Second International Conference on
Geometry, Integrability and Quantization
June 7–15, 2000, Varna, Bulgaria
Ivaïlo M. Mladenov and Gregory L. Naber, Editors
Coral Press, Sofia 2001

A PRIMER ON OBSERVER THEORY

CHETAN PRAKASH

Department of Mathematics, California State University
5500 University Parkway, San Bernardino, CA 92407, U.S.A.

Abstract. This article is a survey which presents the essential ideas of "Observer Theory", a formal theory of perception, developed since the late 80's by Bruce Bennett and Donald Hoffman (both at U.C. Irvine) and myself. First I present the structure of an observer and one type of a framework, within which interactions between observers may be studied. Then I discuss the kinds of dynamics that can arise from such a framework, and how the dynamics can give rise to higher-level or "specialized" observers. Finally I indicate briefly what this says about "true" perception (i. e., perception adapted to the "world" the observer framework is in) and some possible ramifications which could lead to a deeper understanding of the origin of quantum systems and measurement theory. The general reference for this work is Bennett, Hoffman and Prakash [1].

1. Definition and Examples of Observer

We acknowledge certain principles as guiding any formal definition of an observer:

A. Perception is a process of inference.
   That is, an act of perception is a *process* of arriving at conclusions from a set of premises. Premises and conclusions are *propositions,* i. e., statements that can only be either true or false. No special or customary meaning of "inference" is meant beyond this. For example, consider the familiar "Necker Cube" in Fig. 1.
   Here the premises are a set of lines in the plane and, for most people, the conclusion is, at any given moment, a cube in space. So we can say that the conclusion is a 3-dimensional figure, while the premise is a 2-dimensional one — clearly an inference (and an illusory one at that!) and

not a deduction. But the conclusion is actually more than this: since we see *two*, not one, cubes sequentially and with a fairly constant proportion of time for each, we can say that the conclusion is actually a **probability measure** concentrated on these two cubes. So, in order to accommodate the multistability that perception allows, we need to consider conclusions that are probability measures.

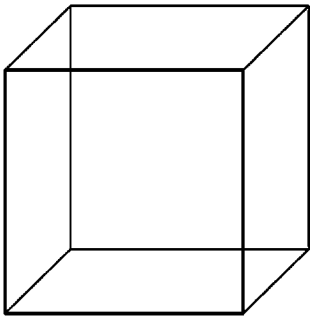

Figure 1. A Necker cube

B. Perceptual inferences are not, in general, deductively valid.
Another example of an illusionary perceptual inference from 2D to 3D is a so-called 3D movie. Yet another example is the "cosine surface" (see e. g. [3]), reproduced below.

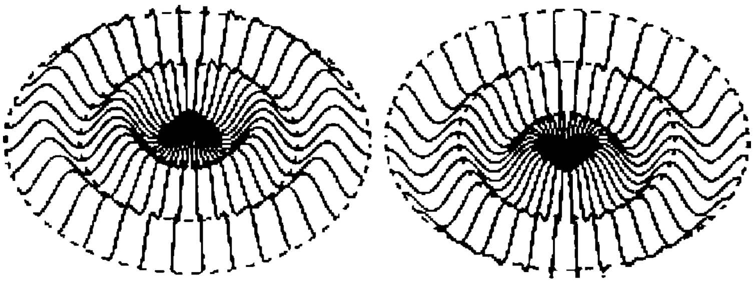

Figure 2. The cosine surface

Here the premise, which is a set of wavy lines in the plane, gives rise to a 3-dimensional figure of hills and valleys (an interesting side-note about this figure: as it is rotated slowly, there comes a point at which the hills and valleys interchange!).
So perceptual inferences may not be deductively valid. However, they may well be "inductively strong", in terms of the world within which the inference is being made. Another characteristic we note is that, at least for some visual inferences, more than one punctual conclusion is possible.

C. Perceptual inferences are biased.
Given a premise, we systematically tend to see one (or a few) of a possible infinity of inductively valid conclusions. (By "inductively valid" I mean a conclusion that, by the laws of geometry and physics, *could* have given rise to the premise).

As examples, consider the two above. There is an infinity of possible 3D figures that could have projected down to the Necker Cube, or to the Cosine surface. Yet we pick one, or a few (and in fact the wrong one, because the stimulus is in fact 2D). Stereo vision is yet another example. There is fascinating account of many other examples and their interpretation and implications for theories of vision in [4].

These principles, then, lead to the definition of an abstract entity called an *observer* and to the *Observer Thesis*:

Every perceptual modality has the structure (in respect to its perceptual characteristics) of an observer.

Just as with, say, the Turing thesis of computer science (that every act of computation may be instantiated as a Turing machine), this hypothesis is falsifiable.

