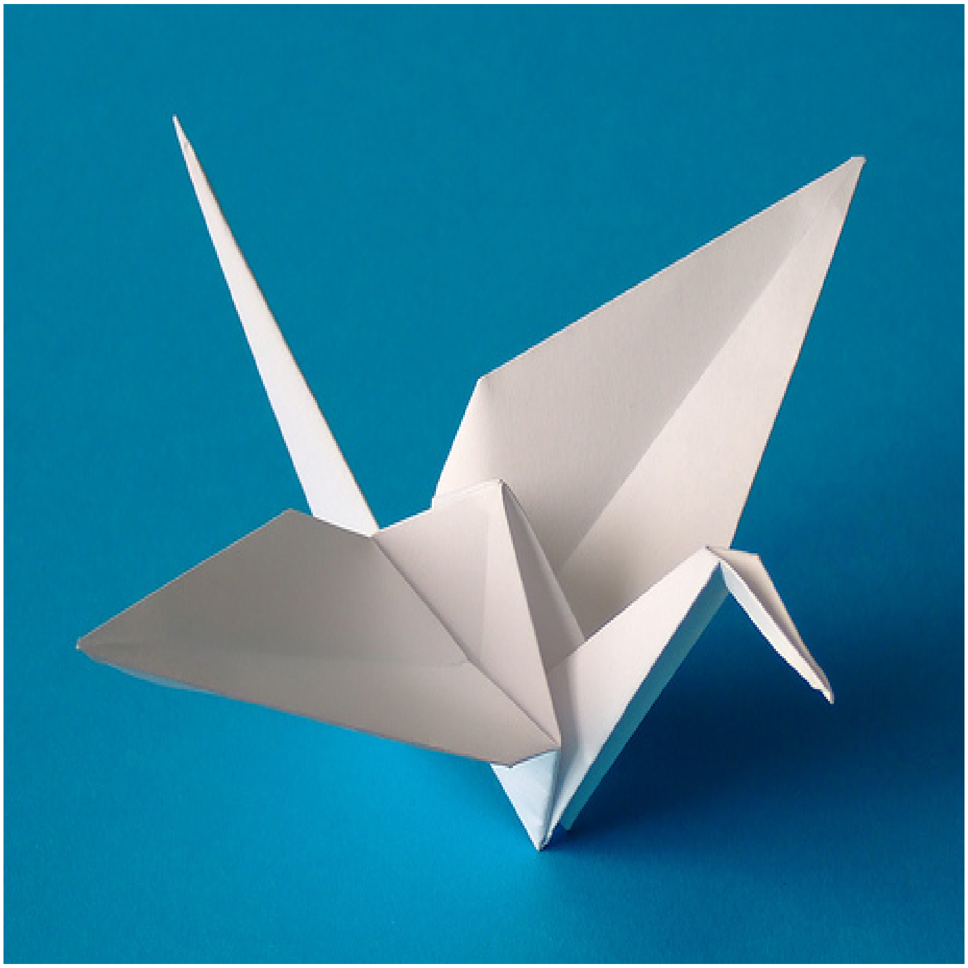
The Living and Intelligent Universe: The Constructivist Understanding of Complex and Non-Trivial Systems.
or
Known Objects as Eigen-Solutions, Kaleidoscopes and Mirror Houses.

Julio Michael Stern
jmstern@hotmail.com
University of São Paulo,
Brazil

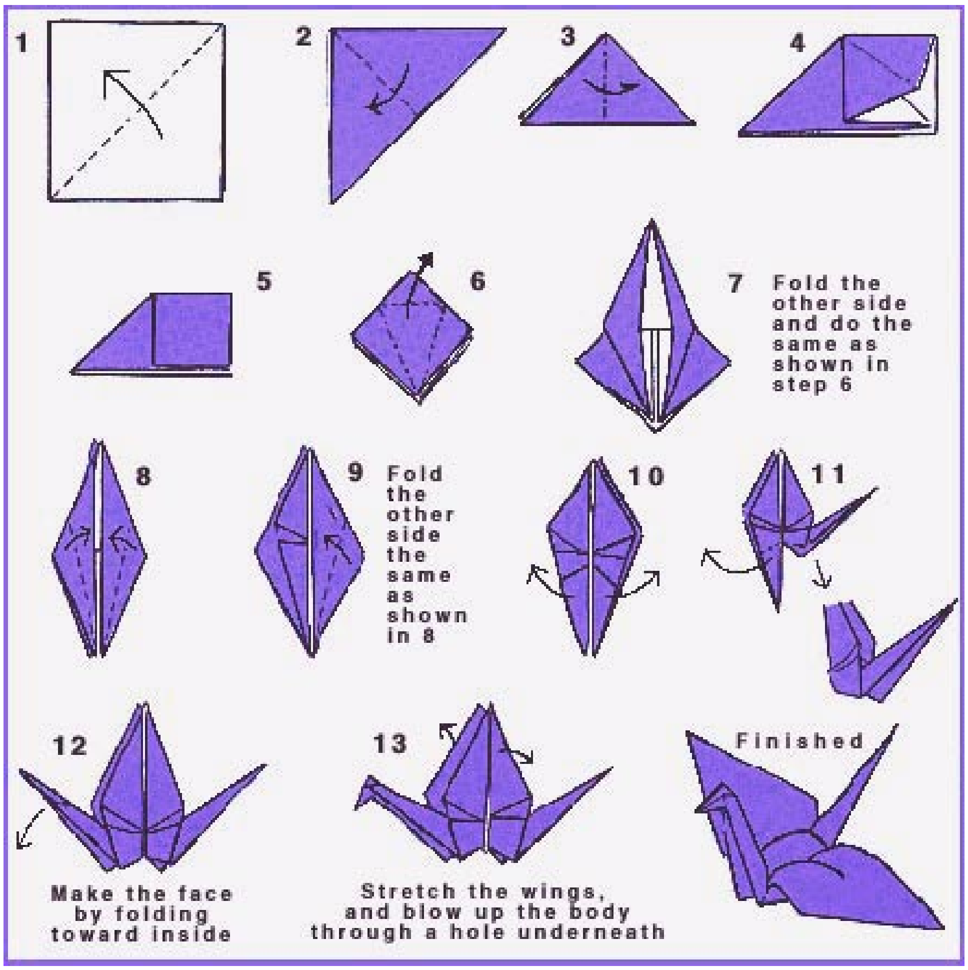
MBR'09 - Model-Based Reasoning
in Science and Technology
UNICAMP, Dec 17-19, 2009.

www.ime.usp.br/∼jstern/slide/Tsuru.pdf



Mneme (meme), from $\mu\nu\eta\mu\eta$ - memory.
Mime, from $\mu\iota\mu\eta\sigma\iota\varsigma$ - imitation.
A trace or unit of retrievable memory, a basic model, a single concept, an elementary idea.
Examples: I-Origami, II-Music, III-Science.

Example-I: "Tsuru Origami"



Tsuru - crane; Ori - fold; Kami - paper.
Instructions: Composition of Foldings.

Richards Dawkins presents the notion of reliable replication mechanisms in the context of evolutionary systems. The example contrasts two versions of the "Chinese Whispers" game using distinct copy mechanisms.

1- Suppose we assemble a line of 20 children. A picture, say, a Crane, is shown to the first child, who is asked to draw it. The drawing, but not the original picture, is then shown to the second child, who is asked to make her own drawing of it, and so on...

A trend will be visible as we walk from one end of the series of drawings to the other, and the direction of the trend will be degeneration...

2- Suppose we set up our Chinese Whispers game again, but this time with a crucial difference. We teach the first child how to make a Tsuru origami, and then ask him to teach the second child, and so on...

In some experiments the line of phenotypes will suffer an abrupt macromutation which will presumably then be copied to the end of the line. But in a good number of experiments, if we lay the 20 origami Tsurus out in order, some will be more perfect than others, but imperfections will not be copied on down the line...

