Projective simulation for autonomous learning agents

Hans J. Briegel

IQOQI, Austrian Academy of Sciences, and
Institute for Theoretical Physics, University of Innsbruck

Collaborations:

Gemma De als Cuevas (→ MPQ Garching)
D Mautner, A Makmal, M Tiersch, D Manzano

Support: FWF (SFB FoQuS), EU (QICS)

Quantum Physics & Biology

- Biology & Information processing

  → DNA, genetic code; protein synthesis;
  → Cell machinery & regulation; error correction; …
  → Brain, learning & behavior

- Does quantum coherence/entanglement play any role?

  → Trivial versus non-trivial quantum effects

  → Photosynthesis; energy transfer efficiency in light harvesting

  → Magneto-reception; radical-pair mechanism; spin chemistry
  Avian magnetic compass

- Can genuine Q-states (entanglement) be maintained at **T = 300K** & **noise**?

  → Non-equilibrium bio-molecular machine: „Entanglement generator“

PRE **82, 021921** (2010)

Quantum Physics & Biology

Artificial intelligence

- Biology & Information processing

  → DNA, genetic code; protein synthesis;
  → Cell machinery & regulation; error correction; …
  → Brain, learning & behavior

- Does quantum coherence/entanglement play any role?

  → Trivial versus non-trivial quantum effects

  → Photosynthesis; energy transfer efficiency in light harvesting

  → Magneto-reception; radical-pair mechanism; spin chemistry
  Avian magnetic compass

- Can genuine Q-states (entanglement) be maintained at **T = 300K** & **noise**?

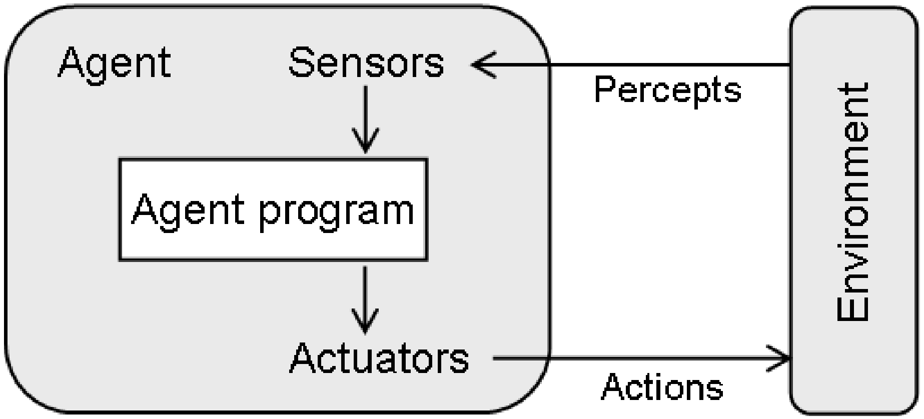
  → Non-equilibrium bio-molecular machine: „Entanglement generator“

PRE **82, 021921** (2010)

Plan of talk

- General & introductory remarks
  - Quantum physics & computation
  - Quantum physics & biology
- Artificial agents
  - Agents *vs*. computers
  - Quantum agents
  - Projective simulation for learning
  - Quantum projective simulation
- Conclusion & Outlook
  - Creative machines
- Open problems & connections to workshop theme

Artificial agents



Emphasis on “embodied aspects“
Embodied approaches in AI & robotics
(Braitenberg, Brooks, Pfeiffer,…)

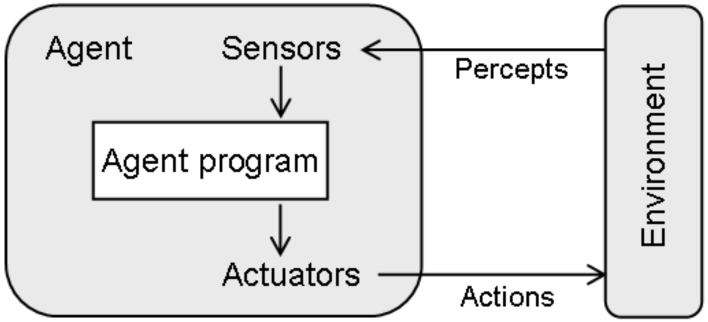
Agent program:
- reflex-type agent
- model-based agent
- utility-based..
- knowledge-based..
- learning agent
- …
- open/closed loop

Various applications:
- traffic control
- remote space
- internet
- robots, nanobots
- models for biological agents
- …

*from: Russel & Norvig, *Artificial Intelligence*, *A modern approach*, Prentice Hall, 2010.