Definition. *An **observer** is a six-tuple,* $\big((X,\mathcal{X}),(Y,\mathcal{Y}),E,S,\pi,\eta\big)$*, consisting of*

1. *a measurable space* $(X,\mathcal{X})$ *called the configuration space (*$\mathcal{X}$ *denotes the measurable structure on* X*, and similarly for other spaces below),*
2. *a measurable space* $(Y,\mathcal{Y})$ *called the premise space,*
3. *a measurable subspace* $(E,\mathcal{E})$ *of* $(X,\mathcal{X})$ *called the distinguished configuration space,*
4. *a measurable subspace* $(S,\mathcal{S})$ *of* $(Y,\mathcal{Y})$ *called the distinguished premise space,*
5. *a measurable surjective function* $\pi\colon X \to Y$ *with* $\pi(E) = S$*;* π *is called the perspective map,*
6. *a Markovian kernel* η *on* $S \times \mathcal{E}$ *such that, for each* s*,* $\eta(s,\cdot)$ *is a probability measure supported in* $\pi^{-1}\{s\} \cap E$*.*

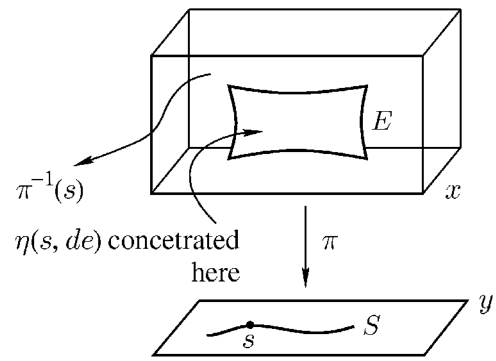

Figure 3. Schematic of an Observer

Note: *Markovian Kernels and Regular Conditional Probability Distributions.* Let $(X,\mathcal{X})$, $(Y,\mathcal{Y})$ be measurable spaces. A **kernel** on X relative to Y or a kernel on $Y \times \mathcal{X}$ is a mapping $N\colon Y \times \mathcal{X} \to \mathbb{R} \cup \{\infty\}$, such that

i) for every y in Y, the mapping $A \to N(y, A)$ is a measure on X, denoted by $N(y,\cdot)$;

ii) for every A in $\mathcal{X}$, the mapping $y \to N(y, A)$ is a measurable function on Y, denoted by $N(\cdot, A)$.

N is called **positive** if its range is in $[0, \infty]$ and **Markovian** if it is positive and, for all $y \in Y$, $N(y, X) = 1$. If $X = Y$ we simply say that N is a *kernel on* X. If N is a kernel on $Y \times \mathcal{X}$ and M is a kernel on $X \times \mathcal{W}$, then the *product* $NM(y, A) = \int_X N(y, \mathrm{d}x)M(x, A)$ is also a kernel. This algebra of kernels comes in handy when attempting to describe true perception later on. It is evident that Markovian kernels are a natural device for the description of probabilistic conclusions (about subsets of the configuration space) made as inferences from punctual stimuli (i. e., points of the premise space).

Remark: *The Interpretation Kernel as a Regular Conditional Probability Distribution.* Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ are measurable spaces. Let $p\colon X \to Y$ be a measurable function and μ a positive measure on $(X, \mathcal{X})$. A **regular conditional probability distribution** (abbreviated **rcpd**) of μ with respect to p is a kernel $m_p^\mu\colon Y \times \mathcal{X} \to [0, 1]$ satisfying the following conditions:

i) m_p^μ is Markovian;

ii) $m_p^\mu(y, \cdot)$ is supported on $p^{-1}\{y\}$ for $p_*\mu$-almost all $y \in Y$;

iii) If $g \in L^1(X, \mu)$, then $\int_X g \, \mathrm{d}\mu = \int_Y (p_*\mu)(\mathrm{d}y) \int_p^{-1}\{y\} m_p^\mu(y, \mathrm{d}x) g(x)$.

It is a theorem that if $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ are standard Borel spaces then an rcpd m_p^μ exists for any probability measure μ [6]. In general there will be many choices for m_p^μ, any two of which will agree a.e. $p_*\mu$ on Y (that is, for almost all values of the first argument). If $p\colon X \to Y$ is a continuous map of topological spaces which are also given their corresponding standard Borel structures one can show that there is a canonical choice of m_p^μ defined everywhere.

Conversely, we can build an appropriate μ as follows: suppose we are given a p.m. λ on S and an interpretation kernel η on $S \times \mathcal{E}$ which is *concentrated on the fibers* $\pi^{-1}(s)$ *of* π. Then the interpretation kernel is a canonical choice of the regular conditional probability distribution of the measure $\lambda\eta(\mathrm{d}e) = \int_S \lambda(\mathrm{d}s)\eta(s, \mathrm{d}e)$ on E, with respect to the map π. Also, λ is the *distribution* $\pi_*(\lambda\eta)$ of $\lambda\eta$ under the map π.

