form of an ensemble of parts – i.e. Specific Objects and their relations and motions – that go into making up the Specific Characteristics of the external physical world as you perceive it.

[210] In rare, quite abnormal circumstances this may not hold. For example, some strong psychedelics may break this parsing down whilst continuing to allow a level of conscious presence to persist. A person may thereby experience themselves as being 'at one with the world' in senses such as that described in Barrett, F.S. and Griffiths R.R. (2018) Classic Hallucinogens and Mystical Experiences: Phenomenology and Neural Correlates *Current Topics in Behavioural Neuroscience 36* 393-430.

[211] In a sense a sphere world diagram can itself be considered a map – a map of an *enduring structure* inherent in the continuously evolving map D[r] builds, updates *and maps itself into*, of the state and state trajectory of B[r] and W[r]\B[r] in the form of the optimally parsed out substates that make up N[i], B[i] and W[i]\B[i].

[212] This can be verified empirically, at the level of phenomenal experience, by once again observing that in normal waking circumstances we *invariably* experience ourselves as spatially located *within* a physical body – at the Centre of a spatially distributed world – in immediate co-location at that Centre with a source of perceived desires and emotions – a Self-as-Centre – which we experience as interfacing at a deep, internal level with our physical body, which we experience as a discrete entity – a Self-as-Body – which we in turn experience as spatially located *within*, but interfacing and interacting with, a surrounding physical world populated by spatially distributed Specific Characteristics.

[213] Noting again (see Fig.11) that what is here being called a 'sphere world' is a label for an unpacked position in Ξ space, where the idea of such an unpacked position is a conceptual construct designed to enable description of a person's experienced moment of a phenomenal world.

[214] Where the Centre is as defined in Section 10.5.2.

[215] Where the term 'members' is used here recalling from Sections 9.3.2 and 9.3.3 of Part 1 the original definitions of N[i], B[i] and W[i]\B[i] as sets which are substates of Ξ, and will have certain substates of Ξ as their members.

At an ‘operational level’, through inferentially identifying itself as D[a], your D[r] will perceive itself – i.e. *you* (as D[r]) will inter alia perceive *yourself* (as D[a]) – to be:

Acting to move

1. the feelings of desire you perceive as sensations arising at the most intimate, innermost part of yourself – at the ‘heart’ of your spatial centre – which you experience as your Self-as-Centre, and:

   (a) whose insides you can perceive as an ensemble of spatially undifferentiated, default indexed substates of Ξ, making up your need image, N[i], parsed out across the interior of the inner spherical shell of the diagram

   (b) and areas of whose outer surface you can perceive as a poorly spatially differentiated distribution of sensations of some of your needs and desires – for example palpable forms of hunger or thirst[216] – which correspond to substates of Ξ soft indexed to the outer surface of the inner spherical shell of your sphere world, corresponding to where the ‘centre’ of your desires/emotions ‘meets’ your physical body, especially in broad locale of your viscera (e.g. as with feelings you perceive to be located in your ‘gut’ or your ‘heart’)

by moving

2. your physical body – which you perceive *as encompassing* the intimate, innermost part of yourself described at (1) – and which you experience as your Self-as-Body, and:

   (a) as a space-filling, three dimensional, physical, Specific Object whose movements you – as the primary force generator D[a] – control *directly* – i.e. ‘*own*’;

   (b) whose insides you can perceive as a relatively blurred spatial distribution of internal sensations – which correspond to the soft indexed substates of Ξ making up your body image, B[i], parsed out across the interior of the intermediate spherical shell in the diagram

   (c) and areas of whose outer surface and near outer surface you can perceive as highly spatially differentiated distributions of mainly visual, but also other, particularly tactile and proprioceptive, sensations – which correspond to substates of Ξ hard indexed to the outer surface of the intermediate spherical shell in the diagram, corresponding to where your physical body meets your physical environment – especially at your mouth parts and your hands

and, by that means, acting to move

3. your physical environment, which you perceive as encompassing your physical body, and:

   (a) as taking up all available three dimensional space other than that occupied by your physical body, out to a perceived horizon,

   (b) populated by Specific Objects other than your physical body – which correspond to hard and some soft indexed substates of Ξ, making up your environment image, W[i]\B[i], parsed out to the interior and sometimes to the surfaces of the outer spherical shell of the diagram,

   (c) whose movements you can control only *indirectly* by interacting with them through movements of your physical body

until the movements you achieve in your physical environment, through moving your physical body in interaction with it, achieve a situation where the feelings of desire you have been perceiving are satisfied.