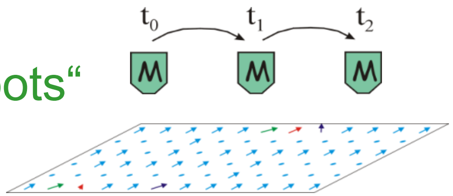
Quantum agents

ASIMO



•Which part should be quantum?

Future “lab robots“



Two examples:

1) Environment = Quantum state $|\psi\rangle$
Actions = Measurements/unitaries on $|\psi\rangle$
Percepts = Measurement results

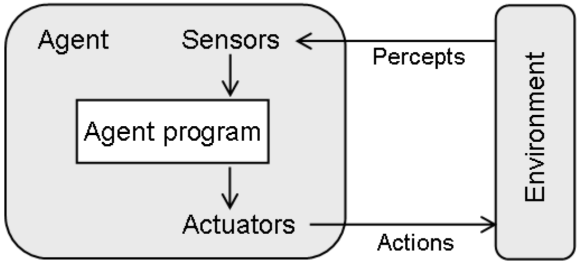
- e.g. One-way quantum computer *Raussendorf & HJB (2001)*
- Quantum error correction
- Quantum-state preparation
- …

2) Environment = Classical
Program = Quantum
Sensors & Actuators = Quantum-classical interfaces

Quantum-enhanced agents/robots
Impact on behavior? (external view)

P.S.: Birds, Drosophila…?

Quantum agents



Problems:

- No ***universal model!***
- What to put in the box? What to quantize?
  - -- Turing machine?
  - -- Neural network?
  - -- … ?
- No well-established theory / integrated view of agents

Self-simulating machine

This talk:

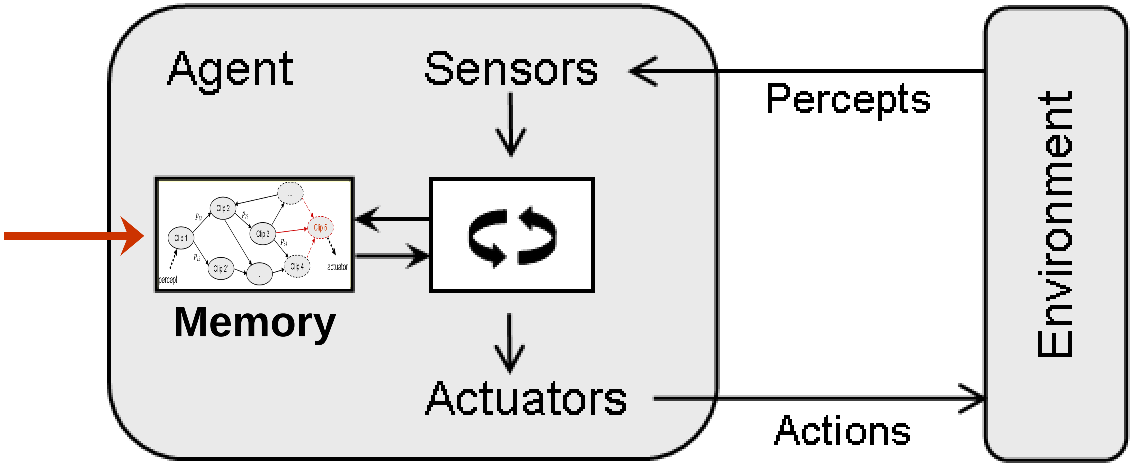
- Proposal of a learning-type agent based on "*projective simulation*".
- Exploit *simulation* -- as a fundamental *physical* concept -- in context of *learning*.
- Our proposal provides a natural basis for quantization:

→ quantum projective simulation.