Suppose the state of affairs in the world is such that the configurations are subject to a probabilistic law μ, i. e., the probability of a stimulus arising from a measurable set A of E is $\mu(A)$. Then we could identify a "truly" perceiving interpretation kernel η as the rcpd of μ with respect to π. A careful study of how the measure μ comes about does indeed allow us to do that, will be discussed further on.

1.1. Examples

The Necker Cube. Here X is the set of 3D line figures with 8 vertices, say, while E is the set of 3D cubes. Correspondingly, Y is the set of 2D figures, while S consists of those 2D figures that project orthographically (as a decent approximation) from the figures in E. So π is orthographic projection. Each fiber of π consists of two points of E and innumerable points of $X - E$. Then, η is the kernel that is concentrated, for each $s \in S$, on those two points of $\pi^{-1}(s) \cap E$.

For a more sophisticated example we have:

Rigid Fixed Axis Motion [5]. This is an example of an observer that identifies a 3D body, rigidly moving in space around a fixed axis, by interpreting 3 views of it in 2D. We assume that the body has 4 (or more) "feature points," identifiable across the 3 views. For the purpose of reconstructing the rotational motion (as against absolute position in space) we may assume that one of the points is always at the origin, so that the body, or its image, can be described in terms of the 3 displacement vectors from that point to the others. Let us denote these displacements by $a_{i,j}$ ($i = 1$, 2, or 3 for the point and $j = 1$, 2 or 3 for the view). So here X is the set of ordered triples of points in $\mathbb{R}^3$, or $\mathbb{R}^{18}$, while Y is the set of ordered triples of points in $\mathbb{R}^2$, or $\mathbb{R}^{12}$. π is, as before, parallel projection, $\mathcal{X}$ and $\mathcal{Y}$ are the Borel algebras and E is the algebraic variety in X defined by six rigidity equations

$$
\begin{aligned}
\vec{a}_{11} \cdot \vec{a}_{11} - \vec{a}_{12} \cdot \vec{a}_{12} &= 0\,, \\
\vec{a}_{11} \cdot \vec{a}_{11} - \vec{a}_{13} \cdot \vec{a}_{13} &= 0\,, \\
\vec{a}_{21} \cdot \vec{a}_{21} - \vec{a}_{22} \cdot \vec{a}_{22} &= 0\,, \\
\vec{a}_{21} \cdot \vec{a}_{21} - \vec{a}_{23} \cdot \vec{a}_{23} &= 0\,, \\
\vec{a}_{11} \cdot \vec{a}_{21} - \vec{a}_{12} \cdot \vec{a}_{22} &= 0\,, \\
\vec{a}_{11} \cdot \vec{a}_{21} - \vec{a}_{13} \cdot \vec{a}_{23} &= 0\,,
\end{aligned}
$$

and two fixed-axis equations

$$
\begin{aligned}
(\vec{a}_{11} - \vec{a}_{12}) \cdot [(\vec{a}_{11} - \vec{a}_{13}) \times (\vec{a}_{21} - \vec{a}_{22})] &= 0\,, \\
(\vec{a}_{11} - \vec{a}_{12}) \cdot [(\vec{a}_{11} - \vec{a}_{13}) \times (\vec{a}_{21} - \vec{a}_{23})] &= 0\,.
\end{aligned}
$$

Finally, we define $S = \pi(E)$. One can deduce (with some effort) that the fibers of π over S have, generically, two points in E and an interpretation kernel will then "output" a measure on these two points.

2. Reflexive Frameworks of Observers

What does an observer observe? A parsimonious theory of perception posits that the essential structure of observer and observed is, at some sufficiently subtle level of analysis, the same. Specifically, we have the **Non-Duality Hypothesis**:

To each perceptual interaction there is a level of analysis at which the object of an observer's perception is another observer, with the same representational structure (X, Y, E, S) but with (possibly) different π and η.

At this level, then, the collection of possible observers which can observe each other make up a framework, in which each observer is indexed by its own particular π and η:

Definition. *A **reflexive observer framework** on (X, Y, E, S) is an injective map*

$$\Pi\colon E \to \mathrm{Hom}(X, Y)$$

such that, for each $e \in E$, we have $\Pi(e)(E) = S$.

Notationally, we write π_e for $\Pi(e)$; let $\mathcal{B}_e = (X, Y, E, S, \pi_e)$ denote the set of *"preobservers"* (i. e., observers without their interpretation kernels specified) with the same perspective map.