[216] An identification of just which substates of Ξ will be optimally parsed to the outer surface of the innermost shell is provided in Working Note A – Part 3. In the meantime readers are again asked to bear in mind the contents of footnote 194 when encountering words such as ‘need’, ‘desire’, ‘appetite’, ‘emotion’, ‘hunger’ and so forth.

Figure 12 illustrates the general case of movement through Ξ space via a *path*, Þ, from Ξ^a – a person's recognition state, at some given time, t = a – to a recognition state, $\Xi^{(a+n)}$, at a later time t = (a+n).[217]

Figure 13 illustrates the specific example provided on the previous page using terms from Section 10.5.1, which is the source of the example.

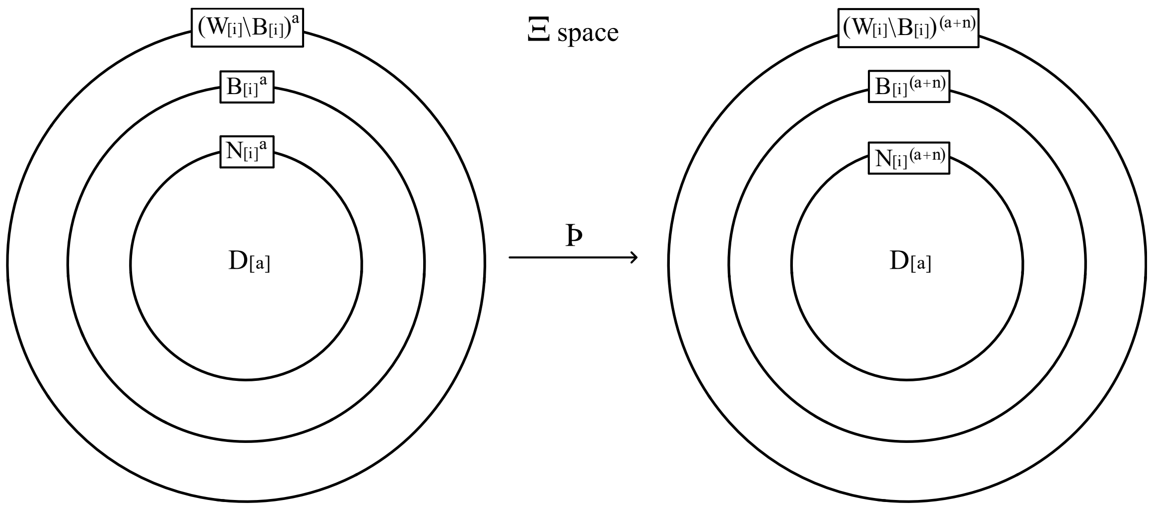

Figure 12

Figure 12 illustrates movement along a path, Þ, through Ξ space from an initial recognition state Ξ^a at time t = a to a recognition state $\Xi^{(a+n)}$ at some later time t = (a+n), where these recognition states have been unpacked into their respective sphere worlds.