Dawkins said: Origami instructions are high fidelity, even if not digital; they are self-normalizing. The code is error-correcting.

I say: Origami instructions (their language) are based on eigen-solutions, i.e. symmetries, invariances, equilibria, fixed points, etc.

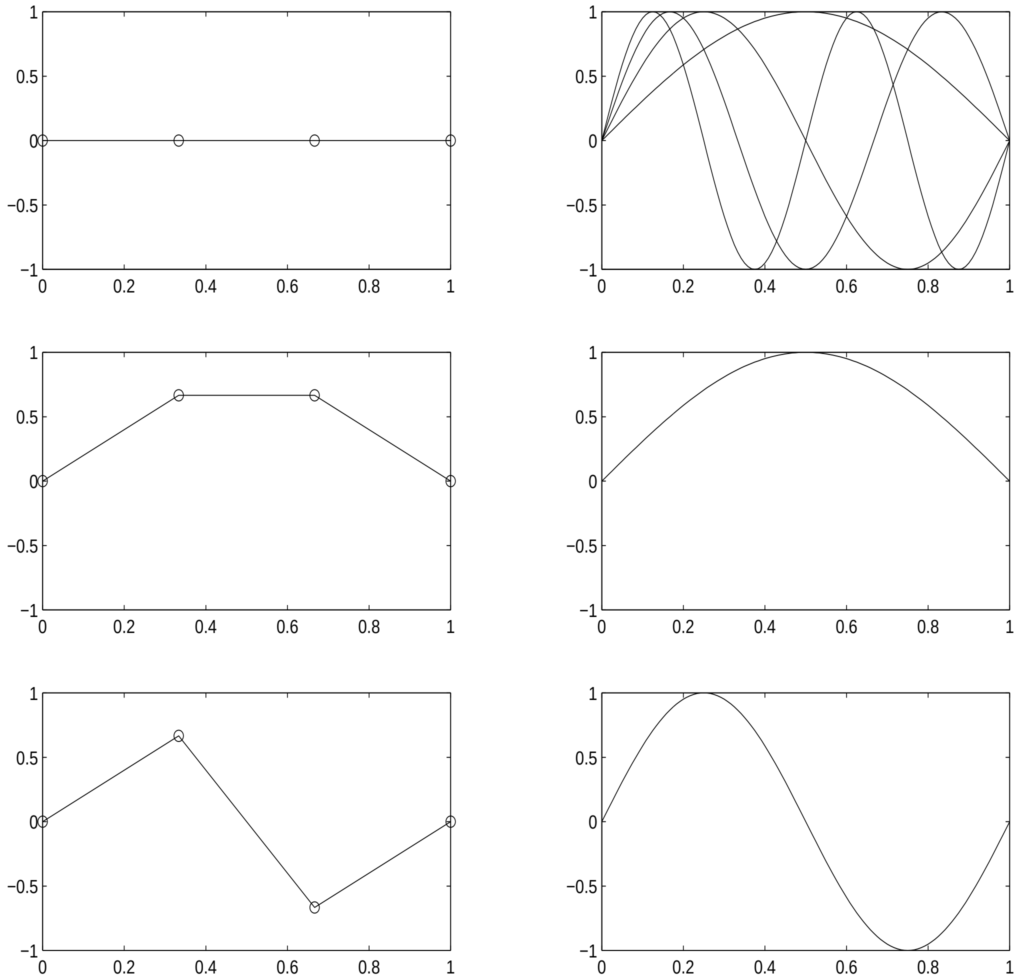
Heinz von Foerster characterized eigen-solutions by four essential attributes: Being discrete (precise, exact), stable, separable and composable.

Precision: The instruction "fold the paper along the diagonal of the square" implies folding at a specific line, a 1-dimensional object in the 2-dimensional sheet of paper. In this sense the instructions are sharp, precise or exact.

Stability: By adjusting and correcting the position of the paper (before making a crease) it is easy to come very close to what the instruction specifies. Even if the resulting fold is not absolutely perfect (in practice it actually never is), it will still work as intended.

Composability and Separability: We can compose or superpose multiple creases in the same sheet of paper. Moreover, adding a new crease will not change or destroy the existing ones. Hence, we can fold them one at a time, that is, separately.

Example–II: Music.



Eigen-Solutions, *Objects* aka Stationary Waves, Normal Modes, or Fourier basis for Continuous and Discrete Vibrating Chords.
- Composition rule: Linear Superposition.

Discrete chord system's dynamics, given by the second order differential equation:

$$\ddot{x} + Kx = 0 \ , \quad w_0^2 = \frac{h}{ms} \ .$$

$$K = w_0^2 \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \ddots & \vdots \\ 0 & 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 2 & -1 \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{bmatrix} \ .$$

Difficult to solve, since the n coordinates in x are coupled by the matrix K. Decoupling operator: Orthogonal matrix Q diagonalizing K, that is, $Q'Q = I$, $Q'KQ = D = \mathrm{diag}(d)$, $d = [d_1, d_2, \ldots, d_n]'$.

Pre-multiplying the eq.above by Q',

$$Q'(Q\ddot{y}) + Q'K(Qy) = I\ddot{y} + Dy = 0 \ ,$$

n decoupled harmonic oscillators, $\ddot{y}_k + d_k y_k = 0$, in the new 'normal' coordinates, $y = Q'x$.

Solutions of (scalar) harmonic oscillators:
$y_k(t) = \sin(\varphi_k + w_k t)$, with
phase $0 \leq \varphi_k \leq 2\pi$, angular freq. $w_k = \sqrt{d_k}$.

The columns of matrix Q are the eigenvectors of matrix K, multiples of the un-normalized vectors z^k. The corresponding eigenvalues, $d_k = w_k^2$, for $j, k = 1 \ldots n$, are given by

$$z_j^k = \sin\left(\frac{jk\pi}{n+1}\right) \ , \quad w_k = 2w_0 \sin\left(\frac{k\pi}{2(n+1)}\right) \ .$$

- Some archetypical or prototypical decoupling operators correspond to Linear Algebra (dense, structured or sparse) matrix factorizations:
LU, Cholesky, QR, SVD, Jordan, NNMF, etc.

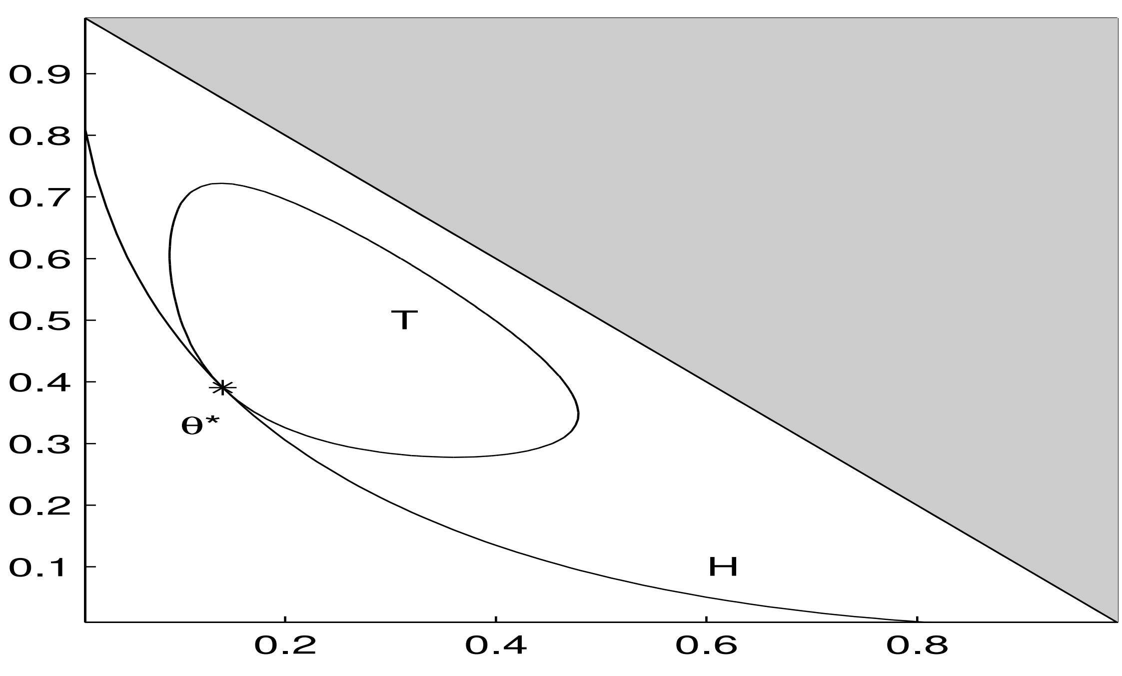
Example-III: Science, Scientific Hypothesis.
Hardy-Weinberg genetic equilibrium law,
the fixed point under panmixia condition.
n , sample size, x_1, x_3 , homozygote,
$x_2 = n - x_1 - x_3$, heterozygote count.

