

Establishing a Roadmap and Metrics for Conscious Machines Development

Raúl Arrabales, Agapito Ledezma and Araceli Sanchis

Computer Science Department

Carlos III University of Madrid

<http://Conscious-Robots.com/Raul>

IEEE International Conference on Cognitive Informatics 2009

PolyU, Hong Kong, 15th June 2009



➔ Contents

- Introduction
- Objectives
- *ConsScale*
- Applying the scale
- Example
- Conclusions

➔ Introduction (I)

- There are many approaches to consciousness.
- Intuitively it involves:
 - Perception.
 - Emotions.
 - Attention.
 - Self-recognition.
 - Theory of Mind (ToM).
 - ...

➔ Introduction (II)

- Existing cognitive architectures cover some of these aspects:
 - BDI Agents.
 - ACT-R.
 - SOAR.
 - ICARUS.
 - Haikonen's Cognitive Architecture.
 - LIDA.
 - CRONOS.

➔ Introduction (III)

- From a functional point of view a human-like mind requires a number of cognitive skills.
 - BUT, they need to be efficiently integrated.
- HOW?
 - Consciousness is (here) considered as the global process that provides the required synergy (“the grand function”).

➔ Objectives

- How to measure the level of consciousness?
 - Quantitatively and Qualitatively.
- To propose a developmental path for the design of conscious machines.
- Avoid the existing controversial issues about assessing consciousness.

➔ *ConsScale* Levels

- Level -1: *Disembodied*
- Level 0: *Isolated*
- Level 1: *Decontrolled*
- Level 2: *Reactive*
- Level 3: *Adaptive*
- Level 4: *Attentional*
- Level 5: *Executive*
- Level 6: *Emotional*
- Level 7: *Self-Conscious*
- Level 8: *Empathic*
- Level 9: *Social*
- Level 10: *Human-like*
- Level 11: *Super-Conscious*

➔ *ConsScale* Levels (Criteria)

- Architectural Components
- Cognitive Skills
- Observed behavior

➔ *ConsScale* (I)

- Level -1: *Disembodied*
 - Behavior: not a situated agent.
 - Phylogeny: amino acid

E

B

➔ *ConsScale* (II)

- Level 0: *Isolated*
 - Behavior: not a situated agent.
 - Phylogeny: isolated chromosome.

E

B

➔ *ConsScale* (III)

- Level 1: *Decontrolled*
 - Behavior: not a situated agent.
 - Phylogeny: dead bacteria.