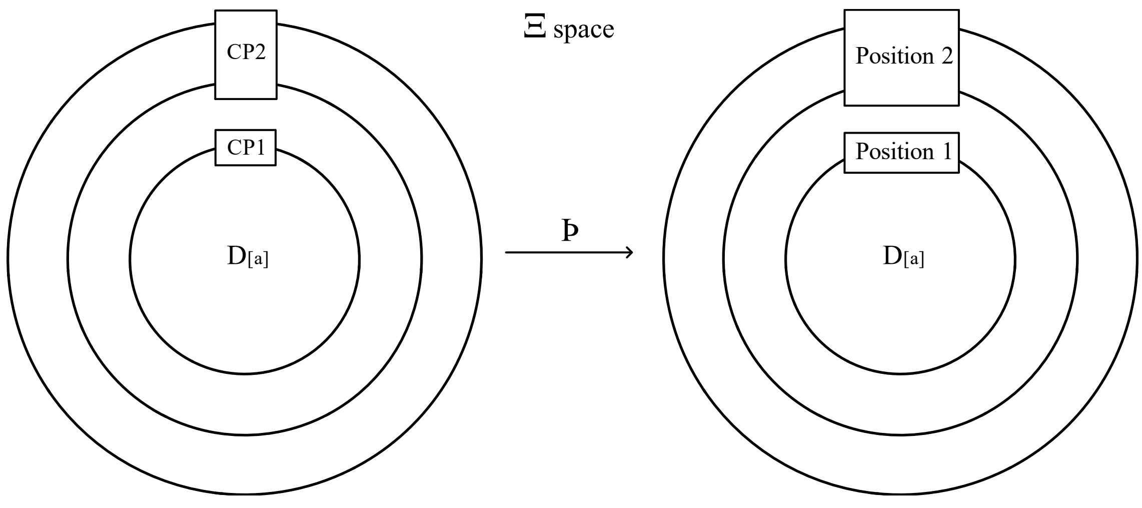

Figure 13

Figure 13 illustrates the example provided on the previous page where, using terms from Sections 10.5.1 and drawing upon Section 10.5.4, the initial recognition state Ξ^a at t = a can be expressed as a position CP2 + CP1 in Ξ space – noting that, in relation to Fig. 12, CP2 is equivalent to $B[i]^a + W[i]\backslash B[i]^a$ and CP1 is equivalent to $N[i]^a$ – and where the final recognition state $\Xi^{(a+n)}$ can be expressed as Position 2 + Position 1 in Ξ space – noting, in relation to Fig. 12, that Position 2 can be set as equivalent to $B[i]^{(a+n)} + W[i]\backslash B[i]^{(a+n)}$ and Position 1 can be set as equivalent to $N[i]^{(a+n)}$. In Fig. 13 the final recognition state, equivalent to Position 2 + Position 1, is a goal state (as per Section 10.5.1, footnote 163, as underlined).[218]

[217] For consistency of notation with respect to future work, say that the movement illustrated in this example is a movement which has taken place in a person's past.
[218] Note also for cross referencing purposes that the following expressions are all equivalent: B[i] + W[i]\B[i], {B[i] + W[i]\B[i]}, B[i] in W[i], and {B[i] in W[i]}, which all describe the parsed form of W[i], where W[i] = {B[i] + W[i]\B[i]}.

10.7 Adding time to space – the structure and contents of a 'now'

To conclude Part 2 of this note and set the stage for Part 3, some ideas first raised in Section 9 of Part 1 will now be revisited and augmented. These ideas go to the subjective experience of time.

Beyond perceiving ourselves as *spatially positioned* at a Self-as-Centre within a Self-as-Body 'looking out onto' a physical environment of Specific Characteristics – where we invariably perceive those contents of perception to be relationally positioned and/or distributed in a single arena of three dimensional space – we also invariably perceive ourselves and our spatially distributed environment to be *temporally positioned* at a *'now'* that is moving forward in A-series time.[219]

In terms of the ideas being developed in this note, this means that

- in operating inferentially to resolve out and map its primary force generator D[a] *to* – and to infer itself to *be* that D[a] *at* – a vantage point at the centre of its own overall mapping of the state and state trajectory of W[r]\D[r][220]

a person's D[r] comes not only to 'look out from' a *spatial centre* within that mapping into a spatially distributed world, but also comes to 'look out from' a *temporal centre* within that mapping – i.e. from a 'now' – into a temporally distributed world – i.e. into a world with a past and a future.

If a person's subjective experience of being in a 'now' is accepted as phenomenologically valid, and it is also accepted that the recognition state Ξ will at any given time provide for the overall contents of a person's phenomenal experience, then to take account of a subjective 'now' the simplest way explicitly to describe this would appear to be to partition any given Ξ extant at some arbitrary time, t = a, into two components, or *aspects*:

- one aspect encoding what a person will perceive to be the phenomenal contents of their past, up to and including their most recently experienced phenomenal moment, Ξ^{a}, and
- the other aspect encoding what a person will perceive to be the phenomenal contents of their future, including their most immediately *predicted* next phenomenal moment, $\Xi^{(a+1)*}$

where both aspects can be considered to come together to form what is subjectively perceived by the person as the contents of their 'now' at t = a.[221]

To further underpin this approach, and reinforce the case made in Section 9.4.1 of Part 1, it is helpful to recognize that from a phenomenological standpoint what we experience as 'now' is invariably *perceptually highly distinct* from how we recall our past or envisage our future. Specifically, as our experience of being in the world unfolds, the preponderance of what we perceive is in our 'now', which exclusively and inexorably is where we experience ourselves to be *present* and to 'look out from' into time. Any past moment we wish to envisage will have a sense of 'being recalled' and will be perceived as being 'less real' than our 'now', while any future moment we wish to envisage will have a sense of 'being anticipated' and will also be perceived as being 'less real' than our 'now'.

Moreover, any specific remembered past moment, or envisaged future moment, may be *more* or *less* strongly perceptually invoked, depending on choice or circumstance, but our ongoing conscious perception of being positioned in the world at a single and entirely *real* 'now' – as that 'now' inexorably 'moves forward' in time – is not subject to choice or circumstance.

