SmartData:

Make the data “think” for itself

Privacy and Security in a Virtual Web-world
IDIS Symposium – Rome
May 27, 2010

Presented by Ken Anderson on behalf of
George J. Tomko, Ph.D.
Expert-In-Residence
Identity, Privacy and Security Institute (IPSI)
University of Toronto, Canada

The Goal of SmartData

- Better privacy is not more security and regulations around an expanding perimeter of collective personal information.
- Better privacy is shrinking that perimeter down to one individual's personal information such that the person and his information are inseparable ….
- And the person is always in control.

Presentation Outline

1. *SmartData: The concept*
2. *SmartData: The security structure*
3. *SmartData: EHR application example*
4. *SmartData: The R & D strategy*
5. *SmartData: Open-source development*
6. *Conclusions*
7. *Discussion*

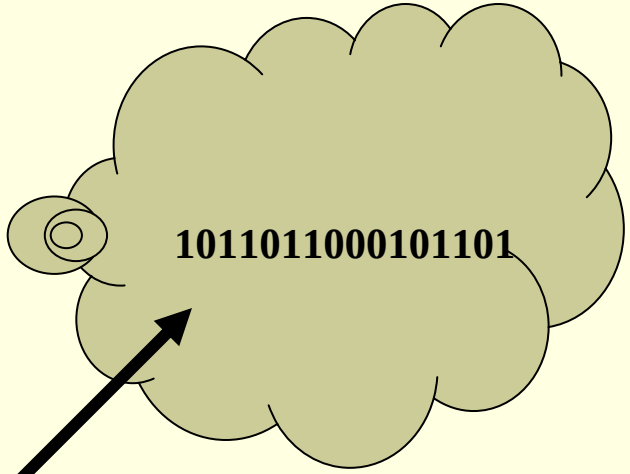
What is SmartData?

A Thought Experiment

Human SmartData

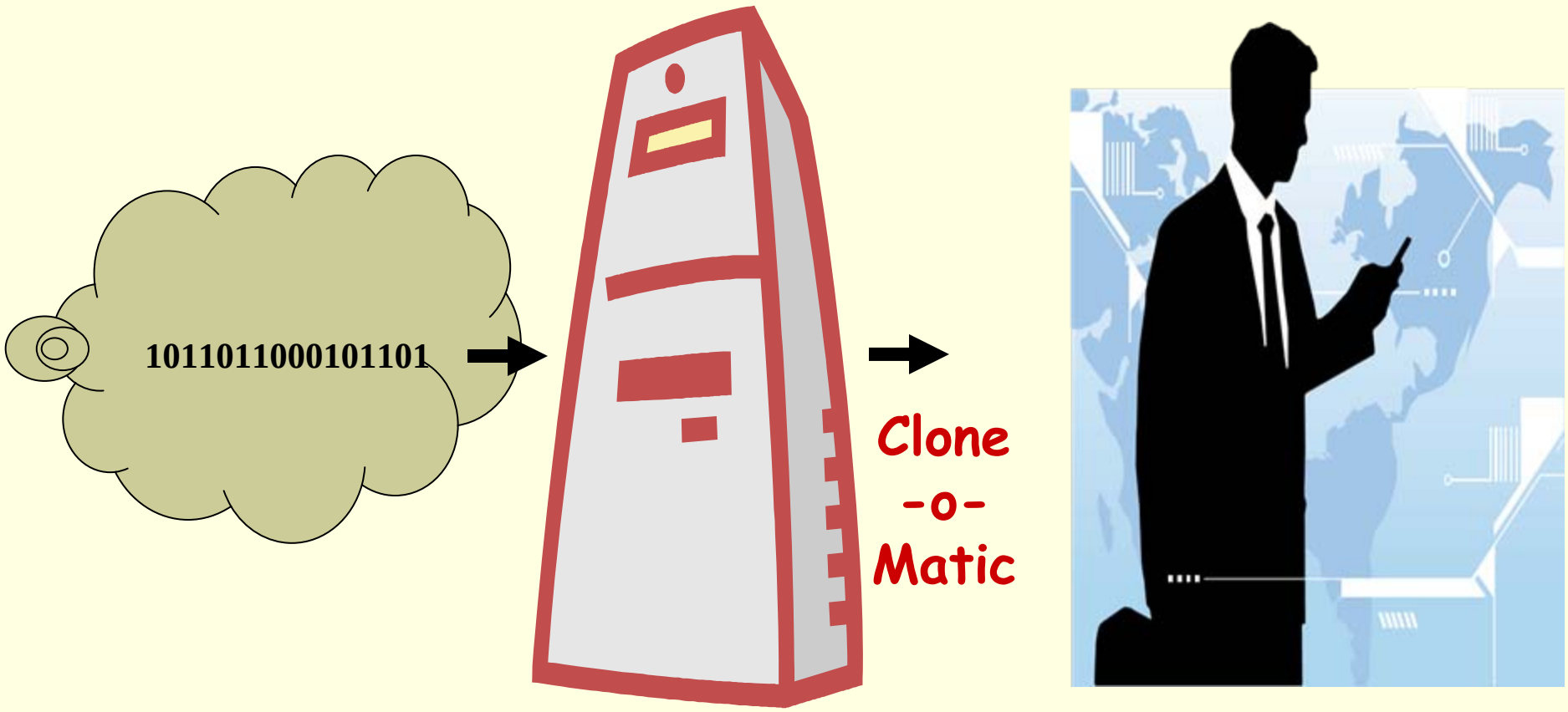
The Digital Human SmartData

Digitize information representing a human into a binary string



Stored in the “cloud”