E

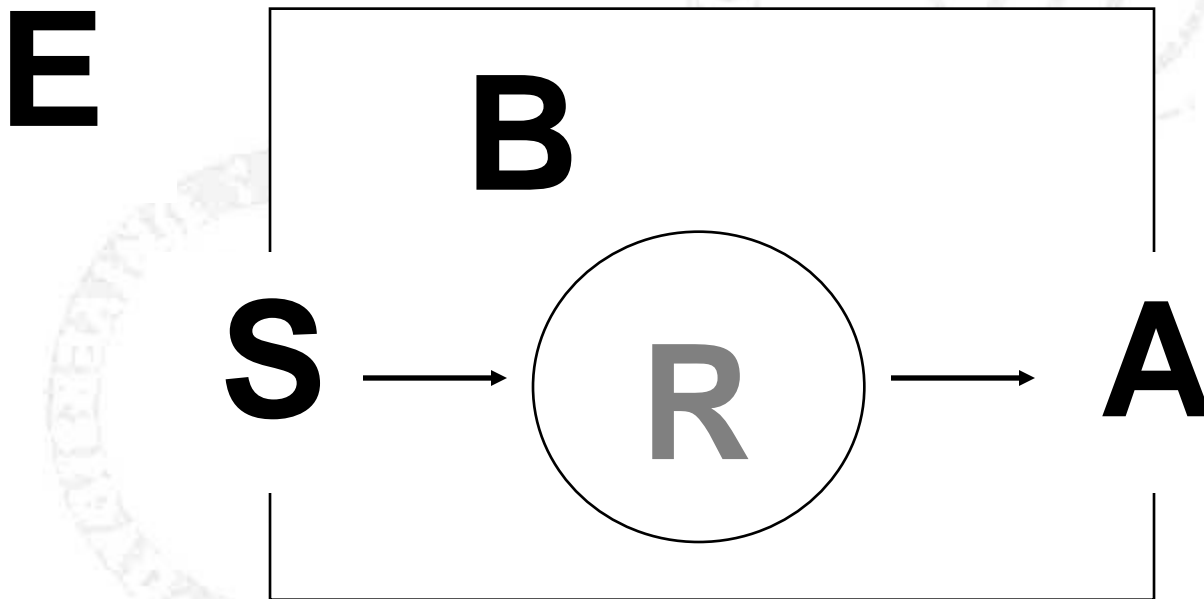
S

B

A

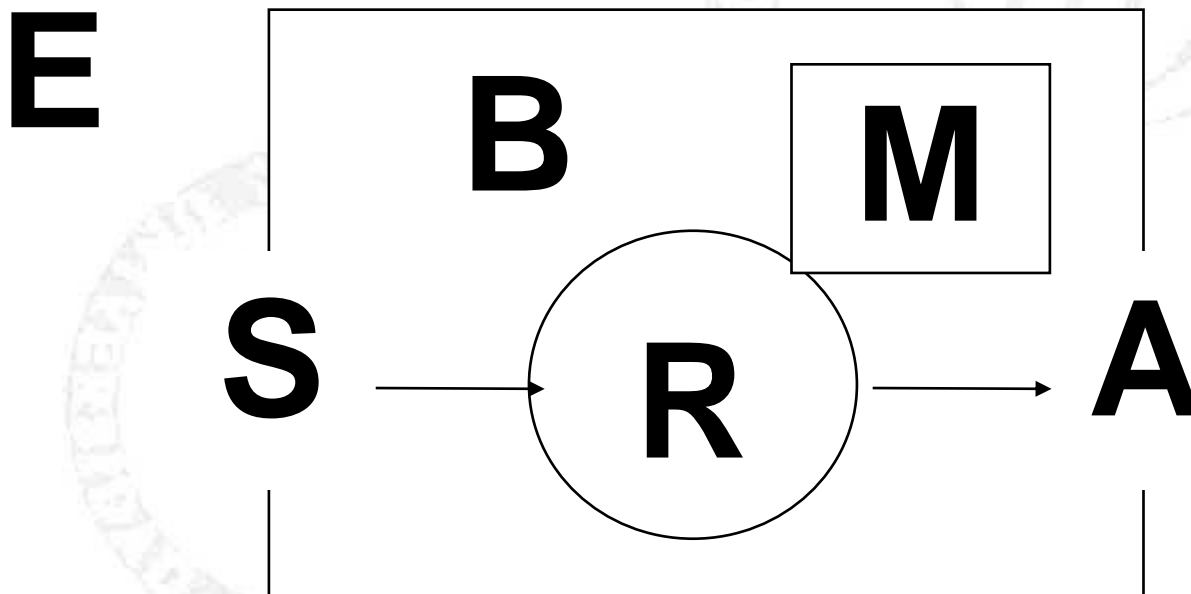
→ *ConsScale* (IV)

- Level 2: *Reactive*
 - Behavior: reflexes.
 - Phylogeny: virus.



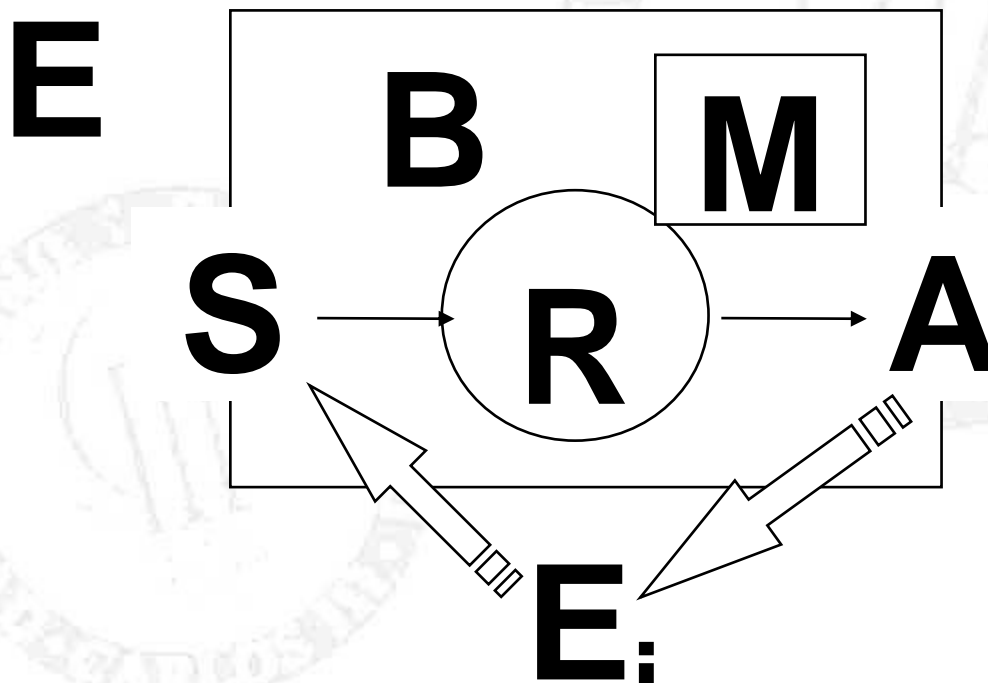
➔ *ConsScale (V)*

- Level 3: *Adaptive*
 - Behavior: basic ability to learn new reflexes.
 - Phylogeny: earthworm.



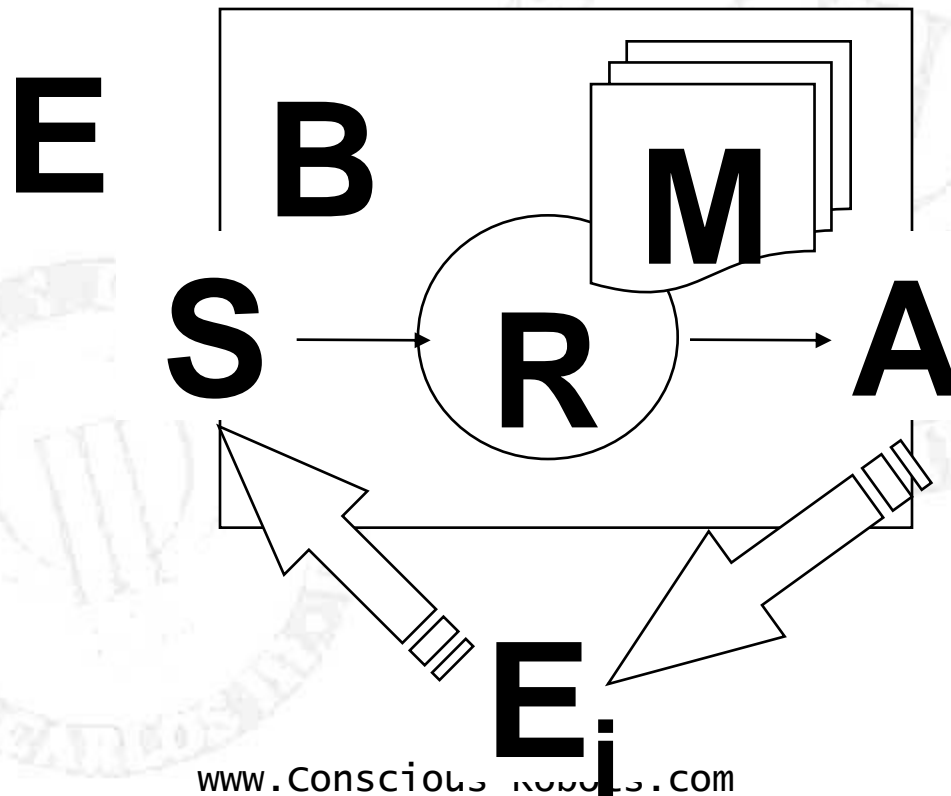
➔ *ConsScale* (VI)