Projective simulation for artificial agents

HJB & Gemma De las Cuevas, *Scientific Reports* **2**, 400 (2012)

Julian Mautner, Adi Makmal, *et al.* preprint (2012)



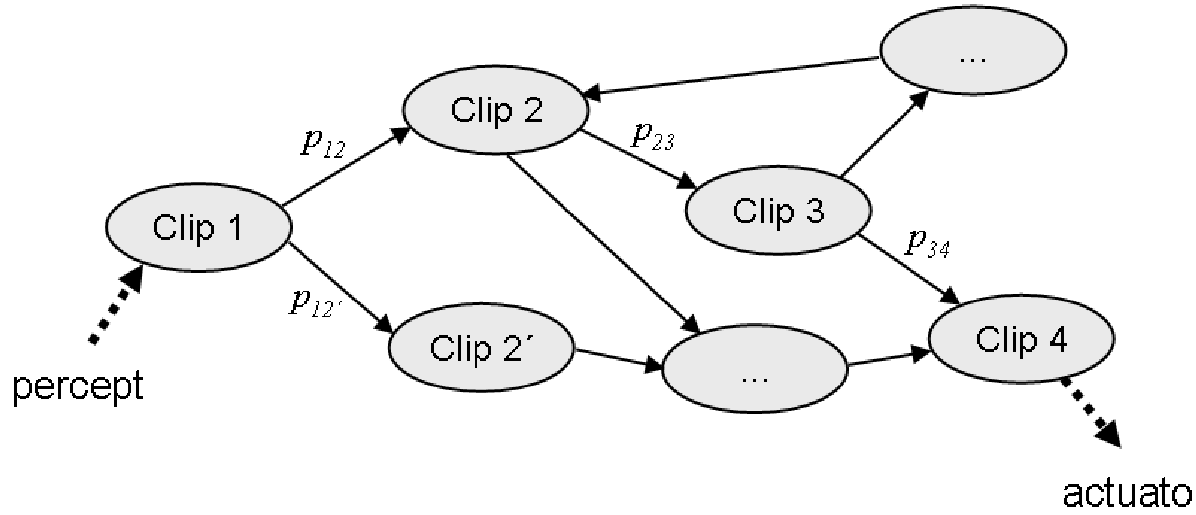
PART 1:
Classical formulation

Learning-type agent:

Operation principle based on
- *projective simulation (PS)*
- *episodic & compositional memory (ECM)*

Episodic & compositional memory (ECM)

Episodic memory:

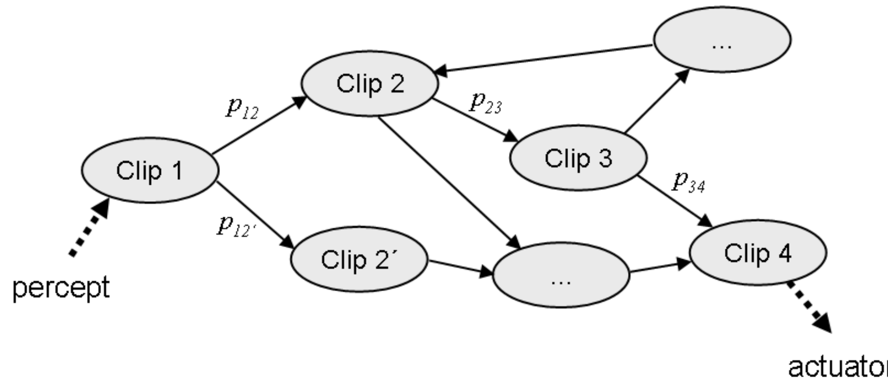


- Agent's past experience stored in form of episodes (short sequences of percepts & actions)
- Basic units: ***Clips*** ~ episodic fragments

  ~ patches of "space-time" memory

  (agent "space" = percept/actuator space)
- Call of memory = ***random walk*** through network of clips

Projective simulation



Simplest version of scheme:

(1) ***Perceptual input*** excites some initial (“percept“) clip. [coupling in]

(2) This triggers a ***random walk*** through clip space.