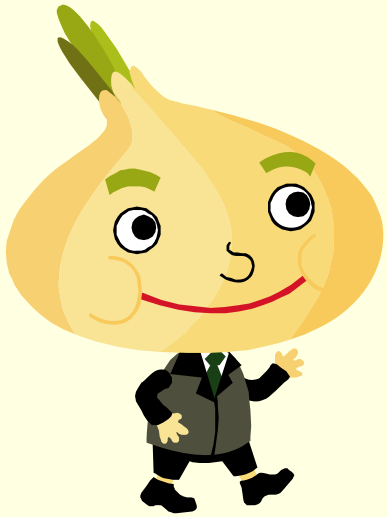
What if we reconstruct the human?



Features of SmartData

- discloses your information only when your personal criteria have been met;
- Protects and secures your personal information;
- Information can be released in a non-digital form;
- Make decisions about whether or not to disclose information based on context.
- Represent your privacy interests on the Web – your surrogate and proxy

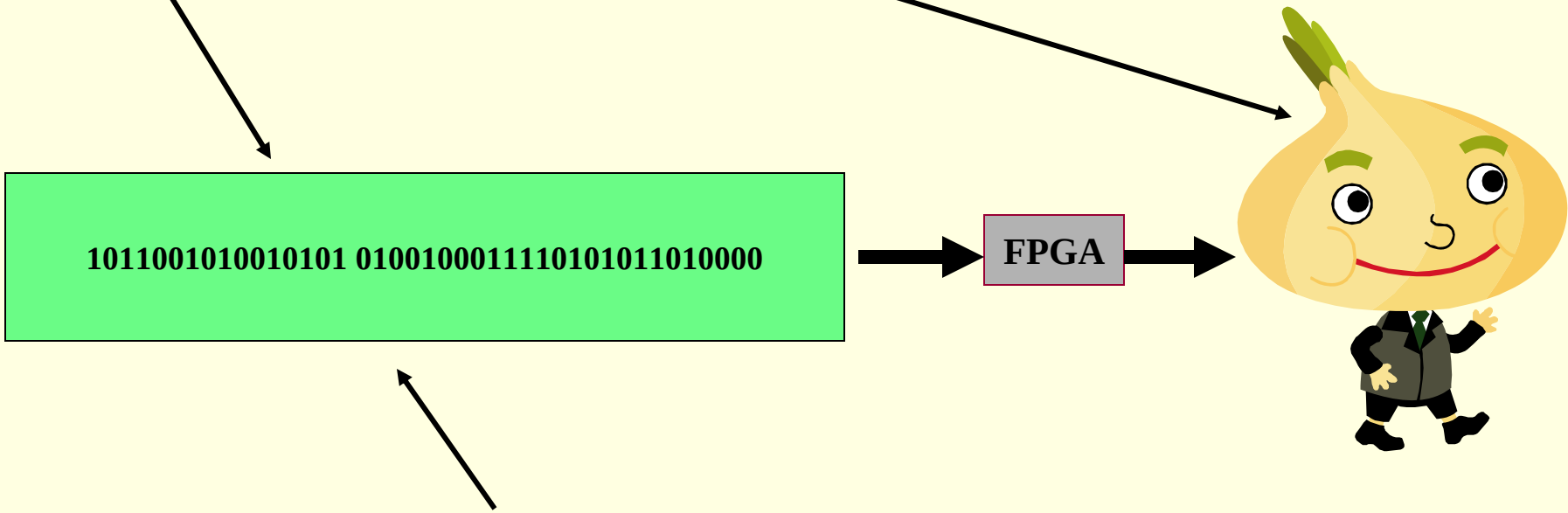
Substitute clone with an intelligent agent



SmartData

Structure of SmartData

Binary String represents the structure of the **SmartData agent**



Personal data “mixed in” with other information

SmartData Security Structure

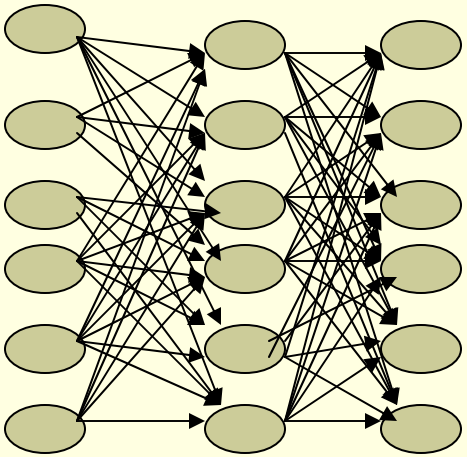
My Personal Data

Upper case: meta-tags
Lower case: Personal data

A[a], B[b], C[c], D[d], …..

a **b** **c** **d**

Meta-tags → A B C D



Neural Network

Contents of lock-boxes "mixed in" with other data across synaptic weights of the neural network