Posterior: $p_n(\theta \mid x) \propto \theta_1^{x_1+y_1} \theta_2^{x_2+y_2} \theta_3^{x_3+y_3}$,

Prior: $p_0(\theta) \propto \theta_1^{y_1} \theta_2^{y_2} \theta_3^{y_3}$, no (low) information.
Uniform: $y = \mathbf{0}$, MaxEnt: $y = -(1/2)\mathbf{1}$.

$$\Theta = \{\theta \geq 0 \mid \theta_1 + \theta_2 + \theta_3 = 1\} ,$$

$$H = \{\theta \in \Theta \mid \theta_3 = (1 - \sqrt{\theta_1}\ )^2\} .$$



Full Bayesian Significance Test (FBST) epistemic or evidence value, supporting and against hypothesis H, $\text{ev}(H)$ and $\overline{\text{ev}}(H)$.

$$s(\theta) = p_n(\theta) / r(\theta) \ ,$$

$$\widehat{s} = s(\widehat{\theta}) = \sup_{\theta \in \Theta} s(\theta) \ ,$$

$$s^* = s(\theta^*) = \sup_{\theta \in H} s(\theta) \ ,$$

$$T(v) = \{\theta \in \Theta \,|\, s(\theta) \le v\} \ , \quad \overline{T}(v) = \Theta - T(v) \ ,$$

$$W(v) = \int_{T(v)} p_n(\theta)\, d\theta \ , \quad \overline{W}(v) = 1 - W(v) \ ,$$

$$\text{ev}(H) = W(s^*) \ , \quad \overline{\text{ev}}(H) = \overline{W}(s^*) = 1 - \text{ev}(H).$$

$s(\theta)$ is the posterior surprise relative to $r(\theta)$. The tangential set, $\overline{T}(v)$, is the HRSS - Highest Relative Surprise Set above level v. The Wahrheitsfunktion or truth function, $W(v)$, is the cumulative surprise distribution.

Reference density - information metric in Θ. $r(\theta)$ represents no (or weak) information in Θ. If $r(\theta) \propto 1$, then $s = p_n(\theta)$, and HRSS is HPDS.

Homogeneous Disjunctive Normal Form, (HDNF): Compound hypotheses, $H^{i,j}$ (i-index) in independent models or structures, (j-index).

$$M^{(i,j)} = \{\Theta^j, H^{(i,j)}, p_0^j, p_n^j, r^j\} \ .$$

$$\text{ev}(H) \ = \ \text{ev}\left(\bigvee_{i=1}^{q} \bigwedge_{j=1}^{k} H^{(i,j)}\right) \ =$$

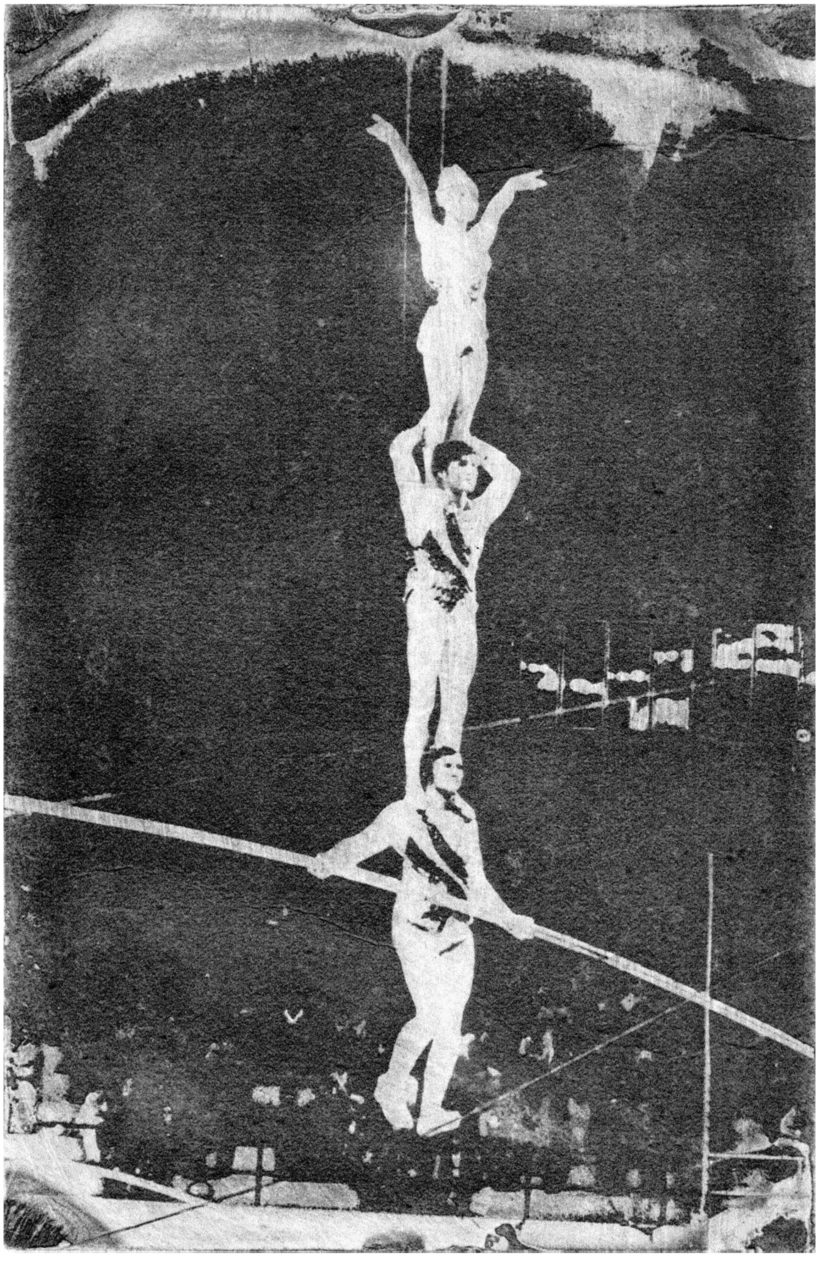
$$\max_{i=1}^{q} \text{ev}\left(\bigwedge_{j=1}^{k} H^{(i,j)}\right) \ =$$

$$W\left(\max_{i=1}^{q} \prod_{j=1}^{k} s^{*(i,j)}\right) \ .$$

$W = \bigotimes_{1 \leq j \leq k} W^j$. Mellin convol. $F(x) \otimes G(y)$ gives the distribution of the product $Z = XY$.

- Wittgenstein's truth value, function and operation: ev(H), $W(v)$ and $\otimes$.
- Possibilistic (partial) support structure: max Summation; arithmetic Multiplication.
- If all hypotheses are very unlikely or likely, that is, all $s^* \approx 0 \vee \widehat{s}$, then ev $\approx 0 \vee 1$, and FBST belief calculus $\approx$ classical logic.