- Level 4: *Attentional*
 - Behavior: attack and escape. Attention + emotion.
 - Phylogeny: fish.



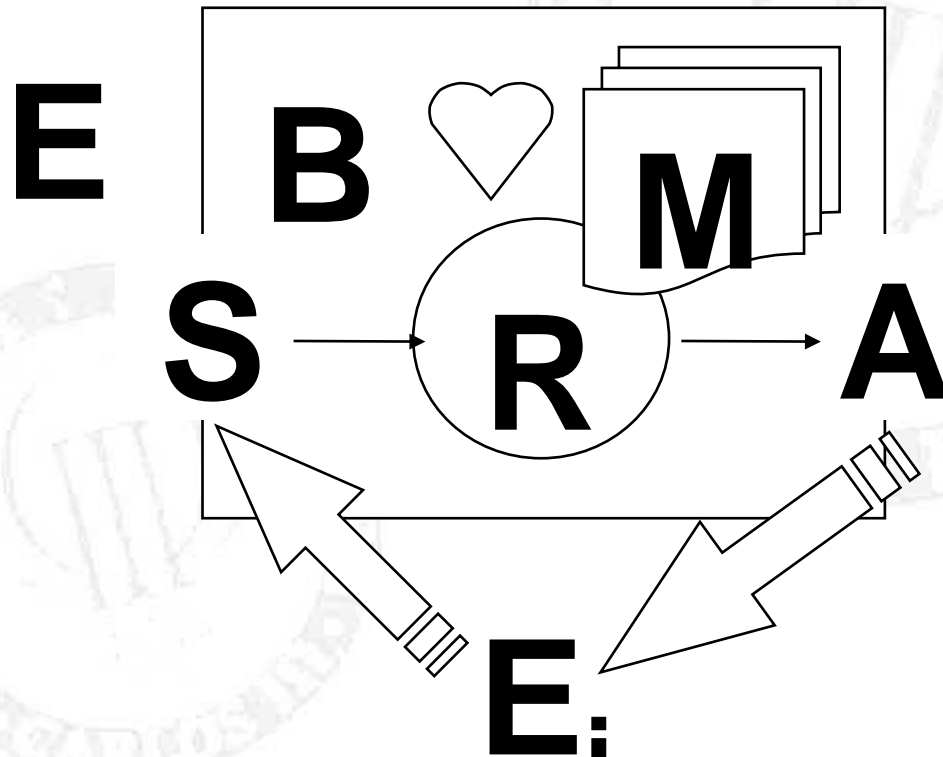
➔ *ConsScale* (VII)

- Level 5: *Executive*
 - Behavior: set shifting. Emotional learning.
 - Phylogeny: quadruped mammal.



➔ *ConsScale* (VIII)

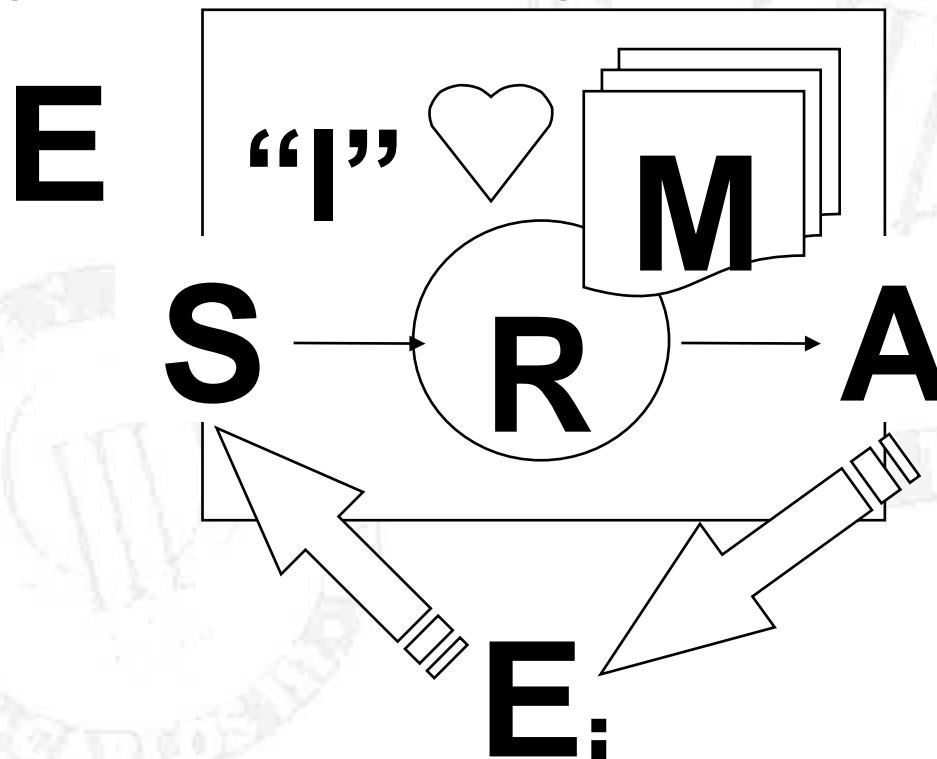
- Level 6: *Emotional*
 - Behavior: feelings influence behavior.
 - Phylogeny: monkey. ToM Stage 1: "I know".