Authenticating

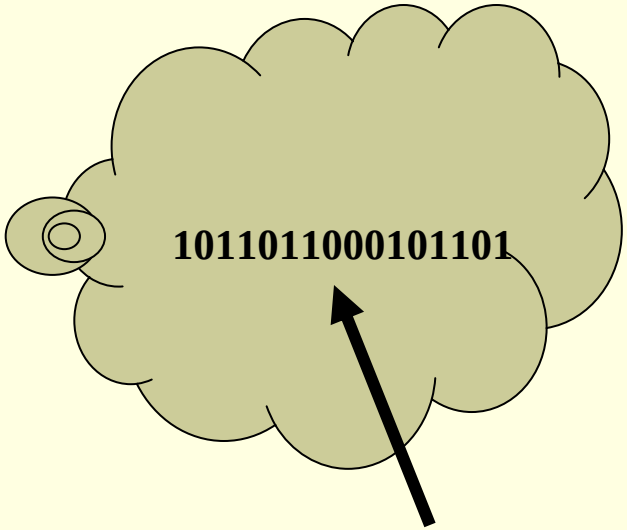
- SmartData: authenticate credentials of requestors
- Requestors: authenticate credentials of SmartData
  - Digital signatures and biometrics

Analog output option

- Digital-to-analog or digital-to-image within SmartData

No Personal information in the cloud: Just SmartData

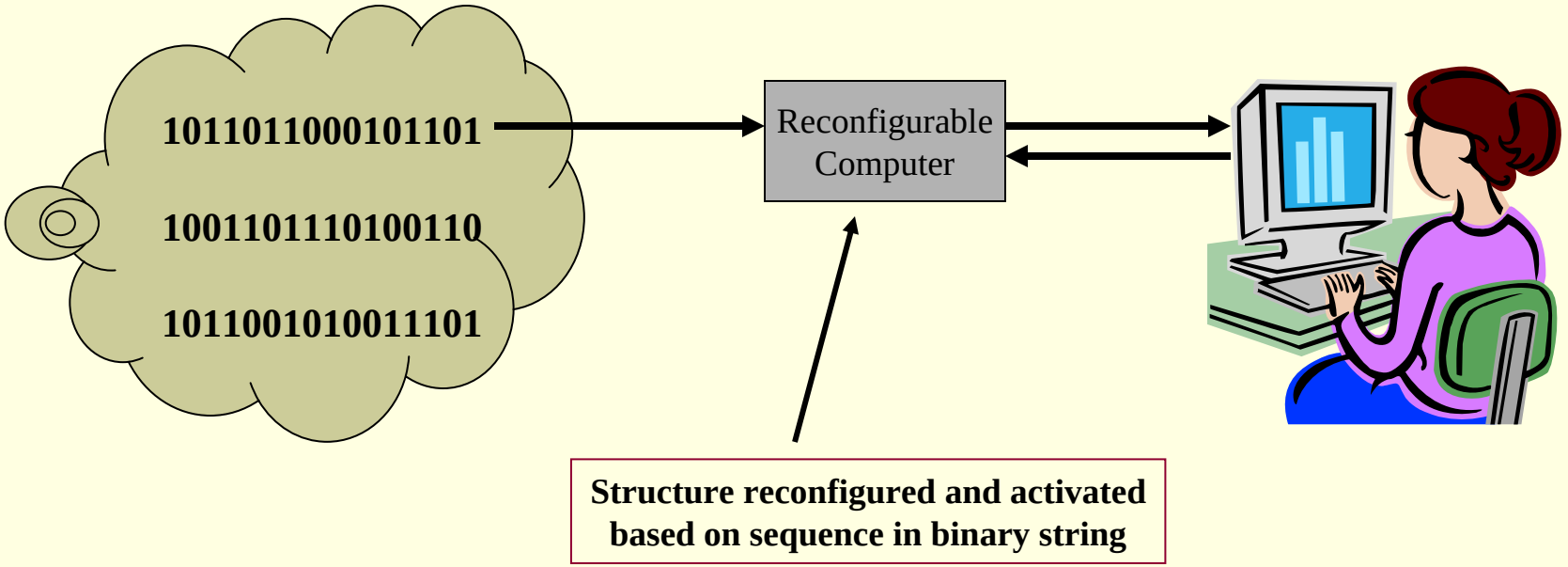
- **Only SD binary string is transmitted**



SmartData binary string – personal information locked inside

- There would be no personal or proprietary “raw” data out in the open.
- It would instead be housed “within” a SmartData agent

SmartData as an Electronic Health Record



Houston, we have a problem!

- Algorithms that incorporate contextual inputs across many domains are subject to computational intractability.
- Computationally practical algorithms are not practical for business purposes.
- Solution: Copy nature and evolve