If our 'now' is never perceived exclusively as one moment within a series of recalled past moments or as one moment within an array of envisaged future moments, it must be situated at the *interface* between such series. There is no other option, no other 'place' for it to be. Nor should it seem controversial to say that from a phenomenological standpoint a person's 'now' must be situated at the *interface* between what they can envisage through memory to have been the contents of their past and what they can envisage through affordance (see below) to be the potential contents of their future.

---

[219] A-series time is described in McTaggart, J.M.E. (1908) The Unreality of Time. *Mind 17*: 457–73. And for further detail relevant to discussion here, see Section 9.4.1 of Part 1 as well as *The Construction of Phenomenal Time* at https://teledyne.com/time.pdf.
[220] Where that mapping is in the contents of the optimally parsed out substates of Ξ that make up N[i], B[i] and W[i]\B[i].
[221] NB: This idea has already been described and applied in Section 9.4.1 on Part 1, where it was proposed that for any time, t = a, the recognition state Ξ should be expressed as $\{\Xi^{a} + \Xi^{(a+1)*}\}$.

Essentially then, what a person momentarily experiences as the phenomenal contents of a 'now' can be described for some arbitrary time, t = a, as a *composite* of the past-facing *surface* of such an interface, in the form of Ξ^a, and the future facing *surface* of such an interface, in the form of $\Xi^{(a+1)*}$.

Together, these two components of a person's phenomenal experience at time, t = a – i.e. their subjective experience at t = a of being a Self-as-Centre in a Self-as-Body in a three dimensional spatial environment made up of Specific Objects in relative motion – can be expressed as $\{\Xi^a + \Xi^{(a+1)*}\}$, where $\{\Xi^a + \Xi^{(a+1)*}\}$ will be a composite of the three dimensional 'face' of a higher dimensional 'block' of all *memory* of their phenomenal past, Ξ^a, and the three dimensional 'face' of a higher dimensional 'block' of all *affordance* of their phenomenal future, $\Xi^{(a+1)*}$. Figure 14 illustrates this description of a person's experience of being in a 'now' at t = a.

Note that Fig. 14 parallels Figure 5 in the note, *The Construction of Phenomenal Time* (TCPT),[222] although TCPT and its figures address only the $\{\{B[i] \text{ in } W[i]\}^a + \{B[i] \text{ in } W[i]\}^{(a+1)*}\}$ substates of $\{\Xi^a + \Xi^{(a+1)*}\}$. An account of the relationship between the former and latter – and of how both may be applied to understanding a person's subjective experience of consciousness as a series of 'now's – is provided in Section 9.4.5 of Part 1.

Note also that here, and implicitly in TCPT, the term 'affordance' is meant in a way related to, but distinct from, the meanings it has variously been given in its usages in the literature.[223] Broadly put, those usages rest upon the idea that inherent in a person's perception of a Specific Object is a perception of the 'opportunities for action' – or 'affordances' – presented by that object. To borrow some examples,[224] a person may perceive a wall as grey and steep, but also as *climbable*, a teapot as *grippable*, or a flying baseball as *catchable*, where in much of the relevant literature the italicized property is considered a perceived affordance.

For the purposes of this note however, it is proposed that this widely used,[224] relatively uncontroversial idea of affordances – call these *proverse affordances*[225] – be extended to encompass a wider range of perceptions, to include:

1. *inverse affordances*, defined here as perceptions of what Specific Objects could do *to* the perceiver, as opposed to what the perceiver could do *with* those Specific Objects – for example, a sun-heated steering wheel may be perceived as *grippable* (proverse affordance) but also as *'sting-hand-able'* (inverse affordance), or an uncovered poisonous spider perceived as '*pull-hand-away-able*' (inverse affordance);

2. *ensemble affordances*, defined here as perceptions of what the perceiver could do with (or have done to them by) an ensemble of Specific Objects, for example a knife and fork, drum kit, pile of jigsaw puzzle pieces, a set of rock climbing holds,[226] or a swarm of angry hornets; and most importantly

3. *block affordance*, defined here as the *array* of proverse, inverse and ensemble affordances that amount to the *full suite* of the perceptions a person has of what they can do, or can have done to them, across *all* of the Specific Characteristics[227] that they perceive to be making up their current, overall physical environment.[228]

For the purposes of this Working Note, the meaning given to the term affordance is that provided at (3) above, for block affordance. This is the intended meaning of the term affordance as it is used above with reference to a person's perception of their phenomenal future, $\Xi^{(a+1)*}$, in Figure 14, and throughout TCPT.