Composing Complex Eigen-Solution Systems
Objects: Waves, Music; Hypotheses, Science.



Triadic (or Semiotic) Wire Walking. Etching
by Alex Flemming (untitled, 1979, PA III/X).

Complex and Non-Trivial Situations.

Autopoietic (Living) Organisms

Origami folding is *allopoietic*, from $\alpha\lambda\lambda o$-$\pi o\iota\eta\sigma\iota\varsigma$ - external production.

Organic morphogenesis is *autopoietic*, from $\alpha\upsilon\tau o$-$\pi o\iota\eta\sigma\iota\varsigma$ - self production.

No external supervision or correction mechanism and high environmental noise imply extra variability. Low precision of the ideally exact (spherical, cylindrical, etc.) symmetries.

In order to the final product to be viable, the process must be self-correcting and redundant.

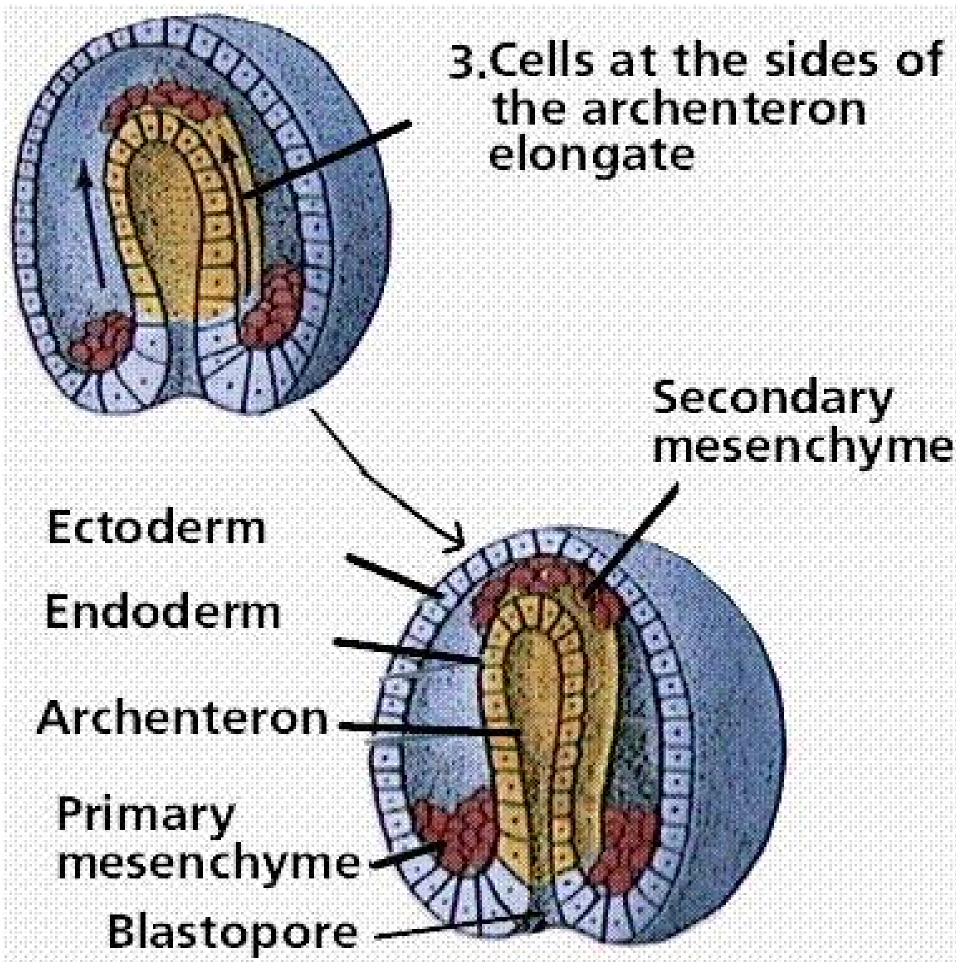
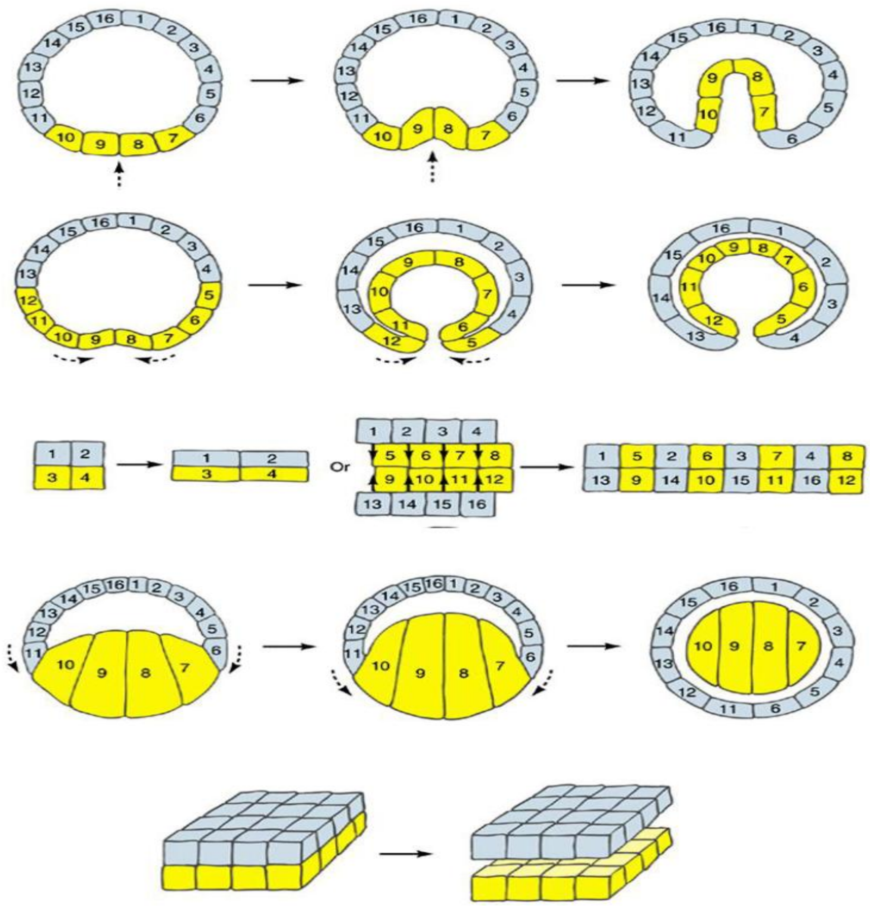


Figure of Gastrulation Process.
Organic morphogenesis:
(Un)Folding the Symmetries of life.



Tissue movements: Invagination, involution, convergent extension, epiboly, delamination.

Centralized vs. Decentralized Control:
In morphogenesis, there is no (external or internal) agent acting like a central controller. The complex forms and tissue movements at a global or macroscopic (supra cellular) scale are the result of collective cellular behavior patterns based on distributed control. The control mechanisms rely on simple local interaction between neighboring cells.

Object Orientation and Code Reuse:
At the microscopic level, cells of organic tissues are differentiated by distinct metabolic reaction patterns. However, the genetic code of any individual cell in an organism is identical, and cellular differentiation at distinct tissues are the result of differentiated (genetic) expressions of this sophisticated program.

Yoyo diagnostic problem: Repeated movements up and down the class hierarchy that may be required when the execution of a particular method invocation is traced.
Yoyo effects from code reuse + bootstapping.

Hypercyclic Bootstrapping.

This metaphor has a bad reputation, from Baron Münchhausen telling how he pulled himself out of a swamp by his own hair (or bootstraps). Nevertheless, it can and often does work!

1- *Tostines mystery: Does Tostines sell more because it is always fresh, or is it always fresh because it sells more?* (cookie brand)
Slogan of a brilliant marketing campaign.
The expression *Tostines mystery* became idiomatic in Brazilian Portuguese, playing the role of *bootstrapping* in English.