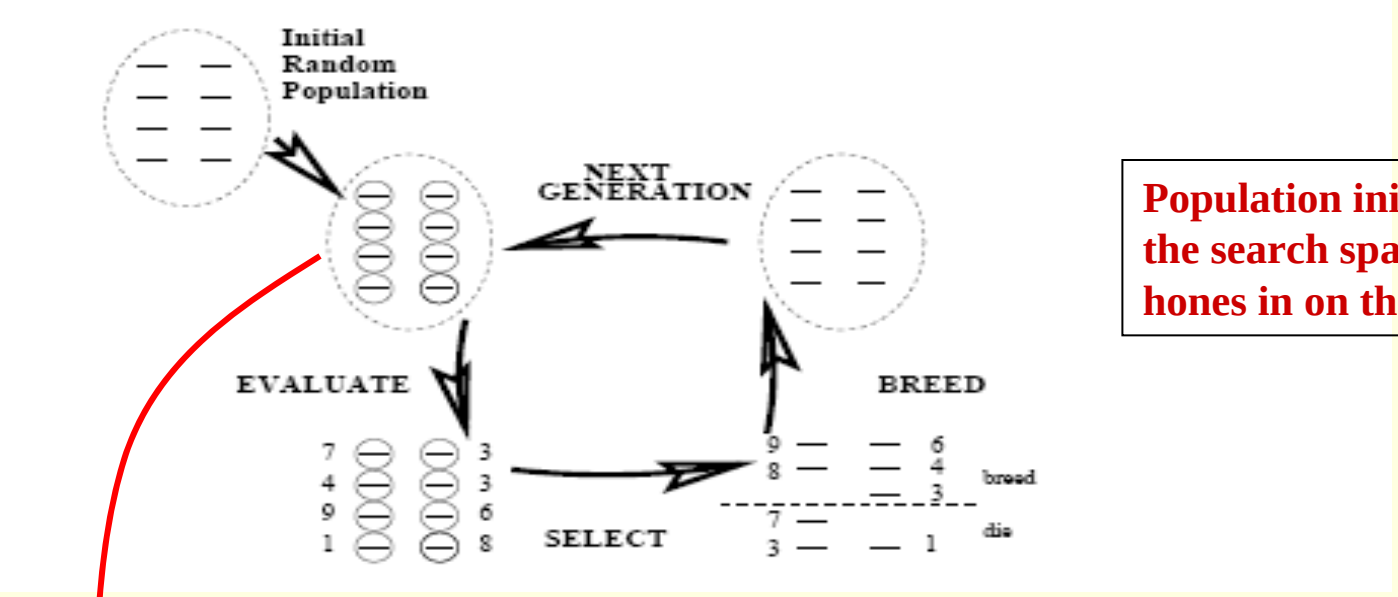
Our Approach

Evolutionary Embodied Cognition within a dynamical systems framework

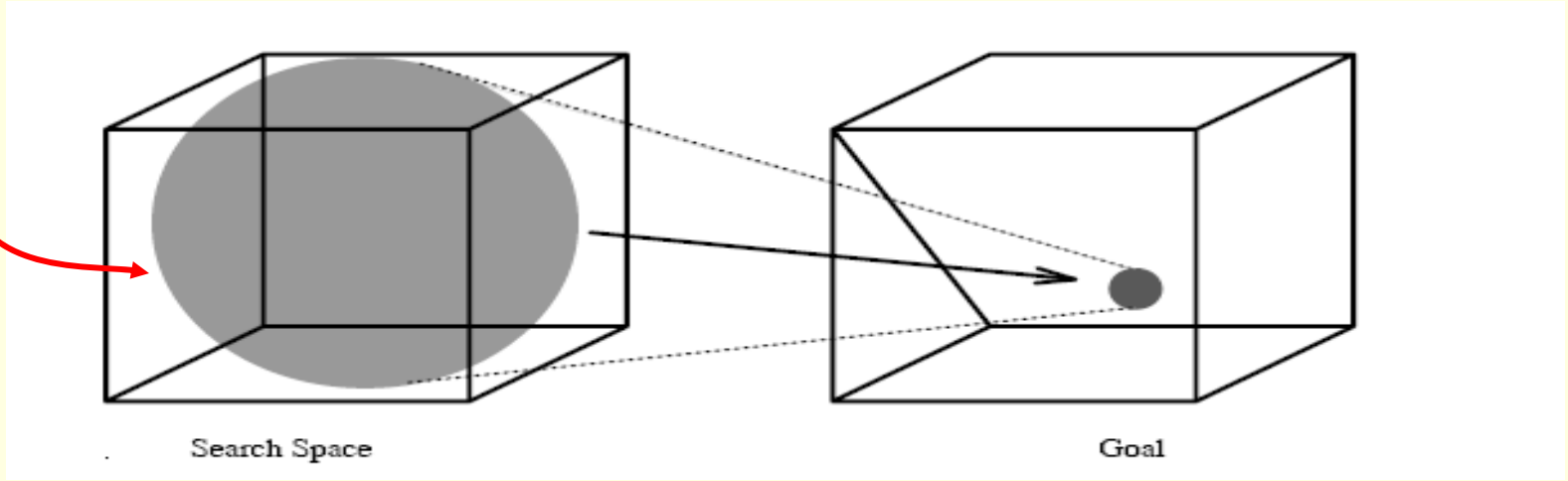
Evolutionary Robotics

- Uses principles of natural evolution to create artificial agents.
- Bottom-up methodology versus top-down as in the field of Artificial Intelligence.
- No initial design – only an initial design objective.