➔ *ConsScale* (IX)

- Level 7: *Self-Conscious*

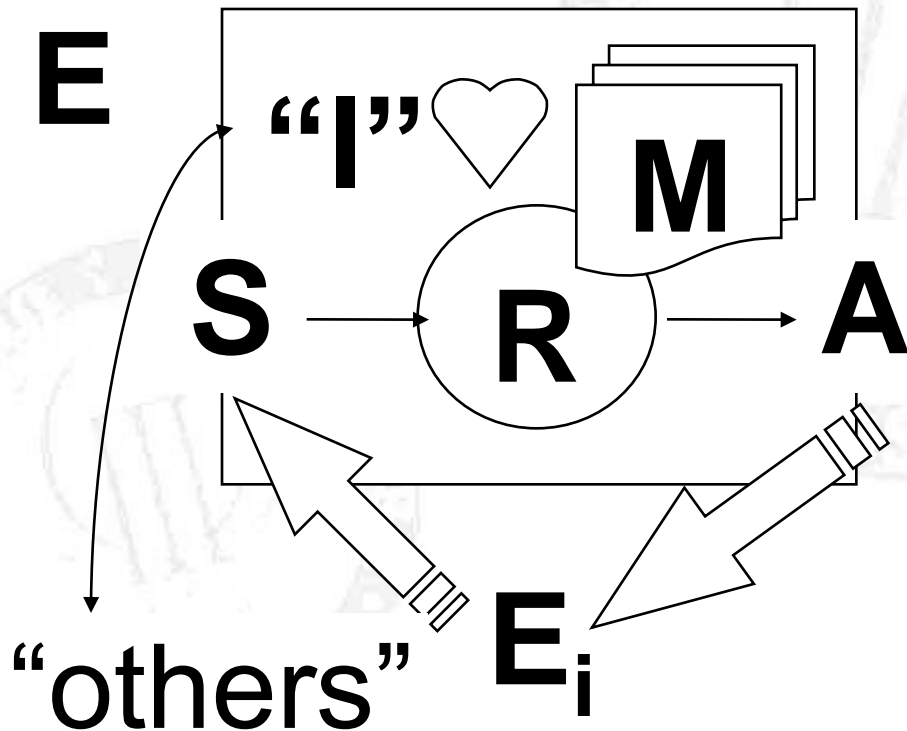
- Behavior: advanced planning. Usage of tools. Mirror test.
- Phylogeny: monkey. ToM Stage 2: "I know I know".



➔ *ConsScale (X)*

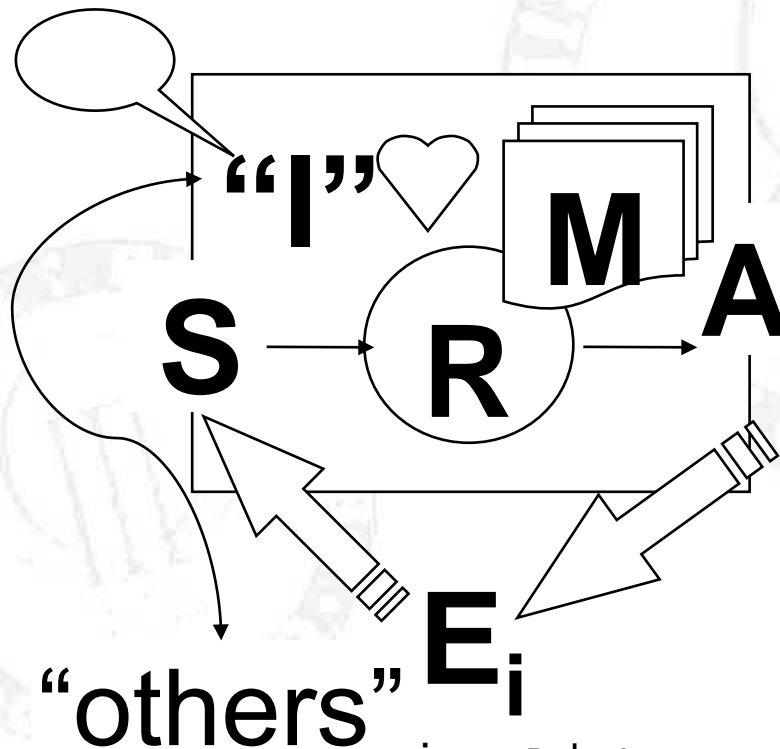
- Level 8: *Empathic*

- Behavior: making of tools. Social behavior.
- Phylogeny: chimpanzee. ToM Stage 3: "I know you know".