2- Bethe-Weizsäcker main catalytic cycle:
Nuclear synthesis of one atom of Helium from four atoms of Hydrogen. Carbon, Nitrogen and Oxygen act as catalysts (main source of energy in many stars). This nuclear physics fusion reaction shows that catalytic cycles are important even at scales much smaller than the typical examples from chemistry or biology.

Consequences of Hypercyclic Organization.

1- Metabolism. Thermodynamic arguments. No Perpetua Mobile. Organisms must extract raw materials, energy and order (information, neg-entropy) from their environment.

2- Exponential growth. Hypercycles (second or higher order auto-catalytic cycles) grow at exponential or even hyperbolic rate.

3- Evolution. Maltus' argument: Populations (finite temperature) outgrowing their resources imply competition, selection and evolution.

4- Hierarchy, Modularity and Building Blocks. Herbert Simon's argument:
- *The time required for the evolution of a complex form from simple elements depends critically on the number and distribution of potential intermediate stable subassemblies.*

Emergence and Asymptotics

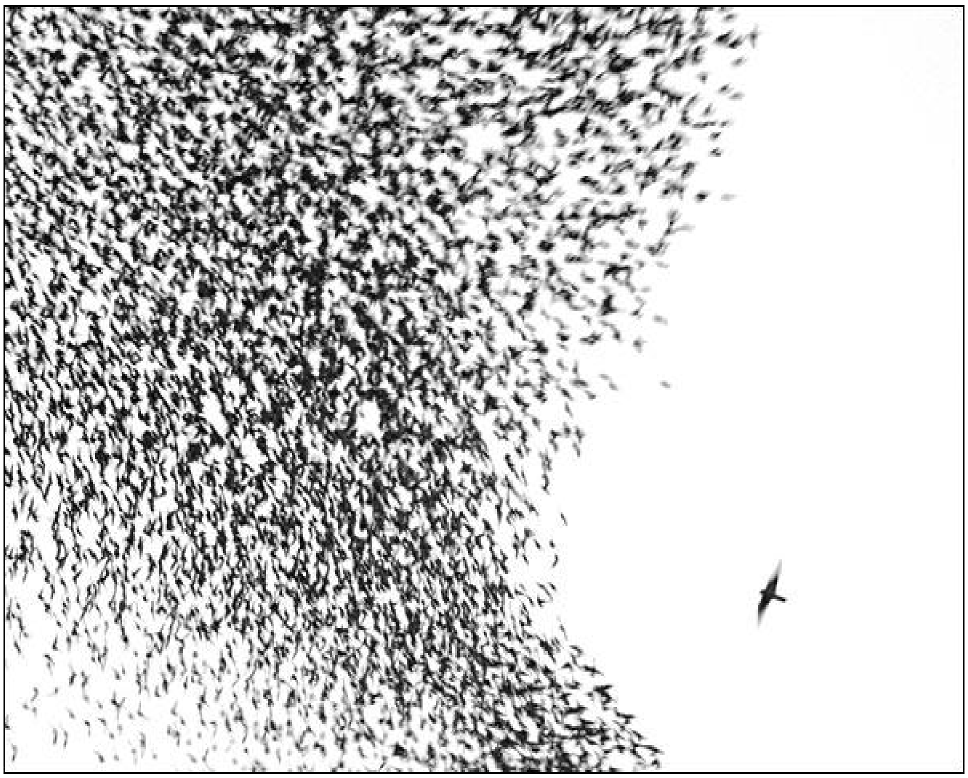
Asymptotic entities emerge from a model describing a local interaction in a small or microscopic scale, as a *law of large numbers*, i.e., as a stable behavior in the case of a system with many (asymptotically infinite) components.

Paradigmatic examples:

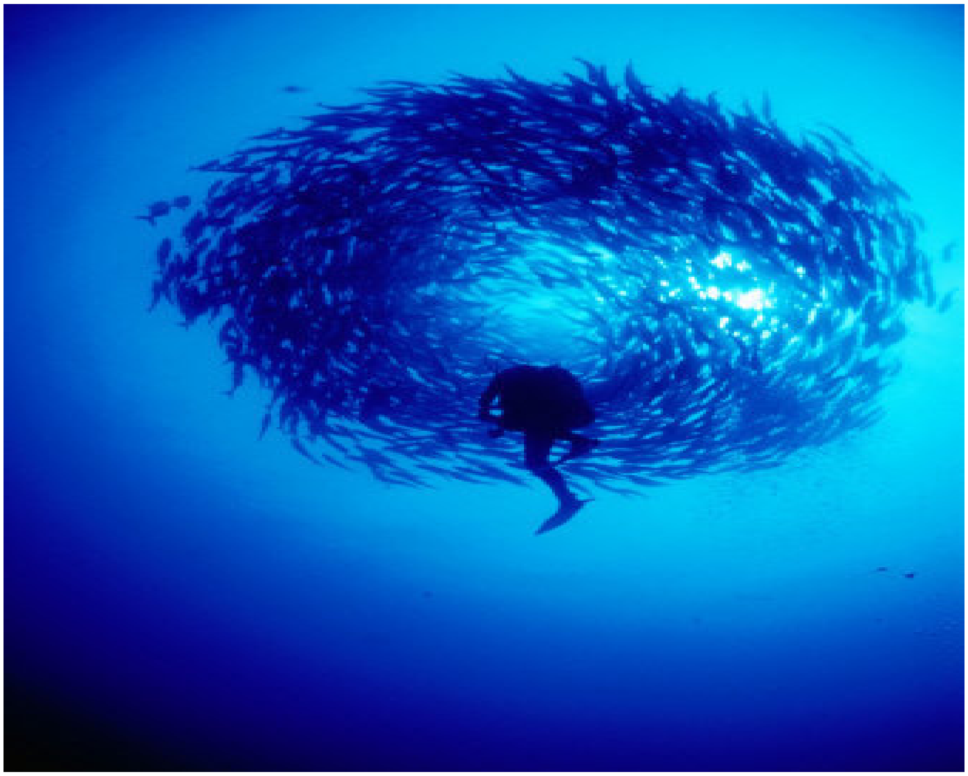
1- Macro-econometric relations on variables describing efficient markets, like aggregated supply, demand and price, form micro-economic models for interacting individual agents.

2- Thermodynamic laws on macroscopic variables, like volume, pressure and temperature of a gas, as asymptotic limits in statistical mechanics models for many interacting particles, like atoms or molecules.

Flocks, Schools and Swarms



- Flocking makes it difficult for a predator to use *local tactics* tracking the trajectory of a single individual, consequently, for a hunter that focus on local variables it is *hard to know* what exactly is going on.



- On the other hand, the same collective behavior creates the opportunity for *global strategies* that track and manipulate the entire flock. These hunting technique may be very efficient, in which case, we can say that the hunters *know very well* what they are doing.

Cognitive Constructivism (CogCon) Epistemological Framework

Experiment				Theory
Operation-alization	$\Leftarrow$	Experiment design	$\Leftarrow$	Hypotheses formulation
$\Downarrow$				$\Uparrow$
Effects observation		True/False eigen-solution		Creative interpretation
$\Downarrow$				$\Uparrow$
Data acquisition	$\Rightarrow$	Mnemetic explanation	$\Rightarrow$	Statistical analysis
Sample space				Parameter space

Figure 1: Scientific production diagram.