The Genetic Algorithm Cycle



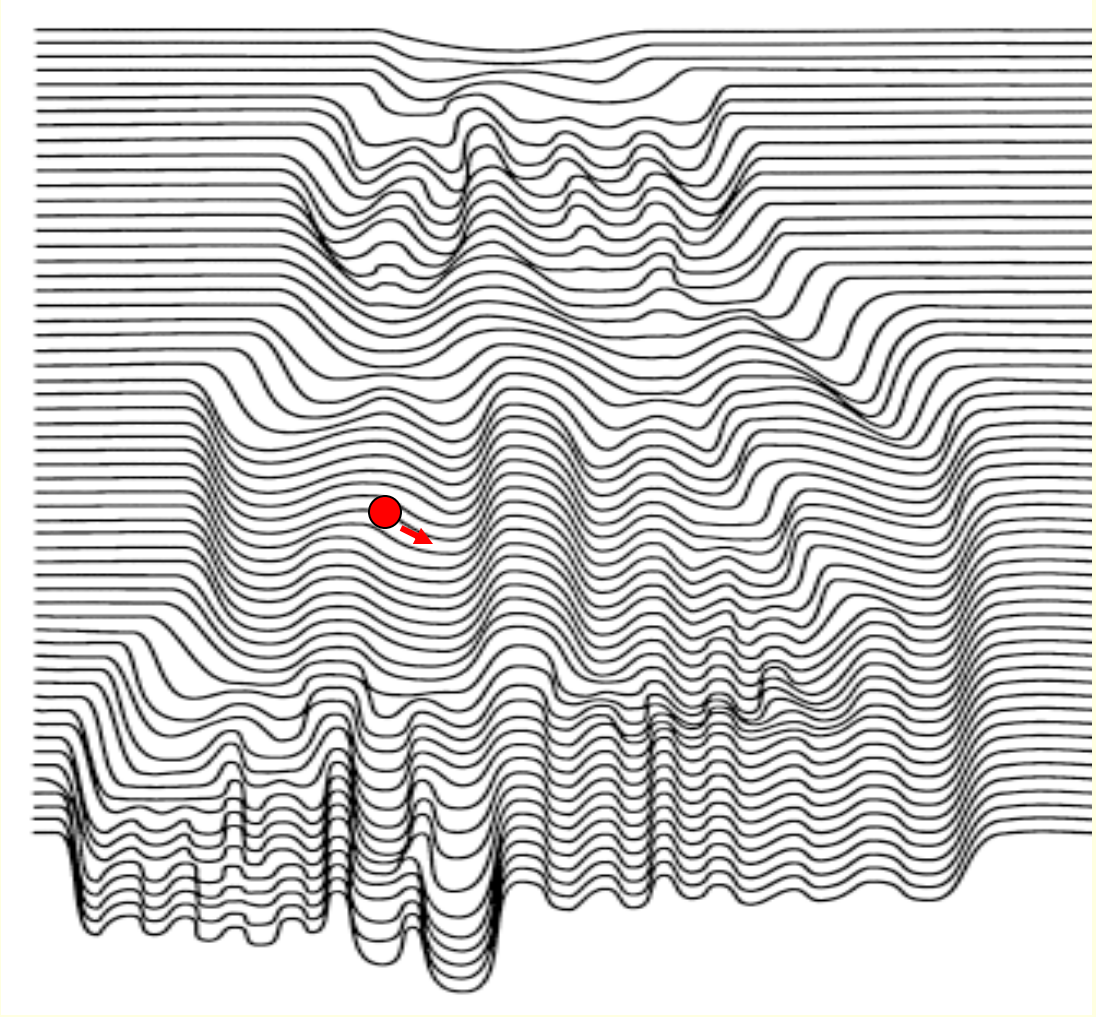
Population initially spans the search space and progressively hones in on the optimum



Embodied Cognition

- Contents and operations of cognition are determined by the whole body and the environment in which the body is situated.
  - Not just the brain alone.
  - Physical, “organismic”, and conceptual embodiment.
- The body is the active interface to the world.
  - transforms physical variables in the environment via the sensors into neural control system parameters.
  - converts neural variables via motor action into environmental parameters.