Example 1. *Take $X = \mathbb{R}^2$ and $E = \mathbb{Z}^2$, so that the observers are indexed by integer points in the plane. Letting $Y = \mathbb{S}^1$, we note that the straight line between any pair of observers, at e and e', say, intersects the circle of unit radius at e at some "rational" point of this circle: we take S to be the set of such rational points on the unit circle $\mathbb{S}^1$.*

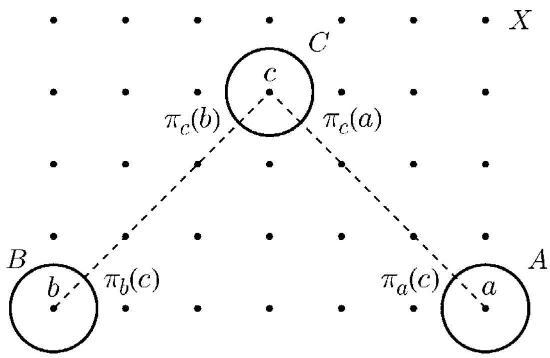

Figure 4. A symmetric observer framework

Example 2. *Let $X = G$, a measurable group with E a measurable subgroup of G. Let H be another measurable subgroup of G and let $Y = G/H$ by the left cosets of H. If we take π to be the canonical map $\pi\colon G \to G/H$, this* *implies that $S = EH/H$: those left-cosets of H in G that intersect E. We see that then $\pi_e(g) = \pi(ge^{-1})$ for any $e \in E$ and $g \in G$.*

$$\begin{array}{ccccc} & E & \subset & G & = X \\ & {\scriptstyle \pi_e|_E}\downarrow & & \downarrow{\scriptstyle \pi_e} & \\ S = & HE/H & \subset & G/H & = Y \end{array}$$

Note that Example 1 is not of this form. There the set of fibers of π_e depend on e, whereas here that set is independent of e.

Example 3. *Again, let $X = G$ be a measurable group with E a measurable subgroup of G. But now allow H to be any measurable group acting on G, on the right. Put $Y = G/H$, the orbits of the action of H on G. If we take π to be the canonical map $\pi\colon G \to G/H$, this implies that $S = EH/H$: those orbits that intersect E. Once again, $\pi_e(g) = \pi(ge^{-1})$ for any $e \in E$ and $g \in G$.*

$$\begin{array}{ccccc} & E & \subset & X & = G \\ & {\scriptstyle \pi|_E}\downarrow & & \downarrow{\scriptstyle \pi} & \\ EH/H = & S & \subset & Y & = G/H \end{array}$$

Note that if H is a subgroup of G, this is just Example 2. However, if H is $\mathbb{R}_+$ acting as dilations on $\mathbb{R}^2$, this is Example 1. Our next example generalizes yet further.

Example 4. *Let G be a measurable group with $\mathcal{H}$ an equivalence relation on G, so that $Y = G/\mathcal{H}$ is the partition of G by $\mathcal{H}$. Take π to be the canonical map $\pi\colon G \to G/\mathcal{H}$. Now suppose J is a measurable subgroup of G and that $\pi(J)$ is also measurable; put $S = \pi(\mathcal{J})$.*

Let X be a G-space: G acts measurably on X on the left. Fix $x_0 \in X$ and set $E = Jx_0$. We need to assume that (i) G acts transitively on X and (ii) Σ_e, the stabilizer of e is, for each e, in a single $\mathcal{H}$-class.

$$\begin{array}{ccccc} E = & Jx_0 & \subset & X & \\ & {\scriptstyle \pi_e|_E}\downarrow & & \downarrow{\scriptstyle \pi_e} & \\ S = & \pi(J) & \subset & G/\mathcal{H} & = Y \end{array}$$

If xe^{-1} denotes *any* element of G such that $ge = x$, then we can now unambiguously define π_e by $\pi_e(x) = \pi(xe^{-1})$.

Note that Example 4 generalizes the previous examples. For example, $\mathcal{H}$ need not arise from a group action; the action of G on X need not be faithful.

This example is significant in that it permits us the full capacity to introduce relativity, by allowing symmetries into the reflexive observer framework. First we make a definition.

Definition. *A symmetric observer framework is a reflexive observer framework $(X, Y, E, S, \pi_\bullet)$ for which there exists a measurable group G, a measurable subgroup $J \subset G$, and a measurable surjective map $\pi\colon G \to Y$ satisfying two requirements:*

i) *G acts transitively and measurably on X, inducing a transitive action of J on E (which is automatically measurable).*
ii) *For all $e \in E$ and $x \in X$, $\pi_e(x) = \pi(g)$, where g is any element in G such that $ge = x$ (i. e., $g =$ "xe^{-1}").*

Then it is easily checked that this is the same as Example 4:

Theorem 1. *The definition of symmetric observer framework is equivalent to Example 4 above.*

Proof: The fibers of the map π form a partition of G. The relation of joint membership in a fiber is an equivalence relation: call it $\mathcal{H}$. Then Y is identified with $G/\mathcal{H}$. Since the action of J on E is transitive, E is identified with Je_0 for any $e_0 \in E$. It remains to verify only (ii) but this is clear. □

Now, in order to create a participator dynamics, we need interpretation kernels for each $e \in E$. But we need to do this with a view to the symmetry.