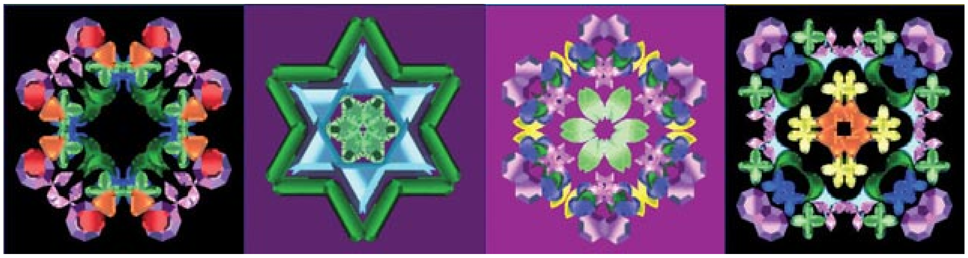
Scientific knowledge as eigen-solutions of a dynamic research process based on a scientific program (code and media, Luhmann). The whole system constitutes an autopoietic unit.

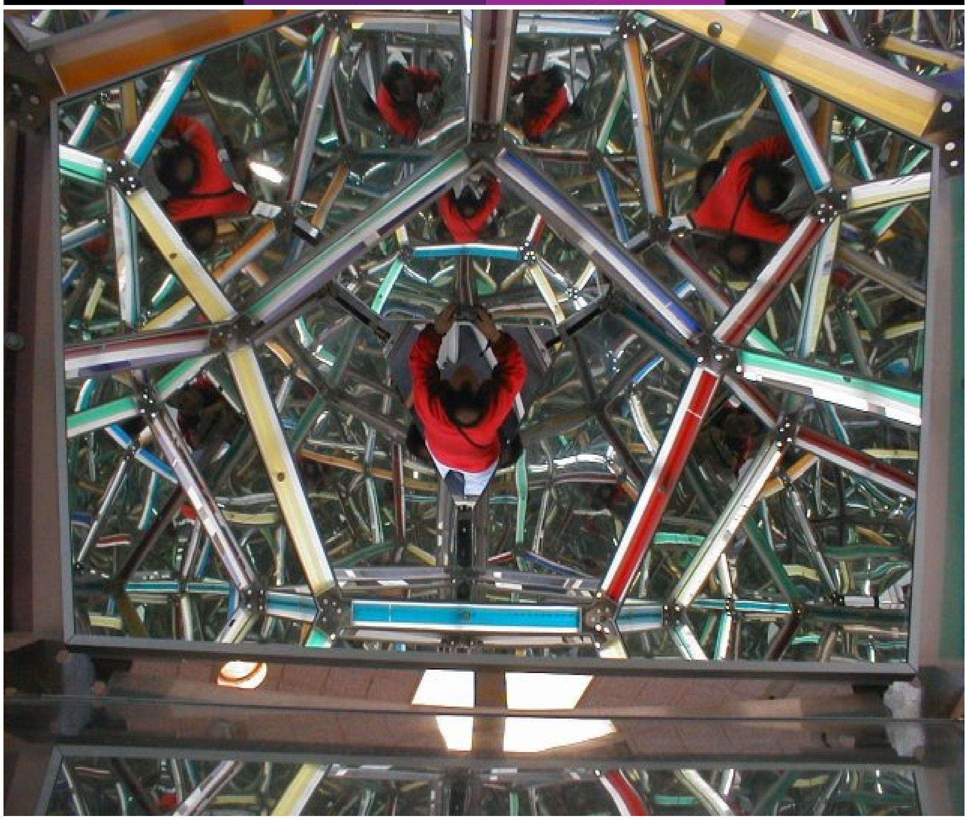
Cognitive Constructvist (CogCon) Ontology:

(1) *Known (knowable) Object:* An actual (or potential) eigen-solution of a given system's interaction with its environment. In the sequel, we may use a somewhat more friendly terminology by simply using the term Object.

(2) *Objective (how, less, more):* Degree of conformance of an object to the essential attributes of an eigen-solution (to be precise, stable, separable and composable).

(3) *Reality:* A (maximal) set of objects, as recognized by a given system, when interacting with single objects or with compositions of objects in that set.





Kaleido
-scopes
and
Mirror
houses

Kaleidoscopes are characterized by exact symmetries derived from Division Algebras (complex, quaternions, octonions). The same algebras describe our motions (rotations and translations) and re-appear in many physical laws.

Mirror Houses in funfairs and amusement parks. Entertainment from misleading ways in which a subject sees himself and other objects, or misperceptions about how or where himself stands in relation to other objects.

Kaleid.and Mirror houses as metaphors for the analysis of interesting philosophical problems.

Mirror House ex: Do we keep finding division algebras everywhere out there when trying to understand the physical universe because we already have the appropriate hardware to see them, or is it the other way around?
We suspect that any trivial choice in the dilemma posed by this trick question (yoyo), will only result in an inappropriate answer!

Kaleidoscope example:
Can Probability be Objective?
Yes! - Probability models generate (-ed as) very objective eigen-solutions (precise, stable, separable and composable). Instances:

1- Statistical distributions can be derived from invariance properties of the model, examples:
-Poisson dist. from time homogeneity.
-Multinomial dist. from exchangeability, etc.
Analogy of DeFinetti and Noether theorems: from symmetries to "real" (physical) objects.

Complementarity models, overcoming conceptual distinctions, dichotomies, dilemmas:

2- Convexification, in a continuous space, of a set of isolated possibilities, in a discrete space.
Ex: Mixed (random) strategies given by a probability, $p = [0, 1]$, of taking the pure strategies 0 or 1 in the odd-even game. The convex version of the game has an equilibrium point, $\{0.5, 0.5\}$, while the discrete version has none.
- Overcome discrete logic flip-flop dilemma.

3- Asymptotic emergence:
Continuous vs.discrete, thermodynamic vs.atomic models in Statistical Physics.

4- Niels Bohr direct complementarity:
Wave vs.particle characteristics in Quantum Mechanics by language of Fourier analysis.

5- Evolution w.reliable reproduction vs.creative innovation by stochastic optimization, using the language of inhomogeneous Markov chains, ex: simulated annealing, genetic programming. Objects? Stable modules or building blocks. Emergent semantics w.consistent interpret.

6- Niels Bohr ideas on complementarity in Biology, Psychology, Sociology etc. Dilemmas on liberty, free will, etc.

- **References:**

- S.J.Blackmore, R.Dawkins (1999). *The Meme Machine.* Oxford University Press.
- N.H.D.Bohr (1987-1999). *The Philosophical Writings of Niels Bohr.* V.I-IV. Ox Bow Press.
- W.Borges, J.M.Stern (2007). The Rules of Logic Composition for the Bayesian Epistemic e-Values. *Logic Journal of the IGPL*, 15, 5-6, 401-420.
- E.C.Colla, J.M.Stern (2008). Sparse Factorization Methods for Inference in Bayesian Networks. *AIP Conference Proceedings,* 1073, 136-143.
- A.Y.Darwiche (1993). *A Symbolic Generalization of Probability Theory.* Ph.D. Thesis, Stanford Univ.
- M.Eigen, P.Schuster (1979). *The Hypercycle: A Principle of Natural Self-Organization.* Springer-Verlag.
- J.N.Franklin (1968). *Matrix Theory.* Prentice-Hall.
- B.Goertzel, O.Aam, F.T.Smith, K.Palmer (2008). Mirror Neurons, Mirrorhouses, and the Algebraic Structure of the Self, *Cybernetics & Human Knowing,* 15,1,9-28.
- B.Goertzel (2007). Multiboundary Algebra as Pregeometry. *Electr.J.of Theoretical Physics,* 16,11,173-186.
- G.J.Klir, T.A.Folger (1988). *Fuzzy Sets, Uncertainty and Information.* Prentice Hall.
- M.Lauretto, C.A.B.Pereira, J.M.Stern, S.Zacks (2003). Full Bayesian Significance Test Applied to Multivariate Normal Structure Models. *REBRAPE,* 17, 147-168.
- M.Madruga, L.Esteves, S.Wechsler (2001). On the Bayesianity of Pereira-Stern Tests. *Test*, 10, 291-299.
- R.Nagpal (2002). Self-assembling Global Shape using Concepts from Origami. p.219-231 in T.C.Hull (2002). Proc.3rd Meet.of Origami Math.Science and Education.