→ corresponds to patchwork-like sequences of virtual experience (“simulation“)

(3) ***Action*** is induced by screening clips for specific features. [coupling out]

→ presence of feature (detection above certain intensity level) triggers motor action.

(4) ***Rewarded action*** leads to

(a) update of transition probabilites and

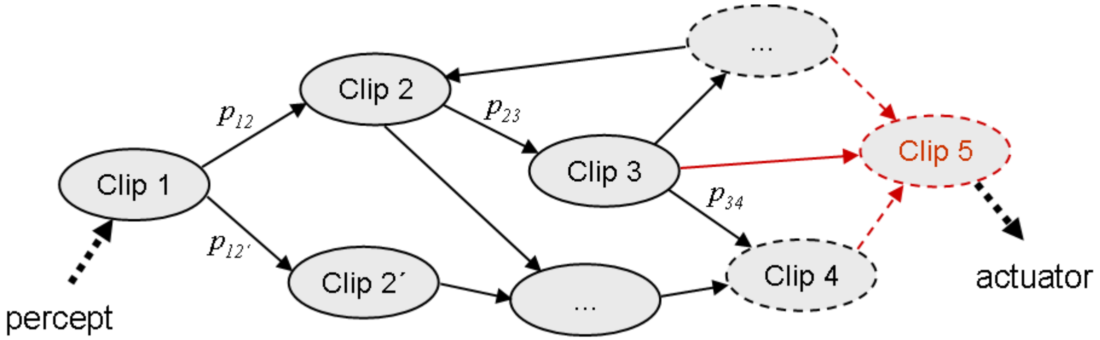
(b) update of ***emotion tags*** associated to transitions

P.S.: Other concepts in AI that are (loosely) related to clips:

- Classifier systems in genetic programming
- Model-based (dyna) planning in reinforcement learning

Refinement of scheme

Compositional memory:



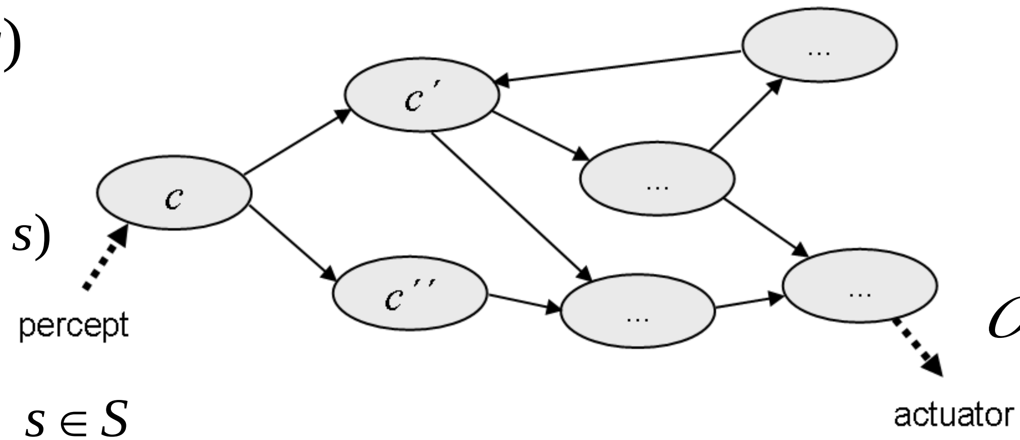
- During the simulation, *fictitious* clips may be *created* that were never actually perceived before, e.g. by random variation or merging.
- Fictitious clips will thereby influence *factual* action of the agent!
- No complicated “computation“.