Definition. *Let a symmetric observer framework $\mathcal{S} = (X, Y, E, S, G, J, \pi)$ be given. A family $\{\eta_e\}_{e\in E}$ of interpretation kernels for $\mathcal{S}$ is said to be* ***symmetric*** *if there exists a Markovian kernel $\eta\colon S \times \mathcal{J} \to [0, 1]$ such that for all $e \in E$, $s \in S$ and $\gamma \in \mathcal{E}$,*

$$\eta(s, \pi^{-1}\{s\} \cap J) = 1,$$

and

$$\eta_e(s, \Gamma) = \eta(s, \Gamma e^{-1})\,.$$

η is then called the ***fundamental kernel*** *of the family η_e.*

One way a family can be symmetric is as follows. Suppose we are given a symmetric observer framework (X, Y, E, S, G, J, π) and a measure ν on J. Since J acts transitively on E, given any $e \in E$ we get a surjective map $c_e\colon J \to E$ by sending the identity element $\imath$ of J to e. c_e identifies E with the quotient space $J/\Sigma_e \cap J$, where Σ_e is the stabilizer of e in G. Let $\nu_e = (c_e)_*(\nu)$; this is the measure ν transported to E by "centering a copy of J at e".

Terminology: With the hypotheses and notation of the previous paragraph, if ν is a measure on J, the family of measures ν_e on E is called the *symmetric family of measures associated to ν*; ν is called the *fundamental measure of the family*. That is, if $\Gamma \in \mathcal{E}$, then $\nu_e(\Gamma) = \nu(c_e^{-1}(\Gamma)) = \nu\{j \in J;\, je \in \Gamma\}$.

Example 5. *Very briefly, we present the Instantaneous Rotation Observer: details may be found in [2]. If we consider an observer that detects rigid motion from two views of 4 points, then it turns out that X and E are principal homogeneous spaces for G and J respectively, where*

$$\begin{aligned} G &= SO(3,\mathbb{R}) \times (\mathbb{S}^1)^{n-1} \times (\mathbb{S}^1)^n \times \mathbb{R}^n \times \mathbb{R}^n \times (\mathbb{R}^*)^n \times (\mathbb{R}^*)^n \, . \\ J &= SO(3,\mathbb{R}) \times (\mathbb{S}^1)^{n-1} \times \mathbb{S}^1 \times \mathbb{R}^n \times (\mathbb{R}^*)^n \, . \end{aligned} \tag{1}$$

This concludes the presentation of the static aspects of observer theory; with this background, we are in a position to examine a dynamical system of "observers" based on a symmetric framework.

3. Participators

Definition. *In a symmetric observer framework $\mathcal{S}$, an action kernel for $\mathcal{S}$ is a Markovian kernel*

$$Q\colon E \times \mathcal{E} \mapsto [0,1]$$

such that, if $\pi(e) = \pi(e')$, $Q(e,\cdot) = Q(e',\cdot)$. The local action kernels are then

$$Q_{\bar{e}}(e,\Gamma) = Q(e\bar{e}^{-1},\Gamma)\,.$$

The action kernels will determine the changes in perspective of participators upon interaction with each other.

Definition. *A* ***participator*** *on the symmetric observer framework $\mathcal{S}$ is a triple $(\xi, \{Q(n)\}, \{\eta(n)\})$, $(n = 0, 1, 2, \dots)$ where*

- *ξ is a probability measure on E (the* ***initial measure*** *of the dynamics);*
- *$Q(n)$ is an action kernel for $\mathcal{S}$ and*
- *$\eta(n)$ is a fundamental interpretation kernel for $\mathcal{S}$.*

The group J (that acts transitively and measurably on E) is called the ***state space*** *of the participators on the framework $\mathcal{S}$.*

If $Q(n)$ is stationary, we refer to the participator as **kinematical**.

Intuitions. ξ is a starting measure for the participator. Each trajectory of the dynamics will result, at "reference time" n, in the participator **manifesting** as an observer

$$A_n = (X, Y, E, S, \pi_{e_n}, \eta(n))\,.$$

If there are k participators, we can think of the state space for the stochastic dynamics as E^k; ξ will then be a k-fold initial measure.

Now participators interact via a **channelling** between their observer manifestations. Successive instants of reference time are marked by successive channellings (between one or more pairs of participators). For any given participator, it's own **proper time**, i. e., the times when it participates in channellings, is a random variable that is a **stopping time** in the canonical chain to be introduced later. Clearly, a channelling is an involution on the set of participators.