Dynamical System Landscape

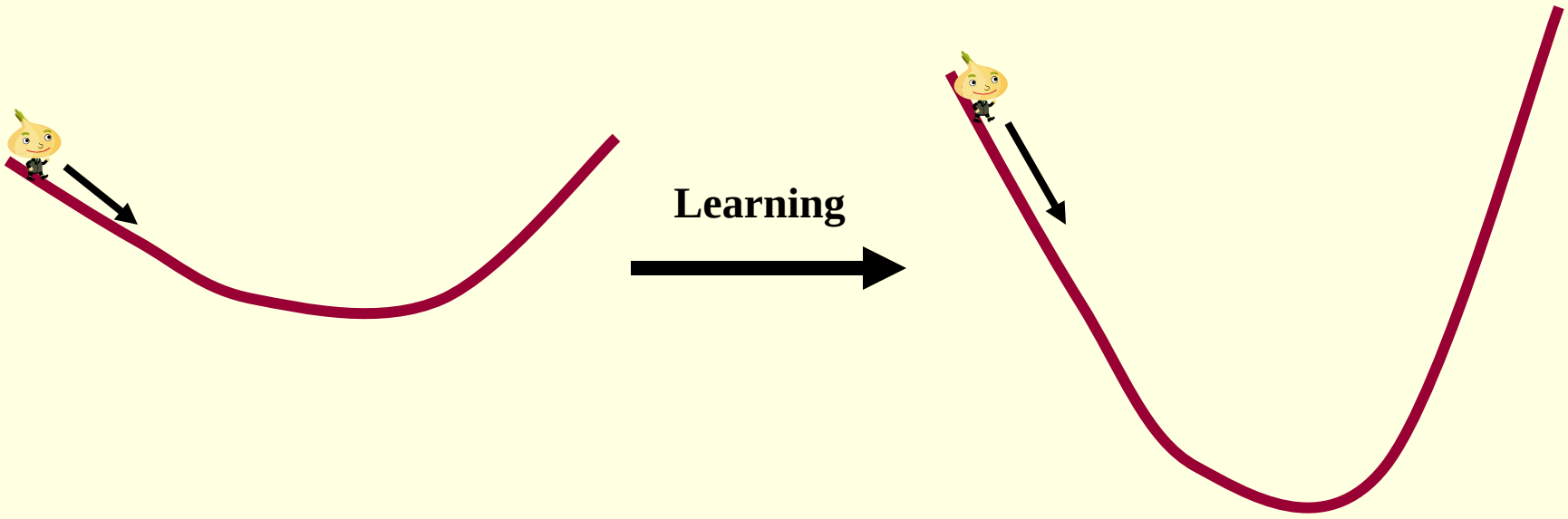


Life is Just a Journey

- In the dynamical systems framework, behaviours (perception, motor, language, thinking, reasoning) and memories are one and the same.
- They are trajectories in an appropriate basin of attraction
- There is no information in the brain, only dynamic flows and operators.

And Learning is …

- about changing shallow basins of attraction into deep basin that are more stable to change



Embodied Dynamical Systems Framework

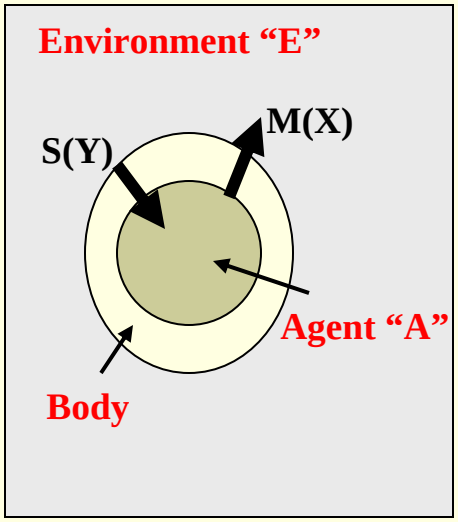
Simulate E

Evolve A

$$\frac{dX}{dt} = A\left(X; S\left(Y\right); U\right)$$

$$\frac{dY}{dt} = E\left(Y; M\left(X\right); V\right)$$

Coupling Parameters



Where:

A = Agent's transition map;

E = Environment's transition map;

X = Output variable of Agent's neurons;

Y = Output variables of environment;

S(Y) = transformation of environment's variables into sensory parameters;

M(X) = transformation of Agent's variables into motor parameters that affect the environment;

U = Agent's internal parameters;