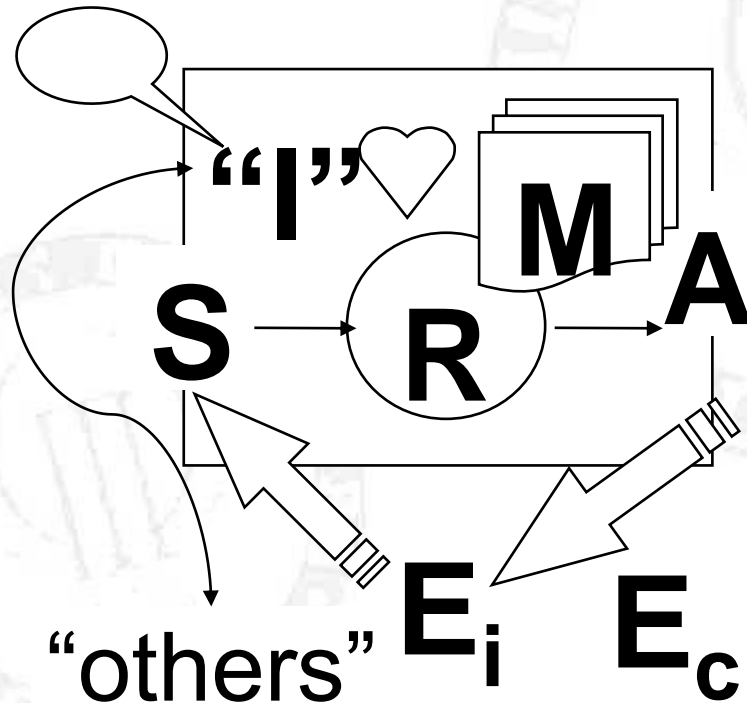
➔ *ConsScale* (XI)

- Level 9: *Social*
 - Behavior: linguistic capabilities. Ability for culture.
 - Phylogeny: human. ToM Stage 4: "I know you know I know".



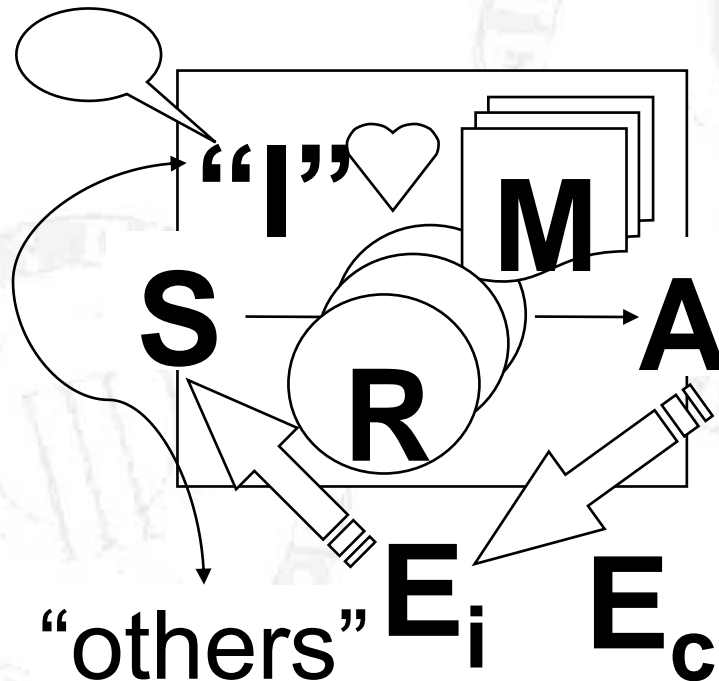
➔ *ConsScale* (XII)

- Level 10: *Human-like*
 - Behavior: accurate verbal report. Culture. Technology.
 - Phylogeny: human. Adapted environment.



➔ *ConsScale* (XIII)

- Level 11: *Super-Conscious*
 - Behavior: several streams of consciousness.
 - Phylogeny: n/a.



➔ Cognitive Skills ($CS_{i,j}$)

Level	Cognitive Skills ($CS_{i,j}$)
2	CS_{2,1} : Fixed reactive responses ("reflexes").
3	CS_{3,1} : Autonomous acquisition of new adaptive reactive responses. CS_{3,2} : Usage of proprioceptive sensing for embodied adaptive responses.
4	CS_{4,1} : Selection of relevant sensory information. CS_{4,2} : Selection of relevant motor information. CS_{4,3} : Selection of relevant memory information. CS_{4,4} : Evaluation (positive or negative) of selected objects or events. CS_{4,5} : Selection of what needs to be stored in memory. CS_{4,6} : Trial and error learning. Re-evaluation of selected objects or events. CS_{4,7} : Directed behavior toward specific targets like following or escape. CS_{4,8} : Evaluation of the performance in the achievement of a single goal. CS_{4,9} : Basic planning capability: calculation of next n sequential actions. CS_{4,10} : Depictive representations of percepts [17].
5	CS_{5,1} : Ability to move back and forth between multiple tasks. CS_{5,2} : Seeking of multiple goals. CS_{5,3} : Evaluation of the performance in the achievement of multiple goals. CS_{5,4} : Autonomous reinforcement learning (emotional learning). CS_{5,5} : Advanced planning capability considering all active goals.
6	CS_{6,1} : Self-status assessment (background emotions). CS_{6,2} : Background emotions cause effects in agent's body. CS_{6,3} : Representation of the effect of emotions in organism (feelings). CS_{6,4} : Ability to hold a precise and updated map of body schema. CS_{6,5} : Abstract learning (learned lessons generalization).

7	CS_{7,1} : Representation of the relation between self and perception. CS_{7,2} : Representation of the relation between self and action. CS_{7,3} : Representation of the relation between self and feelings. CS_{7,4} : Self-recognition capability. CS_{7,5} : Advance planning including the self as an actor in the plans. CS_{7,6} : Use of imaginational states in planning. CS_{7,7} : Learning of tool usage.
8	CS_{8,1} : Ability to model others as subjective selves. CS_{8,2} : Learning by imitation of a counterpart. CS_{8,3} : Ability to collaborate with others in the pursuit of a common goal. CS_{8,4} : Social planning (planning with socially aware plans). CS_{8,5} : Ability to make new tools.
9	CS_{9,1} : Ability to develop Machiavellian strategies like lying and cunning. CS_{9,2} : Social learning (learning of new Machiavellian strategies). CS_{9,3} : Advanced communication skills (accurate report of mental content). CS_{9,4} : Groups are able to develop a culture.
10	CS_{10,1} : Accurate verbal report. Advanced linguistic capabilities. CS_{10,2} : Ability to pass the Turing test. CS_{10,3} : Ability to modify and adapt the environment to agent's needs. CS_{10,4} : Groups are able to develop a civilization and advance culture and technology.
11	CS_{11,1} : Ability to manage several streams of consciousness.