A crucial utility of this way of conceiving of affordances will become clear in ideas to do with imagination, memory, decision-making and emotion proposed in Working Note A – Part 3.

---

[222] See https://teleodyne.com/time.pdf.
[223] For example, as it is used in McClelland, T. (2019) Representing Our Options: The Perception of Affordances for Bodily and Mental Action. *Journal of Consciousness Studies 26*, 155-180 and in the references cited therein.
[224] Ibid.
[225] This is a proposed new usage of the rarely used word 'proverse', which is used here simply to distinguish between the meaning of the term 'affordance' as it is usually applied – where all such affordances are defined here as proverse – and the meaning of 'inverse affordances' as defined in the text above.
[226] Where such holds would be along lines described in Pezzulo, G. et al. (2010) When Affordances Climb into your Mind... *Brain and Cognition 73* 68-73.
[227] Specific Characteristics – defined in Section 10.5.2 – are the entire ensemble of Specific Objects (e.g. chairs, houses, railway lines, apples) – and their positions, spatial relations and motions – that at any given moment a person perceives as surrounding – and, as applicable, interacting with – their Self-as-Body.
[228] For other approaches that extend the concept of affordance to apply to 'whole of perceived environment' see for example Pezzulo, G. & Cisek, P. (2016) Navigating the Affordance Landscape: Feedback Control as a Process Model of Behaviour and Cognition *Trends in Cognitive Sciences 20* 414-424, Thill, S. et al. (2013) Theories and Computational Models of Affordance and the Mirror Systems: An Integrative Review. *Neuroscience & Biobehavioural Reviews 37*, 491-521 and Ch. 6 in Clark, A. *Surfing Uncertainty: Prediction, Action and the Embodied Mind* Oxford University Press, New York, USA 2016.