V = Environment's parameters

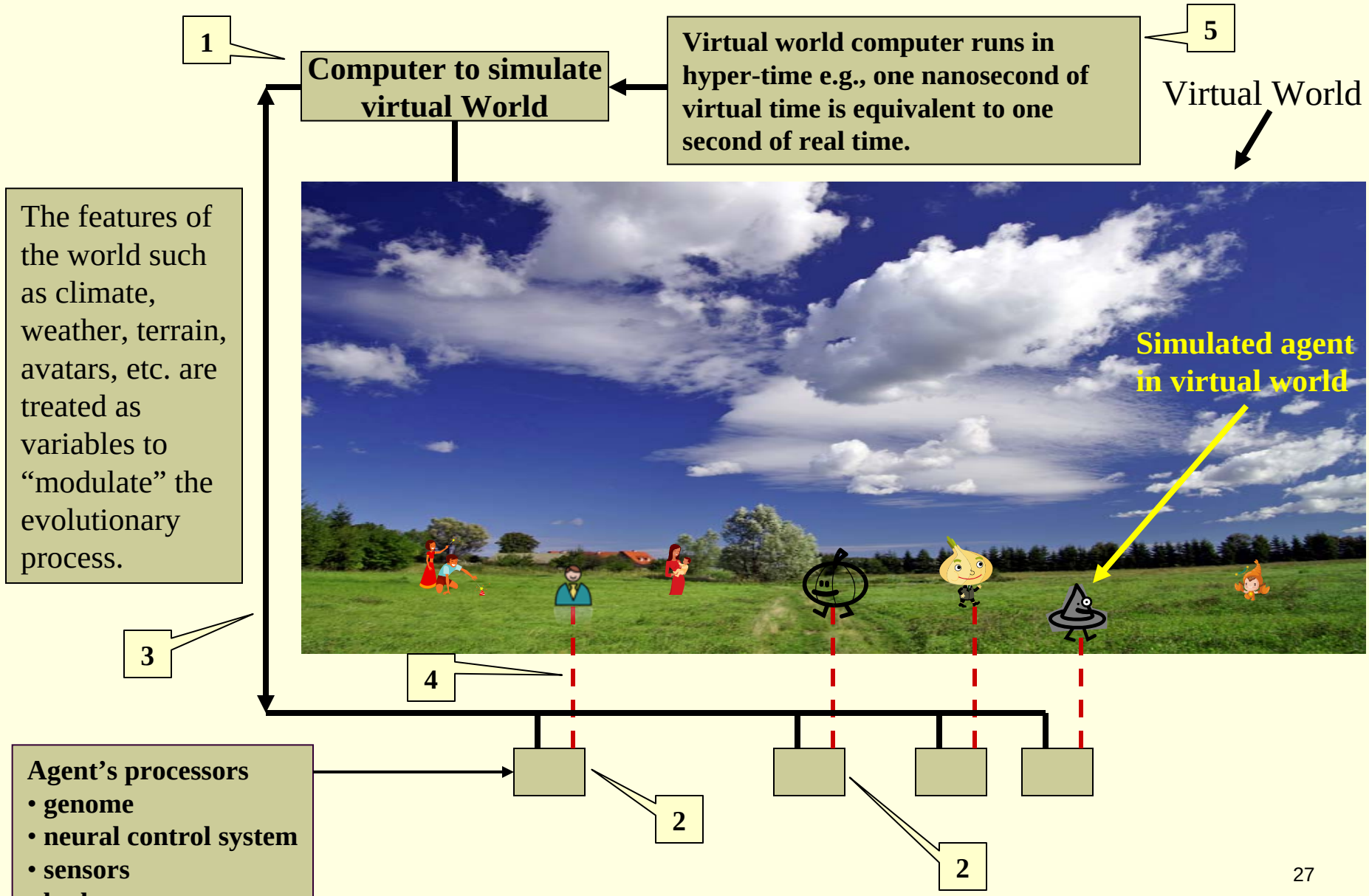
Simulation of Virtual Worlds

- Original internet (text) ---- One dimensional.
- World Wide Web (images) --- Two dimensional.
- Virtual worlds --- Three dimensional
- Humans familiar with 3-D world – social ways of exchanging information.
- Demands for privacy and security will escalate dramatically.

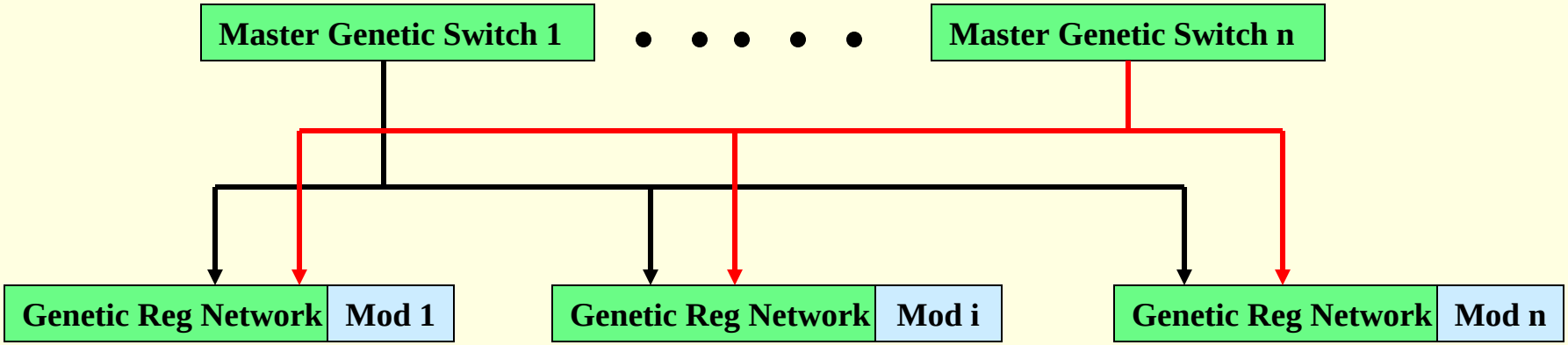
The World Selects the Cognitive Structure

- Evolutionary process is a knowledge gaining process of the world.
- The world selects the cognitive structures.
- Therefore, must “build-into” and organize the virtual world such that it will select the necessary structure for SmartData.

The Matrix of Virtual Evolution



Evolution by Modifying Design



Mod 1 = neuron (w,x,y,z…)
Where:
w = type of neuron;
x = number of neurons;
y = transfer function;
z = rules for LTP;

Mod i = sensor (x,y,x,z…)
Where:
w = type of sensor;
x = number of sensors;
y = location of sensors;
z = resolution;

Modules conserved;
GRN controls variables (,w,x,y,z);
GRN and MGS will undergo mutation