- C.A.B.Pereira, J.M.Stern, (1999). Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses. *Entropy Journal*, 1, 69–80.
- C.A.B.Pereira, J.M.Stern (2008). Special Characterizations of Standard Discrete Models. *REVSTAT Statistical Journal*, 6, 3, 199-230.
- C.A.B.Pereira, S.Wechsler, J.M.Stern (2008). Can a Significance Test be Genuinely Bayesian? *Bayesian Analysis,* 3, 1, 79-100.
- M.Springer (1979). *The Algebra of Random Variables.*
- J.M.Stern (2003). Significance Tests, Belief Calculi, and Burden of Proof in Legal and Scientific Discourse. Laptec-2003, *Frontiers in Artificial Intelligence and its Applications,* 101, 139–147.
- J.M.Stern (2004). Paraconsistent Sensitivity Analysis for Bayesian Significance Tests. SBIA'04, *Lecture Notes Artificial Intelligence,* 3171, 134-143.
- J.M.Stern (2006). Language, Metaphor and Metaphysics: The Subjective Side of Science. Tech.Rep. MAC-IME-USP-2006-09.
- J.M.Stern (2007). Cognitive Constructivism, Eigen-Solutions, and Sharp Statistical Hypotheses. *Cybernetics & Human Knowing*, 14, 1, 9-36.
- J.M.Stern (2007). Language and the Self-Reference Paradox. *Cybernetics & Human Knowing*, 14,4,71-92.
- J.M.Stern (2008). Decoupling, Sparsity, Randomization, and Objective Bayesian Inference. *Cybernetics & Human Knowing*, 15, 2, 49-68.
- J.M.Stern (2008). *Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses.* Tutorial book for MaxEnt 2008, Boracéia, Brazil.

FBST Decalogue (Ten Commandments):

1- Give an intuitive and simple measure of significance for sharp hypotheses, a **probability** defined in the original **parameter space**.

2- Have an intrinsically geometric definition, independent of any non-geometric aspect, like:
- The hypothesis (manifold) parameterization,
- The coordinate system on the parameter space,
i.e., be an **invariant** procedure.

3- Give a smooth measure of significance, i.e. **continuous and differentiable**, on the hypothesis parameters and sample statistics, under appropriate regularity conditions.

4- **Likelihood principle**, i.e., the information gathered from observations should be represented (only) by the likelihood function.

5- Be able to provide an **exact** procedure, making no use of "large sample" asymptotic approximations.

6- Require **no ad hoc prior** information that could lead to judicial contention, like a positive probability mass on a zero measure set, or a belief ratio between hypotheses, etc.

7- Be able to provide a **consistent** test for a given sharp hypothesis, in the sense that increasing sample size should make it converge to the right accept/reject decision.

8- Allow, (only) if desired, the incorporation of previous experience or expert opinion via a **subjective prior** distribution.

9- Provide a **possibilistic** support function.

10- Provide **compositionality** operations in complex models.

Invariance:
Reparameterization of H (of $h(\theta)$): Trivial.
Reparameterization of Θ, (regularity cond.= bijective, integrable, a.s.cont.differentiable)

$$\omega = \phi(\theta) \quad , \quad \Omega_H = \phi(\Theta_H)$$

$$J(\omega) = \left[\frac{\partial\,\theta}{\partial\,\omega}\right] = \left[\frac{\partial\,\phi^{-1}(\omega)}{\partial\,\omega}\right] = \begin{bmatrix} \frac{\partial\,\theta_1}{\partial\,\omega_1} & \cdots & \frac{\partial\,\theta_1}{\partial\,\omega_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial\,\theta_n}{\partial\,\omega_1} & \cdots & \frac{\partial\,\theta_n}{\partial\,\omega_n} \end{bmatrix}$$

$$\widetilde{s}(\omega) = \frac{\widetilde{p}_n(\omega)}{\widetilde{r}(\omega)} = \frac{p_n(\phi^{-1}(\omega))\,|J(\omega)|}{r(\phi^{-1}(\omega))\,|J(\omega)|}$$

$$\widetilde{s}^* = \sup_{\omega\in\Omega_H} \widetilde{s}(\omega) = \sup_{\theta\in\Theta_H} s(\theta) = s^*$$

hence, $T(s^*) \mapsto \phi(T(s^*)) = \widetilde{T}(\widetilde{s}^*)$, and

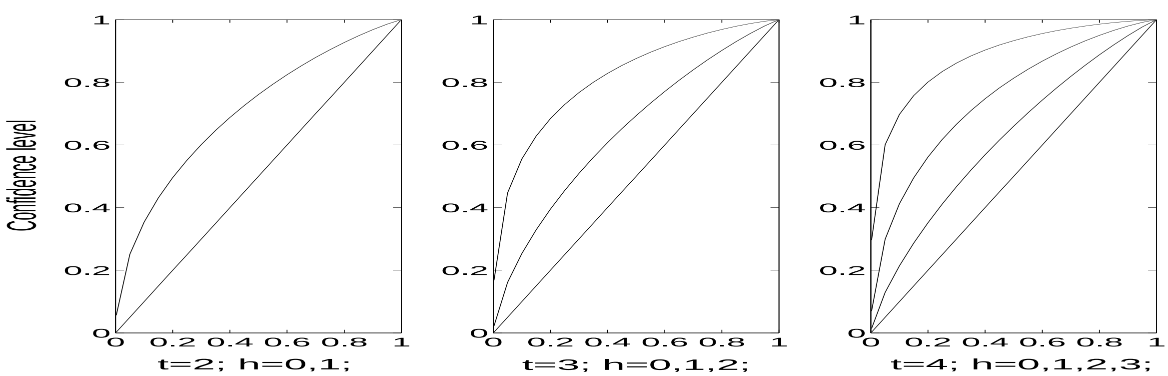
$$\widetilde{\mathsf{Ev}}(H) = \int_{\widetilde{T}(\widetilde{s}^*)} \widetilde{p}_n(\omega) d\omega =$$

$$\int_{T(s^*)} p_n(\theta) d\theta = \mathsf{ev}(H) \quad , \quad \text{Q.E.D.}$$

Critical Level and Consistency:
$\overline{V}(c) = \Pr(\overline{\text{ev}} \leq c)$, the cumulative distribution of $\overline{\text{ev}}(H)$, given θ^0, the true parameter value.
Let $t = \dim(\Theta)$ and $h = \dim(H)$.
Under appropriate regularity conditions, for increasing sample size, $n \to \infty$,

- If H is false, $\theta^0 \notin H$, then $\overline{\text{ev}}(H) \to 1$
- If H is true, $\theta^0 \in H$, then $\overline{V}(c)$, the confidence level, is approximated by the function

$$\mathsf{Q}\left(t-h, \mathsf{Q}^{-1}(t,c)\right) \ .$$



Test τ_c critical level vs. confidence level

Alternative approaches: Empirical power analysis Stern and Zacks (2002) and Lauretto (2004); Decision theory, Madruga (2001); Sensitivity analysis (paraconsistent logic), Stern (2004).

Decision-theoretic (orthodox) Bayesian view, Dubins and Savage (1965):

"Gambling problems embrace the whole of (decision) theoretical statistics."

Epistemic questions about H are questions on How to Gamble on H (against an H').

A significance test is legitimate if and only if it can be characterized as an Acceptance (A) or Rejection (R) decision procedure defined by the minimization of the posterior expectation of a loss function, Λ.

FBST loss funct. based on indicator functions of θ being or not in the tangential set $\overline{T}$:

$$\Lambda(R,\theta) = a\, I(\theta \notin \overline{T})\ , \quad \Lambda(A,\theta) = b + d\, I(\theta \in \overline{T})\ .$$

Note that Λ depends on the observed sample (via the likelihood function), on the prior, and on the reference density, stressing the important point of non separability of utility and (prior) probability.