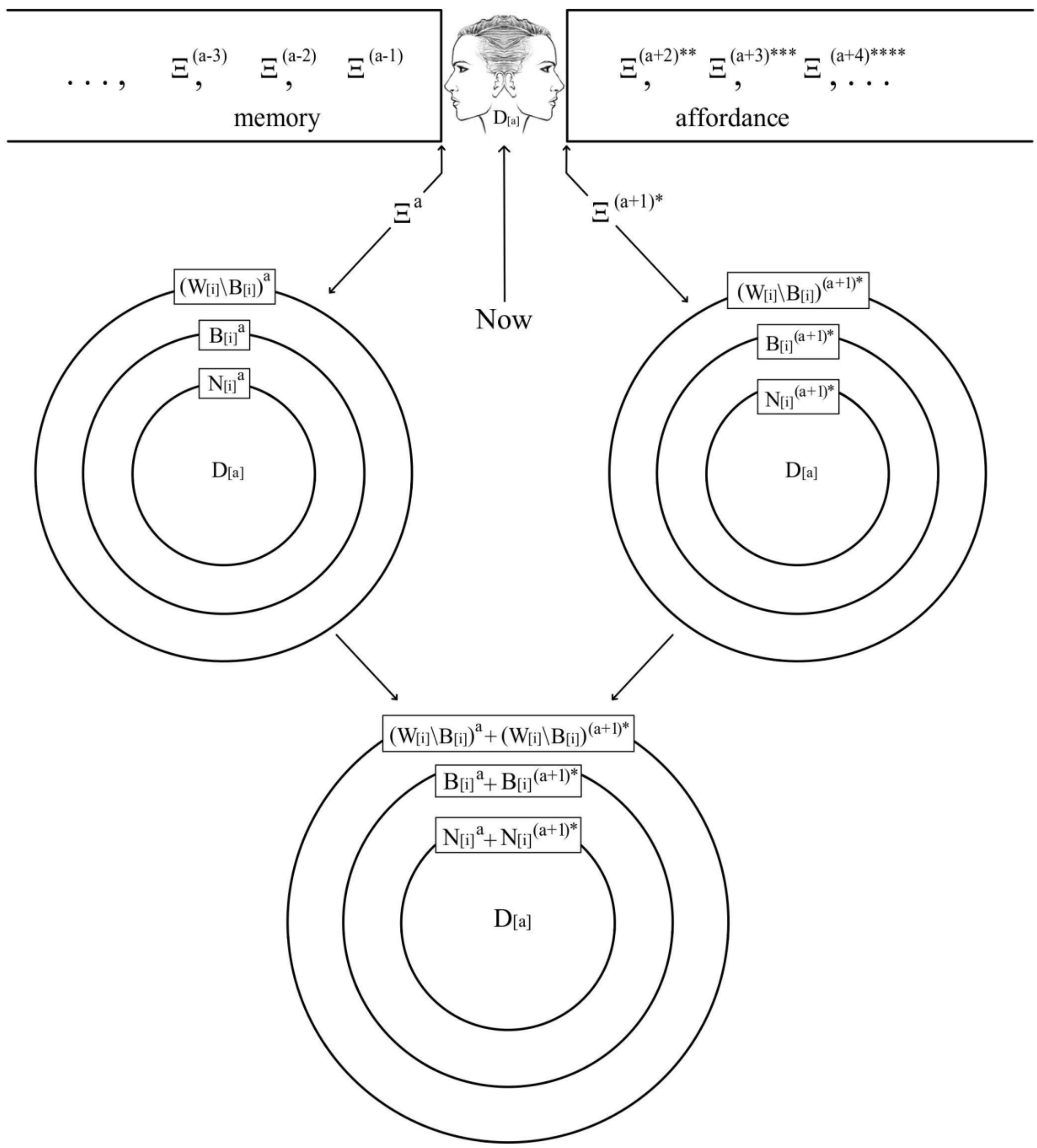

Figure 14

Figure 14 illustrates the phenomenal structure of a subjective 'now' extant for a D[a] at some arbitrary time, t = a. At the top of the diagram the three dimensional 'faces' of memory and affordance – Ξ^{a} and $\Xi^{(a+1)*}$respectively – are represented by the vertical edges to the left and right of the Janus profile, where that profile represents the seat of D[r]'s inference of itself as D[a], looking out into time from D[a]'s vantage point at the *centre* – the 'now' – of the fourth, temporal, dimension in D[r]'s overall mapping of the state and state trajectory of W[r]\D[r]. Here then, the two faces Ξ^{a} and $\Xi^{(a+1)*}$ can each be considered to map a *state* of W[r]\D[r], encoded as a *static* three dimensional distribution of optimally parsed substates of Ξ – i.e. encoded as substates parsed into the contents of 'spatial containers' W[i]\B[i], B[i] and N[i] – with the *state trajectory* of W[r]\D[r] then mapped through forming both of those static maps into a composite able to encode motion in the state of W[r]\D[r] as motion among those three dimensional distributions. This is illustrated in the centre of the figure where the faces Ξ^{a} and $\Xi^{(a+1)*}$ are shown as parsed into two respective sphere worlds. These can be thought of as the result of rotating each of the faces, Ξ^{a} and $\Xi^{(a+1)*}$ respectively, through 90 degrees around a vertical axis. The bottom of the diagram puts these two sphere worlds into a composite designed to illustrate D[r]'s perspective – as it inferentially experiences itself as D[a] – of looking out into a three dimensional spatially distributed environment *in motion* at the 'now' (i.e. at the inflection point *i* in the recognition cycle) for time, t = a. The increasing number of asterisks going to the right in the top segment of the diagram signifies increasing levels of uncertainty/indeterminacy in the predictions made in D[r] as it applies its generative and recognition models to 'look further' into – i.e. to predict – its phenomenal future. Fig. 14 is a parallel expression and extension of the idea illustrated in Figure 5 of the note *The Construction of Phenomenal Time* (see text).

The account provided above is consistent with the approach adopted in Section 9.4 in terms of a recognition cycle where – drawing on Fig. 9 and related text, plus observations made in Section 10.1 – time evolution of Ξ from some time, t = a (where t is in unit beats of the recognition cycle) has previously been expressed as the series:

$$\{\Xi^{a} + \Xi^{(a+1)*}\},\ \{\Xi^{(a+1)} + \Xi^{(a+2)*}\}, \ldots, \{\Xi^{(a+n)} + \Xi^{(a+n+1)*}\} \qquad \text{R1}$$

Note here that – as reflected in the sphere world diagrams shown in Fig. 14 – R1 can be considered an expression of the sum, R3 + R4 + R5, of the optimally parsed series:

$$\{N[i]^{a} + N[i]^{(a+1)*}\}, \{N[i]^{(a+1)} + N[i]^{(a+2)*}\}, \ldots, \{N[i]^{(a+n)} + N[i]^{(a+n+1)*}\} \qquad \text{R3}$$

$$\{B[i]^{a} + B[i]^{(a+1)*}\}, \{B[i]^{(a+1)} + B[i]^{(a+2)*}\}, \ldots, \{B[i]^{(a+n)} + B[i]^{(a+n+1)*}\} \qquad \text{R4}$$

$$\{(W[i]\backslash B[i])^{a} + (W[i]\backslash B[i])^{(a+1)*}\}, \{(W[i]\backslash B[i])^{(a+1)} + (W[i]\backslash B[i])^{(a+2)*}\}, \ldots, \{(W[i]\backslash B[i])^{(a+n)} + (W[i]\backslash B[i])^{(a+n+1)*}\} \qquad \text{R5}$$