➔ *ConsScale* as a Roadmap

- Consider consciousness as an integrator that puts a mind together.
- Consider synergy instead of isolated cognitive abilities.
- Evolutionary inspired developmental path.

→ CQS

- *ConsScale* Quantitative Score.

- L_i

- Particular score for level i .

- *CLS*

- Cumulative Level score.

- *CQS*

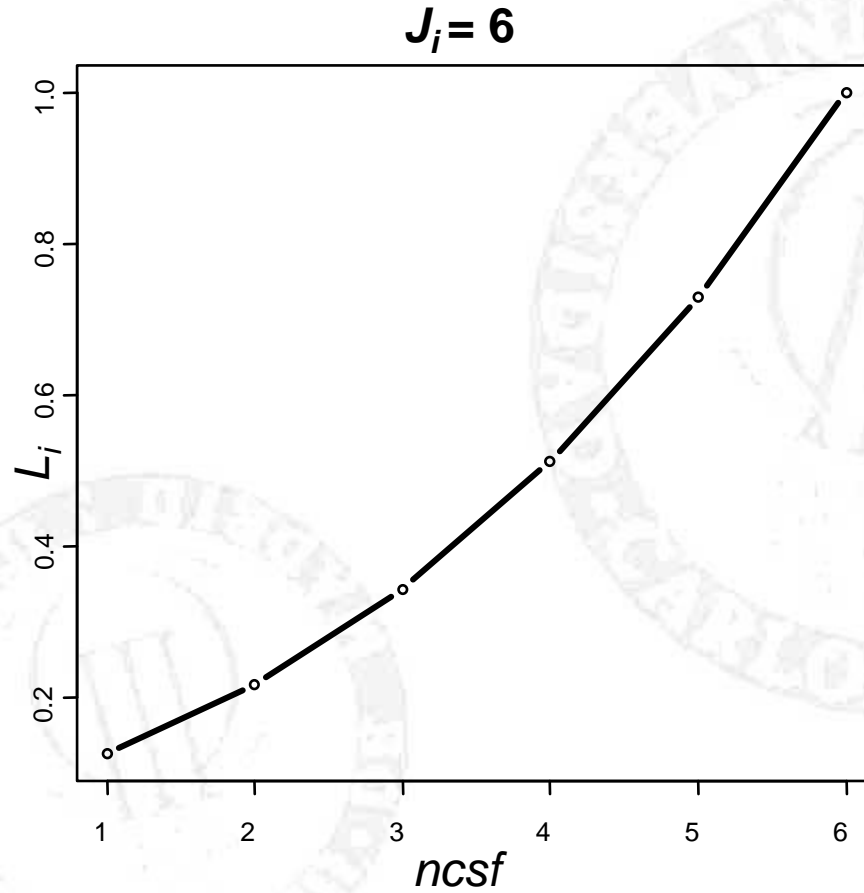
- ConsScale Quantitative score.

➔ L_i

$$L_i = \begin{cases} 0 & \text{if } ncsf \text{ is } 0 \\ \frac{ncsf + (J - J_i)}{10^3} & \text{otherwise} \end{cases}$$

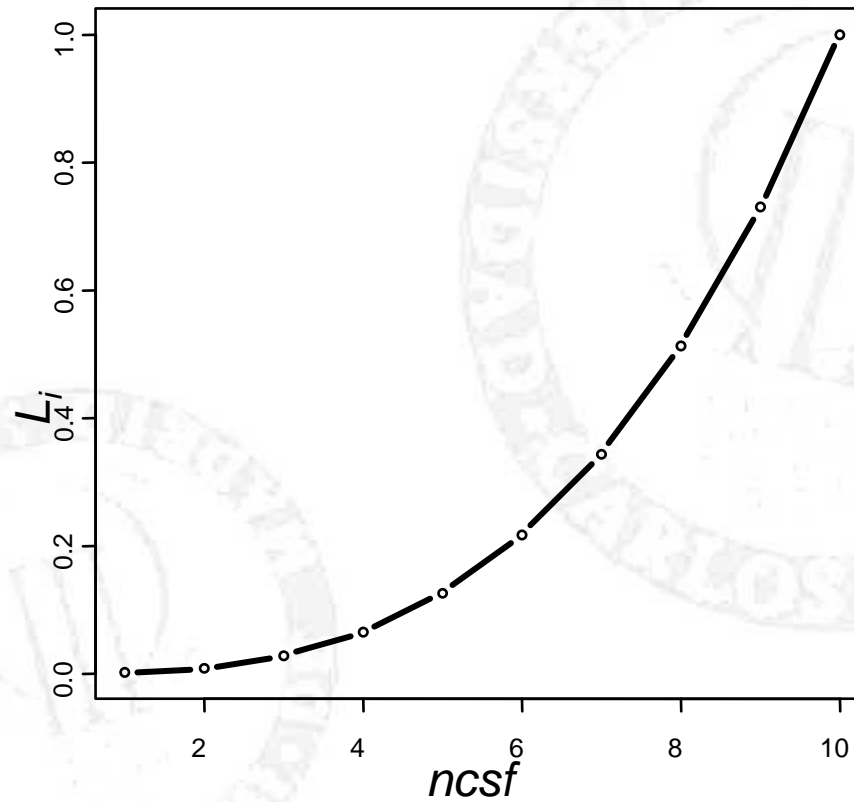
- $ncsf \rightarrow$ Number of cognitive skills fulfilled.
- $J \rightarrow$ Maximum number of CS .
- $J_i \rightarrow$ Total number of CS in level i .