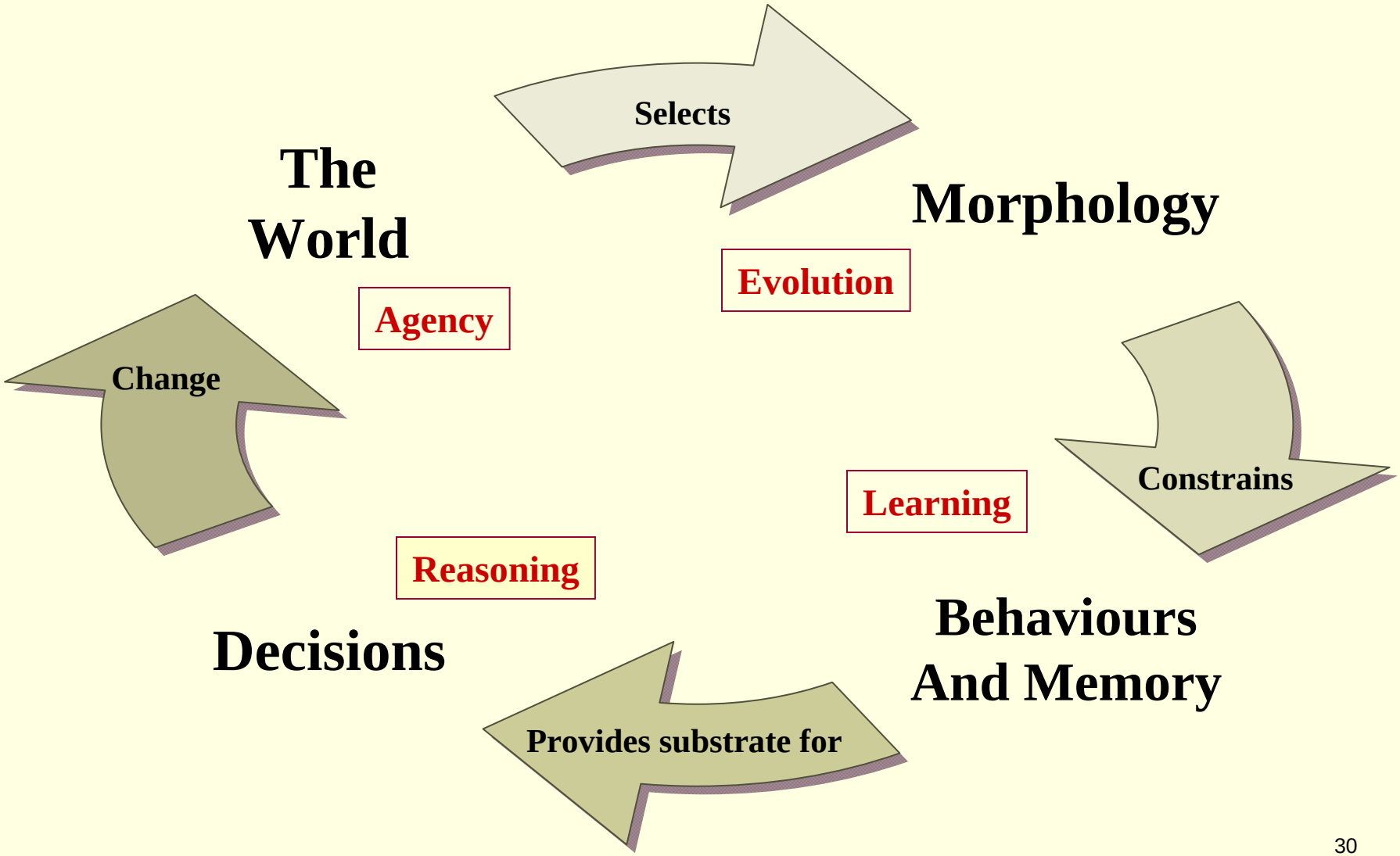
Nested Loops: Evolution, Learning, Reasoning

- Evolution: Constrained by the environment.
  - Time span of thousands to millions of years
  - Evolves physical and “basic” mental morphologies.
  - Mental morphologies are derivative from the physical.

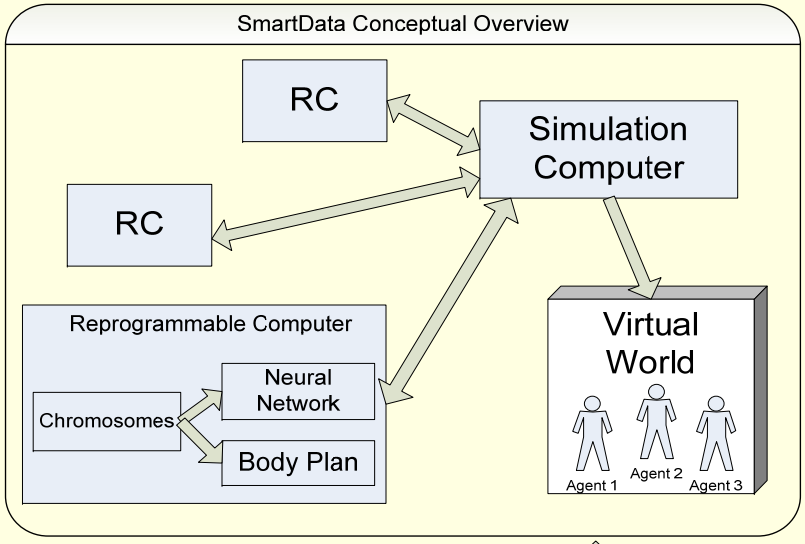
- Learning: Constrained by feelings which are a product of evolution.
  - Time span is the lifetime of the individual.
  - Includes behaviours (sequence of motor actions) and various types of memories.

- Reasoning: Constrained by learning and by feelings.
  - Time span is very short.
  - Encompasses a temporal horizon.
  - Utilizes stored information and rules of logical inference.

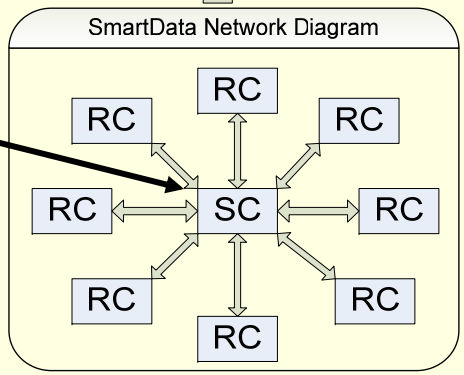
The Nested Loops of Artificial Agency



Open Source Development



RC = Reprogrammable Computer
SC = Simulation Computer



Virtual world as a hub on GENI's Internet II

Conclusions

- Current-day protections will become largely ineffective.
- Empowering virtual, cognitive agents to act on our behalf to protect the data entrusted to them.
- The ultimate embodiment of Privacy by Design.
- SmartData – an innovative approach to protecting privacy and security