And note also that R4 + R5 can, as in Sections 9.4.5 and 10.1, be expressed in the form:

$$\{\{B[i] \text{ in } W[i]\}^{a} + \{B[i] \text{ in } W[i]\}^{(a+1)*}\}, \{\{B[i] \text{ in } W[i]\}^{(a+1)} + \{B[i] \text{ in } W[i]\}^{(a+2)*}\}, \ldots, \{\{B[i] \text{ in } W[i]\}^{(a+n)} + \{B[i] \text{ in } W[i]\}^{(a+n+1)*}\} \qquad \text{R2}$$ [229]

Interim foreword to Part 3 – the Explanatory Gap and 'What it is like'

Part 1 of Working Note A set out to recast in terms of the predictive processing paradigm the highly speculative conceptual frameworks presented in its predecessors, the Essay, The Introductory Summary, *How is Free Will Possible?* and *The Construction of Phenomenal Time* (TCPT).[230]

Part 2 has used ideas developed in Part 1 as a basis to address in more detail the means by which a proposed processor D[r] within a brain[r] can come to infer that its own physical agency – or 'physical self' – is an entity[i] 'within' a B[i] – i.e. that it is a primary agent D[a] – and, with this, how a D[r] can come to subjectively, consciously experience itself as a person (i.e. as a D[a]) looking out from, and operating from, a centre/point of view onto what is described here as a phenomenal space populated with Specific Objects, and from a centre/point of view – called a 'now' – onto what has been described here, and in TCPT, as phenomenal time.[231]

In undertaking this task, Part 2 has sought to go some way towards describing how a D[r], within a brain[r] – itself within a B[r] embedded in W[r] – can have an ongoing subjective, conscious experience of itself within the realms of 'what it is like' to experience being a person in same the way that you or I have such an experience.[232]

To then go more directly to the question of 'what it is like': If it is accepted that a D[r] – through operation of all of the *noumenal-level* arrangements and processes described above – can inferentially experience itself as a D[a], and that the form and general contents of that experience will be phenomenal experience, including in the overall form illustrated through use of sphere world diagrams, then it can be proposed at the most general level that:

- When a D[r] inferentially experiences itself to be a D[a], this is 'what it is like' to be an operating D[r]

where experiencing itself to be a D[a] will be to experience being a 'self'[233], and will – from the perspective of such a self – encompass experiences of 'what it is like' to have all of the conscious, subjective, phenomenal experiences that a person has.

---

229 Given that, from Section 9.3.2, the term {B[i] in W[i]} is equivalent to the term {B[i] + W[i]\B[i]}.

230 See https://teleodyne.com.

231 For a sense of the conceptual task undertaken, and whether the conjectures made here in the course of seeking to describe the nature of phenomenal space are radical or conservative, readers may wish to review for example: McGinn, C. (1995) Consciousness and Space pp149-163 in Metzinger, T. (Ed.) *Conscious Experience* Ferdinand Schöningh, Paderborn Germany; Fingelkurts, A.A. et al. (2010) Natural World Physical, Brain Operational, and Mind Phenomenal Space-Time *Physics of Life Reviews 7* 195-249; and Chapter 9 in Revonsuo, A. *Inner Presence: Consciousness as a Biological Phenomenon* MIT Press, Cambridge MA, USA 2006. The insights of Adam Safron are also relevant, e.g. as provided in the first three sections of Safron, A. (2021) Integrated World Modeling Theory (IWMT) Expanded: Implications for Theories of Consciousness and Artificial Intelligence (preprint) DOI: 10.31234/osf.io/rm5b2 and at https://www.researchgate.net/publication/352582207.

232 Recalling the challenge identified in Section 10.1 in relation to (1) Nagel, T. (1974) What is it Like to be a Bat? *The Philosophical Review 83* 435–450 and bearing in mind the observations made in (2) Papineau, D. (2011) What Exactly is the Explanatory Gap? *Philosophia 39* 5-19.

233 Where, more precisely, being a D[a] can be viewed as being a 'self-model' of {D[r] in W[r]} run at the noumenal level by D[r] within D[r], by means described in the body of the text, and where D[r] – when making reports about itself – will, as proposed in Section 10.5.1, refer to itself (to its self-model) as "I".