➔ $L_i (J_i = 6)$



➔ $L_i (J_i = 10)$

$J_i = 10$

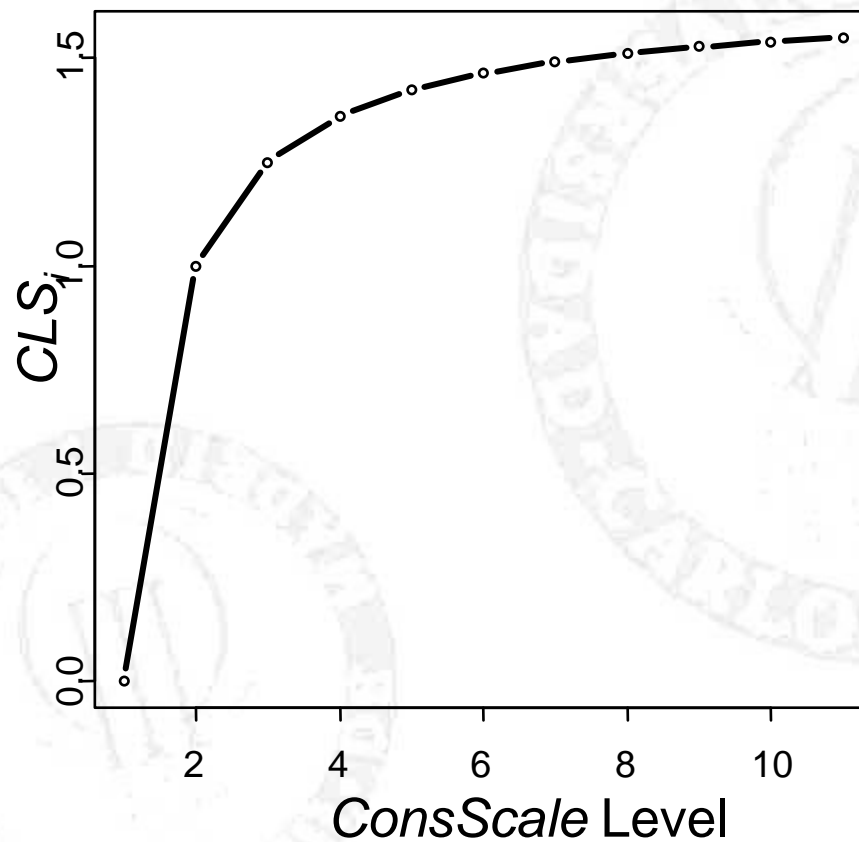


➔ *CLS*

$$CLS = \sum_{i=2}^{11} \left(\frac{L_i}{i-1} \right)^2$$

- $L_i \rightarrow$ Level i partial score.
- $i \rightarrow$ Level index.

➔ CLS possible values



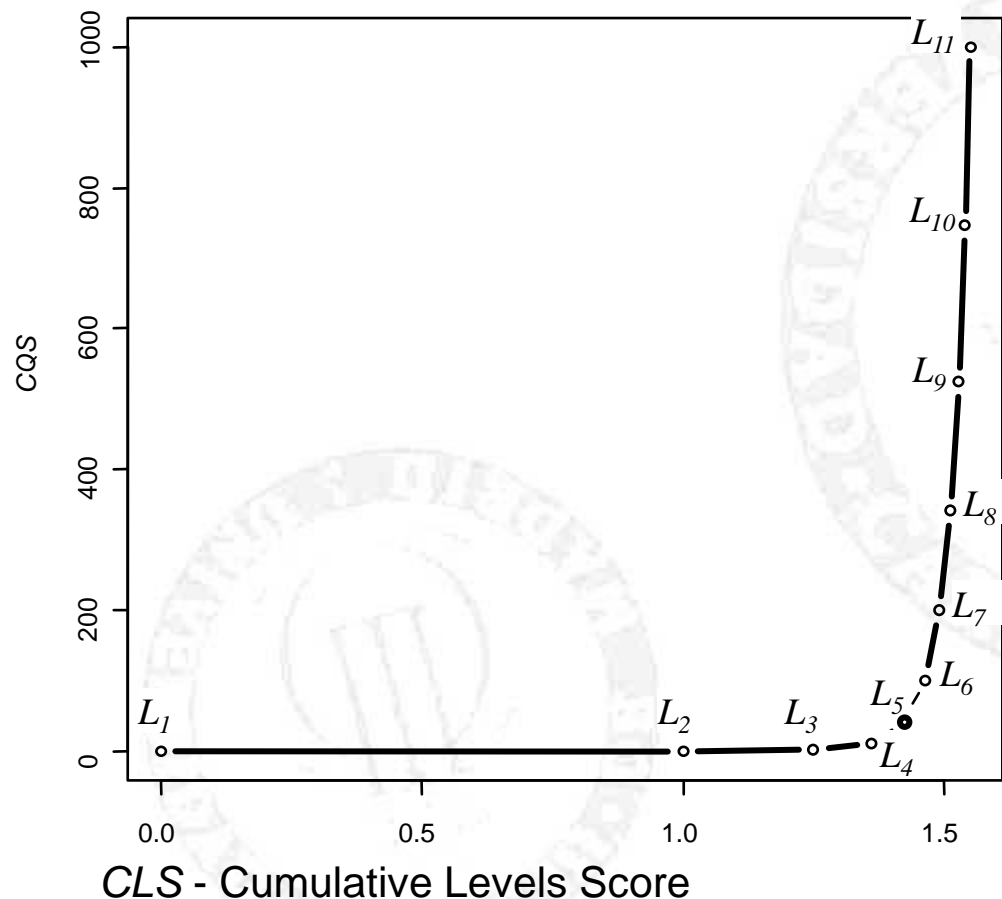
➔ *CQS* (0 to 1000)

$$CQS = \frac{e^{(CLS^5/K)} + a}{10}$$

- *CLS* → Cumulative level score.
- *K* → Constant value (~0'97).
- *a* → Constant value (-1).