Comparative example:
Independence in 2×2 contingency table.

$$H : \theta_{1,1} = (\theta_{1,1} + \theta_{1,2})(\theta_{1,1} + \theta_{2,1}) \ .$$

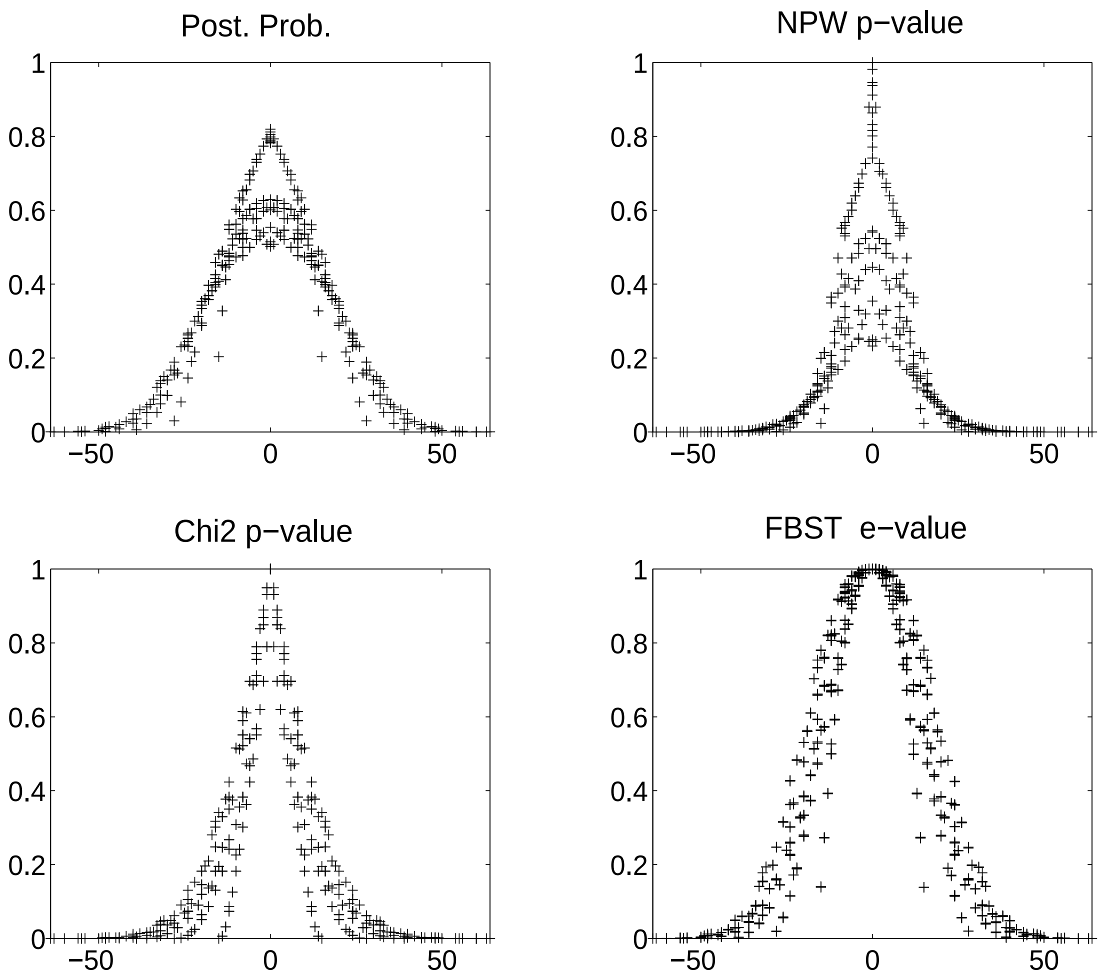
Next Figure compares four statistics, namely,
-Bayes factor posterior probabilities (BF-PP),
-Neyman-Pearson-Wald (NPW) p-values,
-Chi-square approximate p-values, and the
-FBST evidence value in favor of H.

Horizontal axis: $D =$ diagonal asymmetry,
is the unormalized Pearson correlation,

$$D = x_{1,1}x_{2,2} - x_{1,2}x_{2,1} \ ,$$

$$\rho_{1,2} = \frac{\sigma_{1,2}}{\sigma_{1,1}\sigma_{2,2}} = \frac{\theta_{1,1}\theta_{2,2} - \theta_{1,2}\theta_{2,1}}{\sqrt{\theta_{1,1}\theta_{1,2}\theta_{2,1}\theta_{2,2}}} \ .$$

Wish list:
- Full symmetry in X gives H full support.
- ev(H) continuous and differentiable.
- Compositionality rules.



Independence Hypothesis, n=16

- NPW p-value: Gives support for X given H, not for H given data! Consequence: Different supports (< 1) for X exhibiting full symmetry.
- Post.Prob: Maximum support, $s_{max} < 1$, needs to be calibrated by ad hoc (Lebesgue) measures, atomic masses, special priors, etc. Does not work! Lindley paradox, more priors...

Abstract Belief Calculus - Darwiche (1993)

$\langle \Phi, \oplus, \oslash \rangle$, Support Structure,
Φ , Support Function, for statements on $\mathcal{U}$.
Null and full support values are **0** and **1**.
$\oplus$, Support Summation operator,
$\oslash$, Support Scaling or Conditionalization,
$\langle \Phi, \oplus \rangle$, Partial Support Structure.

$\oplus$, gives the support value of the disjunction of any two logically disjoint statements from their individual support values,

$$\neg(A \wedge B) \Rightarrow \Phi(A \vee B) = \Phi(A) \oplus \Phi(B) \ .$$

$\oslash$, gives the conditional support value of B given A from the unconditional support values of A and the conjunction $C = A \wedge B$,

$$\Phi_A(B) = \Phi(A \wedge B) \oslash \Phi(A) \ .$$

Unscaling: $\Phi(A \wedge B) = \Phi_A(B) \otimes \Phi(A)$.

Support structures for some belief calculi,
$a = \Phi(A)$, $b = \Phi(B)$, $c = \Phi(C = A \wedge B)$.

$\Phi(\mathcal{U})$	$a \oplus b$	$\mathbf{0}$	$\mathbf{1}$	$a \preceq b$	$c \oslash a$	
$[0,1]$	$a+b$	0	1	$a \leq b$	c/a	Pr
$[0,1]$	$\max(a,b)$	0	1	$a \leq b$	c/a	Ps
$\{0,1\}$	$\max(a,b)$	0	1	$a \leq b$	$\min(c,a)$	CL
$\{0..\infty\}$	$\min(a,b)$	∞	0	$b \leq a$	$c-a$	DB

Pr= Probability, Ps= Possibility,
CL= Classical Logic, DB= Disbelief.

In the FBST setup, two belief calculi are in simultaneous use: ev constitutes a possibilistic partial support structure coexisting in harmony with the probabilistic support structure given by the posterior probability measure in the parameter space, see also Zadeh (1987).
See Klir (1988) on nesting property of $T(v)$.

Benefit of the Doubt, In Dubito Pro Reo, Onus Probandi, Presumption of Innocence, Safe Harbor Liability Rule:

- By this principle, one must consider the defendant's claim, H, in the most favorable way.
- This principle establishes that there is no liability as long as there is a reasonable basis for belief, placing the burden of proof on the plaintiff, who must prove H to be false.
- Prevents the plaintiff of making any assumption not explicitly stated by the defendant, or tacitly implied by existing law or regulation.

Hence: *Possibilistic* belief calculus, requiring *nesting focal sets*, entailing the FBST: "Witnesses" against H are the Focal Sets

$$\overline{T}(v) = \{\theta \in \Theta \,|\, s(\theta) > p_n(\theta^* \,|\, X)\}.$$

Most favorable θ is $\theta^* \in \arg\max_H p_n(\theta \,|\, X)$.

We Can Not integrate on H.
It is a classic, not a quantic defendant :-)