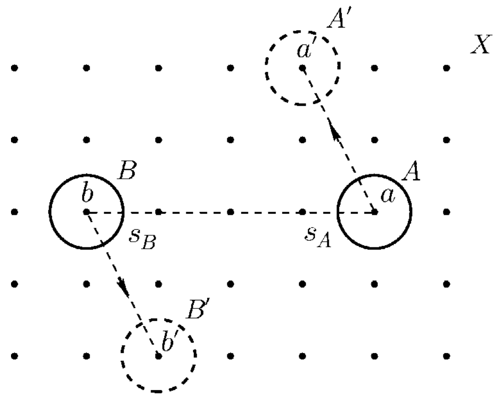

Figure 5. Channelling between participators and their movement to new perspectives

Dynamical Assumptions

1. *Closed System.* We assume that the number of participators is conserved (though any individual may not channel at any given instant of reference time).
2. *Independent Action.* Given the current set of perspectives and the channelling involution, the *next* set of perspectives to manifest is an independent set of random variables.
3. *Channelling Distribution.* The channelling at any given instant is stochastic and is governed by a Markovian kernel: we will discuss this in detail later.

As an example of the Independent Action postulate above, suppose A, B and C are three participators, two of which channel. Where will they be after the channelling? The channellings and the action kernels determine this as follows:

$$P\left(A \in \Gamma_A, B \in \Gamma_B, C \in \Gamma_C \,\middle|\, \begin{array}{l} A \text{ at } e_A, B \text{ at } e_B, C \text{ at } e_C, \\ (A,B) \text{ channel}, C \text{ does not} \end{array}\right)$$
$$= Q_{e_A}(e_B, \Gamma_A) Q_{e_B}(e_A, \Gamma_B) 1_{\Gamma_C}(e_C) .$$

A simple example of a Markovian participator dynamics may be presented in the instance where we have a pair of kinematical participators and an Abelian group J. Then it is a straightforward exercise to see that we have the following theorem.

Theorem 2. *With exactly two kinematical participators channelling with Abelian state space J, with action kernels Q and R, the dynamics is that* *of a Markov chain, with state space J and transition probability*

$$P(e,\Gamma) = \int_{J\times J} 1_\Gamma(e-k+h)Q(e,\,\mathrm{d}h)R(-e,\,\mathrm{d}k)\,.$$

Example 6. (Asymptotics of a two-participator integer dynamics) *Consider $E = \mathbb{Z}$, the integers, $Y = S = \{1,0,-1\}$, $X = G = \langle \mathbb{R},+\rangle$ and $J = \langle \mathbb{Z},+\rangle$. Take the fundamental map π to be $\pi(x) = \mathrm{sign}(x)$. Consider a pair of participators "moving on" (i. e., whose perspective maps are indexed by) the integers. We will assume that there is no self-channelling, and that the fundamental action kernel is given by the delta function $Q(0,\cdot) = \delta_0(\cdot)$ and that*

$$Q(r,x) = \begin{cases} \rho\,, & \textit{if } x = \mathrm{sign}(r); \\ 1-\rho\,, & \textit{if } x = \mathrm{sign}(-r); \\ 0\,, & \textit{otherwise}\,. \end{cases}$$

Thus each participator jumps, upon a channelling, towards the other participator with probability ρ and away from it with probability $1-\rho$.

$1-\rho \leftarrow \bullet \rightarrow \rho$ $\qquad\qquad$ $\rho \leftarrow \bullet \rightarrow 1-\rho$

$* * *|* * *|* * * \bigcirc * * *|* * *|* * *|* * *|* * * \bigcirc * * *|* * *|* * *$

Participator 2 $\qquad\qquad$ Participator 1

Figure 6. A Markovian two-participator dynamics with $E = \mathbb{Z}$

Then it is an exercise to see that for $0 \le \rho \le \frac{1}{2}$ there is a unique stationary measure $\nu(q) = \delta_0(q)$, while for $\frac{1}{2} < \rho < 1$ there is a **one-parameter family** of stationary measures

$$\nu(q) = \begin{cases} d\left(\dfrac{1-\rho}{\rho}\right)^{|q-1|}, & \text{if } q \text{ is odd and } q > 0\,; \\ d\left(\dfrac{1-\rho}{\rho}\right)^{|q+1|}, & \text{if } q \text{ is odd and } q < 0\,; \\ 0, & \text{if } q \text{ is even and } q \neq 0\,; \\ \dfrac{2\rho - 1 - 2\rho^2 d}{2\rho - 1}, & \text{if } q = 0\,. \end{cases}$$

The range of allowed values of the parameter d is contained in the closed interval $[0,1]$. For fixed $\rho \in (\frac{1}{2},1]$ the range is $[0,(2\rho-1)/2\rho^2]$. For a proof of this result (see [1]).