➔ CQS possible values

CQS for levels 1 to 11.

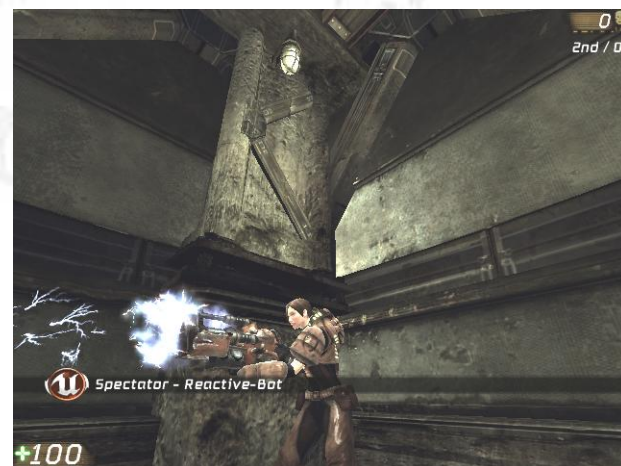
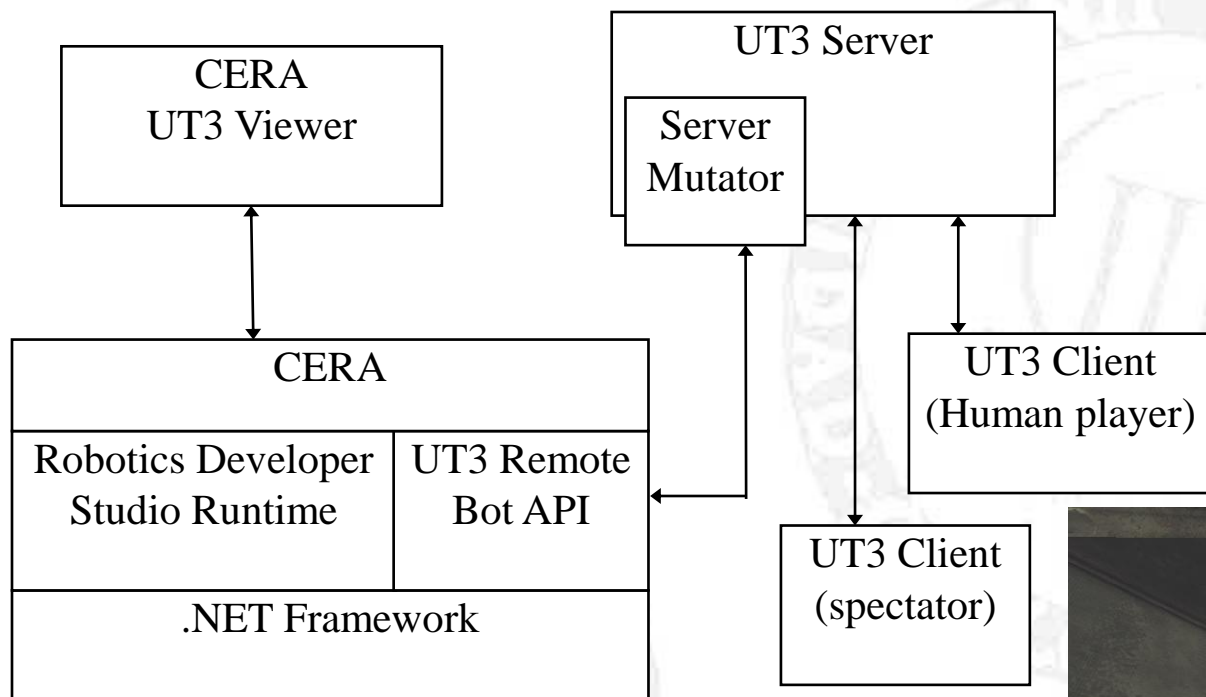


i	Description	CQS
1	<i>Decontrolled</i>	0.00
2	<i>Reactive</i>	0.18
3	<i>Adaptive</i>	2.21
4	<i>Attentional</i>	12.20
5	<i>Executive</i>	41.23
6	<i>Emotional</i>	101.08
7	<i>Self-Conscious</i>	200.02
8	<i>Empathic</i>	341.44
9	<i>Social</i>	524.54
10	<i>Human-Like</i>	745.73
11	<i>Super-Conscious</i>	1000.00

➔ CQS online calculator

<http://conscious-robots.com/consscale>

➔ Example (I)



➔ Example (II)

Agent	L_2	L_3	L_4	CLS	CQS
<i>Reactive-Bot</i>	1	0	0.000	1.000	0.18
<i>Adaptive-Bot</i>	1	1	0.000	1.250	2.22
<i>Attentional-Bot</i>	1	1	0.216	1.255	2.38

➔ Conclusions

- Cognitive approach to artificial consciousness metrics.
- Framework for evaluation.
- The general scale needs to be instantiated.

➔ Thank you