Note that the preceding key paragraph describes a way to derive 'conscious kinds' from 'material kinds', thereby providing a materialist solution to the Explanatory Gap consistent with, for example, the relatively recent analysis, approach and requirements developed by Papineau.[234]

It also meets Nagel's seminal call to:[235]

> "…acknowledge that a physical theory of mind must account for the subjective character of experience…"

and his requirement that:

> "If mental processes are indeed physical processes, then there is something that it is like, intrinsically, to undergo certain physical processes."

where a proposed candidate set of physical processes[236] has been described centrally and at length through Working Note A Parts 1 and 2.

(And for the record, the approach here is also consistent with Nagel's famous concern that a person can never really know what it is like to be a bat.)[237]

Part 3 of Working Note A will now extend the work of Part 2 by addressing the nature of imagination, recollection, decision-making and emotion, to culminate in an examination of the question of free will in a deeper and more complete sense than could be attempted in the note *How is Free Will Possible?*

---

[234] Footnote 232 ref (2).

[235] Ibid. ref (1).

[236] Where here, 'physical processes' are constituted by proposed architecture and processes at the noumenal level of D[r] within X[r] within B[r] within W[r] where these are considered to be available to understanding in the form of an understanding of W[z], as is being achieved to an ever increasing degree through the practices of modern science (see Appendix 1 at https://teleodyne.com/working_note_A_appendix_1.pdf for relevant definitions, and for further discussion of the relationship between W[r] and W[z] see *The Construction of Phenomenal Time*, including at page 1 and footnote 5, at https://teleodyne.com/time.pdf.)

[237] Apropos, beginning with an elaboration regarding the key paragraph above: 'what it is like' will include for any for any given D[r], an experience – from the inferential perspective of 'what it is like' to be a D[a] (self) – of 'what it is like' to be that D[a] (self) experiencing, for example, seeing a rainbow, hearing a gunshot, driving a car, or experiencing from the perspective of a Self-as-Centre any other of what are, at the noumenal level of D[r], elements of that D[r]'s inference of itself as a D[a] which, as described, will be inferred *in terms of its evolving relations and interactions with* the optimally parsed and evolving substates of some given time series of Ξ. More specifically, what is being described here is experience from the perspective of a D[a] (self) of a rainbow[i], gunshot[i], car[i] etc. where all such perceived entities[i] will be unique to any given individual, and will have subjectively perceived qualities dependent upon the unique neurology and history of learning pertaining to the recognition and generative models operating within that individual's D[r].

From this perspective, the fact that *any* commonality – let alone the great commonality humans experience – can exist between what one individual and another individual perceive in experiencing some given shared environment[r] *must* inter alia be due to evolutionarily driven common hardwired inferential architecture[r] within respective human D[r]s – for example the 3-space inference – giving rise to a common overall structure for subjectively perceived worlds – as per the sphere world diagram – along with far more at the level of common overall neurological predispositions to certain forms of optimal parsing, with all this then further 'synchronised' – including at an evolutionary scale – through social experience, including common education, cooperative physical activity and, most importantly, through language use. It will be this commonality that allows one person to 'identify' with 'what it is like' to be another person in some given common situation.

Here Nagel's observations of human inability to identify with 'what it is like' to be a bat can, fully consistent with the case he makes, be seen as due to an expected massive relative difference between a putative bat D[r], and a bat's resultant experience of itself as a D[a], and the differences between one human D[r], and its experience of itself as a D[a], and this experience for another human D[r]. (An exploration of how individuals' respective W[i]'s can be 'synchronised', i.e. brought into 'register', is provided in the Essay at Section 8.2. The core ideas presented there still hold, though much of the architecture and processes referred to in that Section have now been superseded by those proposed in Working Note A – Part 1.)

In relation to Nagel's interest in bat sonar, and its implications for bat phenomenal experience, there certainly seem to be some interesting questions around what kind of parallels to the 3-space inference might be under implementation in the sonar using animals. For example, unlike in humans dolphin auditory and visual cortical areas are adjacent, an arrangement that likely facilitates fast heteromodal hierarchical processing integrating upward propagating cueing/error inputs from those areas perhaps leading, for a putative dolphin D[a], to perception of objects within some form of phenomenal space that may be somewhat, or perhaps highly, different to the phenomenal space that humans experience (see Herman, L.M. (2010) What Laboratory Research has Told Us about Dolphin Cognition *International Journal of Comparative Psychology 23* 310-330). What then might the sphere world diagram look like for a dolphin, an orca or a bat? More provocatively, is it conceivable that it might one day be possible – through comparative anatomy – to derive such diagrams by studying the neurological architecture and processes attendant to hierarchical processing as it operates within the living brains of these animals?