What is the significance of these stationary measures, and of the asymptotics of these participator Markov chains in general? Firstly, we are in a position to clearly define what is meant by *true perception*. Recall that many Markov chains (in particular the Harris chains [7]) possess asymptotic **absorbing sets** with their own stationary measures: once the chain enters an absorbing set, it stays there and is governed by the law of the stationary measure. Suppose, then, that there is a stationary measure ν of the (full) dynamics of participators (upon the nature of which we will expand a bit in the next section). Consider a particular participator, which experiences a "subjective" chain, called it's **trace** chain, and so experiences a subjective stationary measure which we will call $\mathcal{D}\nu$. Then we may propose the

Definition. *If a participator Markov dynamics is asymptotically in a given absorbing set with stationary measure ν, then a given participator is said to* ***perceive truly*** *if its fundamental interpretation kernel is (a version of) the regular conditional probability distribution of the trace measure $\mathcal{D}\nu$ with respect to its fundamental map π.*

Here the trace is essentially projection onto the participator's subjective space E. Further details on the participator dynamics will be provided after we have discussed the channelling distribution.

4. Channelling and the Markovian Dynamics

Notation. If D denotes a subset of $\{1, \ldots, k\}$ and χ denotes an involution on D, we let $\mathcal{I}_k$ denote the set of all such involutions (D, χ) on subsets of $\{1, \ldots, k\}$.

Definition. *A τ-distribution is a family $\tau = \{\tau_k\}_{k=1}^{\infty}$ where each τ_k is a Markovian kernel on $E^k \times 2^{\mathcal{I}_k}$ satisfying*

i) Consistency: *Given $k' < k$, let $S' = \{1, \ldots, k'\}$, $S = \{1, \ldots, k\}$. Then, (with the notation above) for any*

$$(y_1, \ldots, y_{k'}, z_{k'+1}, \ldots, z_k) \in E^k,\ \chi \in \mathcal{I}(k'),$$

we have

$$\tau_{k'}(y_1, \ldots, y_{k'}; \chi) = \frac{\displaystyle\sum_{\chi'' \in \mathcal{I}(S-S')} \tau_k(y_1, \ldots, y_{k'}, z_{k'+1}, \ldots, z_k; \chi \cup \chi'')}{\displaystyle\sum_{\substack{\chi' \in \mathcal{I}(S') \\ \chi'' \in \mathcal{I}(S-S')}} \tau_k(y_1, \ldots, y_{k'}, z_{k'+1}, \ldots, z_k; \chi' \cup \chi'')}.$$

ii) Symmetry: *Recall that we are in a symmetric framework. If $x_i x_l^{-1} = y_i y_l^{-1}$ then $\tau(x; \cdot) = \tau(y; \cdot)$.*

This is also called **configurational symmetry**. Note that permutation symmetry is not required. Another possible symmetry condition, **translational symmetry** is discussed in [1]. Intuitively, the consistency condition allows us to relate the τ_k for different k's by asserting the independence of the channelling distribution within a subset, given that there is no channelling between that subset and its complement.

We are now in a position to discuss the various dynamics associated with a participator system.

The **Augmented Dynamics** of an ensemble of k participators on the symmetric framework $\mathcal{S}$ is the Markov chain $\{\xi_\tau, Q_i(n)\}_{i=1}^k$ with state space $E^k \times \mathcal{I}_k$ and with initial measure and transition probability defined as follows. Let $\xi_1, \ldots, \xi_k$ be initial probability measures defined on E^k. Let

$$\xi_\tau(\Delta_1 \times \cdots \times \Delta_k \times \{\chi\}) = \int\limits_{\Delta_1 \times \cdots \times \Delta_k} \xi(\mathrm{d}y_1 \cdots \mathrm{d}y_k)\tau(y_1, \ldots, y_k; \chi)\,.$$

We assume that the participators are distributed independently initially; to this end let the **initial measure** of the chain be

$$\xi = (\xi_1 \otimes \cdots \otimes \xi_k)_\tau$$

and let the **transition probability** be given by

$$\hat{N}_t(e, \chi_0; \Delta \times \{\chi_1\}) = \int\limits_{\Delta} N_{t,\chi_0}(e_1, \ldots, e_k;\, \mathrm{d}y_1 \cdots \mathrm{d}y_k)\tau(y_1, \ldots, y_k; \chi_1)\,,$$

where

$$\begin{aligned}&\hat{N}_t(e, \chi_0; \Delta \times \{\chi_1\}) \\ &= \int\limits_{\Delta} \Big[\prod_{i \in D(\chi_0)} Q_{ie_i}(t)\,(e_{\chi_0(i)};\, \mathrm{d}y_i) \prod_{j \notin D(\chi_0)} \epsilon_{e_j}(\mathrm{d}y_j) \Big] \tau(y_1, \ldots, y_k; \chi_1)\,.\end{aligned}$$

Various chains descend from this one: our objective is to define the **phenomenal** or **subjective reality** chain for a given participator. Firstly, we take the marginal distributions which sum the augmented dynamics over the channelling distributions: this may be termed the **noumenal** chain or **source** chain. If we then *relativise* (i. e., ignore absolute perspectives) after which we take a *trace* (i. e., ignore instants when the given participator is not involved in a channelling), we finally arrive at the participator's **phenomenal** chain. This is the chain we referred to above in our definition of true perception.

From a mathematical point of view, it is interesting that these various forgetful processes, performed in *any* order starting with the augmented dynamics, *preserve the Markovian nature of the dynamics* . These reductions and some general results are discussed in Chapter 7 of [1].

5. Specialization, Perceptual Hierarchy, Possible Connections with Quantum Measurement

Finally, I want to very briefly mention some ideas about the emergence of levels of perception and implications for both perception as well as physical science.

We may ask: How can an ensemble of participators in a symmetric framework give rise to a **higher level observer**, one which observes qualitatively different phenomena from those detected by its constituents? In order to develop a non-reductive theory of such a **specialization process**, i. e. a theory without "hidden variables" at the new level, we posit some principles:

1. The *premises* of the new observer should be deducible from the *sequence of conclusions* in the dynamics of the participators in the ensemble.
2. The sequence of lower-level conclusions should be reliable.
3. A channelling between two such specialized observers occurs as a result of the interactions between their instantiating ensembles, and in such a way as to preserve the identity of those instantiations.
4. The premise of a specialized observer is a stable perturbation of the asymptotic behaviour of its instantiated Markov chains.

It is this last which ensures the non-reductive nature of specialization. We can thus imagine a hierarchy of specializations, each level being the instantiation, or "physical," level for the next, specialized or "perceptual" level. Work is going on to tease out the mechanism of interaction between loosely interacting ensembles of instantiations. The idea is that if two such instantiations have arrived at absorbing sets of their respective dynamics and then interact weekly (i. e, with relatively few cross-channellings), they will tend to shift into new absorbing sets. One could say that each is then the instantiation of a specialized observer whose premise is the shift from one absorbing set to another: This represents the absorbing set of the other observed instatiation, or the *configuration* of the other specialized observer. The asymptotics is being used to observe the asymptotics, at the higher level.

Lastly, here is a brief note on quasicompact chains [7] and how their asymptotic behaviour is suggestive of quantum phenomena. Recall that for a Markov chain, the **asymptotic algebra** $\mathcal{A}$ is defined as

$$\{A \in \mathcal{F};\, \forall n\ \exists A_n \in \mathcal{F}_n : \theta^{-n}(A_n) = A\}\,,$$

where θ is the shift operator and $\{\mathcal{F}_n\}$ is the filtration of σ-algebras of the chain. The **invariant σ-algebra** $\mathcal{I}$ is

$$\{A \in \mathcal{F};\, \theta^{-1}(A) = A\}\,.$$

A transition probability P is **quasicompact** if there is a compact operator K such that, for some n_0-th iteration of P, $\|P_{n_0} - K\| < 1$. It is then a fact that if P is quasicompact, the asymptotic algebra $\mathcal{A}$ is finite a.s. and partitions the base space (i. e., the set of trajectories) $\mathcal{B}$ of the Markov chain. In particular, there exist natural numbers d_ρ and **cyclic classes**

$$E_{\rho,\sigma} = \{x \in \mathcal{B};\, P(x, E_{\rho,\sigma+1}) = 1\}\,,$$

for $1 \le \sigma \le d_\rho$ (modulo d_ρ). The absorbing sets are then the various $E_\rho = \cup_\sigma E_{\rho,\sigma}$. There are invariant probability measures m_ρ supported on the E_ρ. Where it gets interesting is that the eigenvalues of the transition probability P are the so-called **cyclic functions** $f_{\rho,k} = \mathrm{e}^{2\pi \mathrm{i} k/d_\rho}$, $1 \le k \le d_\rho$. Finally P behaves as a unitary time evolution on $\mathcal{L}^2(E_\rho, m_\rho)$. This is clearly redolent of the quantum mechanical situation with finite pure-point spectrum; it is entirely conceivable that appropriately weakening the hypothesis of quasicompactness will result in behaviour similar to continuous spectra. Coupling these ideas with those of specialization discussed above should lead to some interesting connections between perception and quantum mechanics at a much deeper level than has been hitherto realized.

References

[1] Bennett B., Hoffman D. and Prakash C., *Observer Theory: A Formal Theory of Perception*, Academic Press, 1989.

[2] Bennett B., Hoffman D. and Prakash C., *Structure from Two Orthographic Views of Rigid Motion*, J. Opt. Soc. Am., **6** (1989) 1052–1069.

[3] Hoffman D., *The Interpretation of Visual Illusions*, Sci. Am. 249, 6, (1983) 154–162.

[4] Hoffman D., *Visual Intelligence: How We Create What We See*, W.W. Norton & Company, 1999.

[5] Hoffman D. and Bennett B., *The Computation of Structure from Fixed Axis Motion: Rigid Structures*, Biol. Cyb. **54** (1986) 71–83.

[6] Parthasarathy K. R., *Introduction to Probability and Measure*, MacMillan, New Delhi 1968.

[7] Revuz D., *Markov Chains*, North Holland, Amsterdam 1984.