→ *arXiv:1104.3787*
Sci. Rep. **2**, (2012)

Mathematical description

Internal description: $p^{(t)}(c'|c)$

$\mathcal{I}(c|s)$



$\mathcal{O}(a|c)$

$s \in S$

$S = S_1 \times \ldots \times S_N$

$a \in A$

$A = A_1 \times \ldots \times A_M$

External description: $P^{(t)}(a|s)$

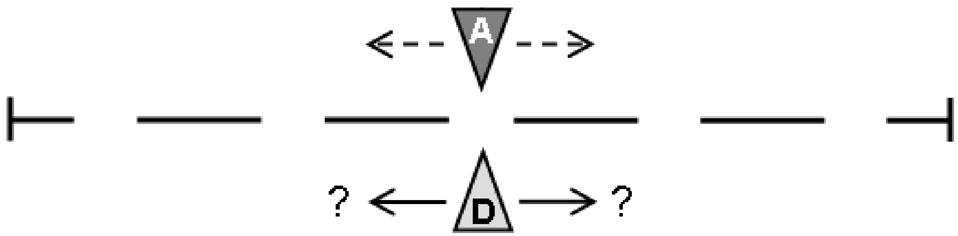
Clips: $c = (c_1, c_2, \ldots, c_L) \in C$

$c_l \in \mu(S) \cup \mu(A)$

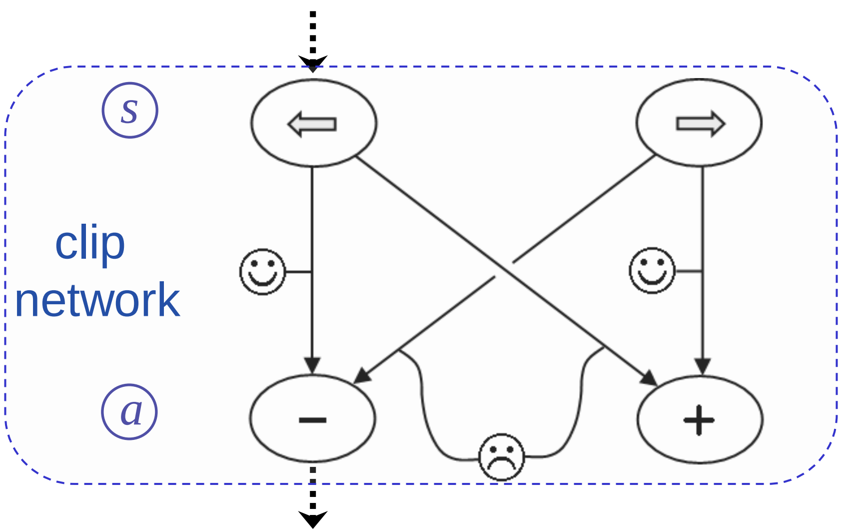
Simplest case: $c = \mu(s) = (s)$

(clip length=1) $c = \mu(a) = (a)$

Illustration: Invasion game



s sensor input



a motor output

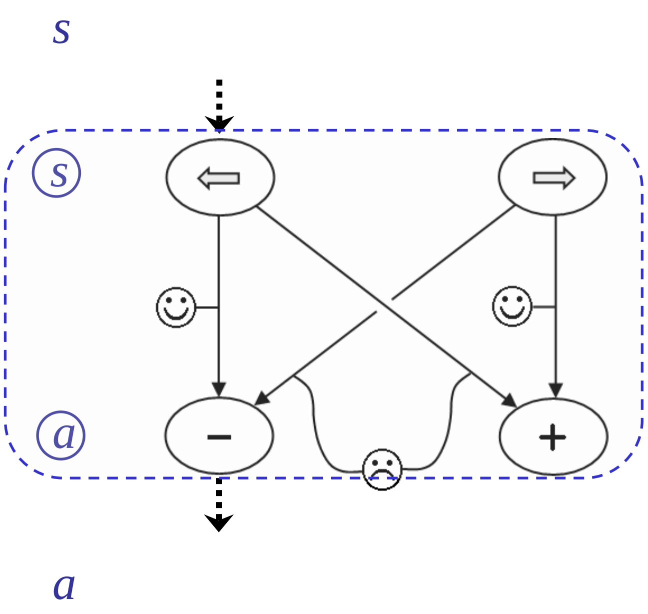
$$p^{(n)}(a \mid s) = \frac{h^{(n)}(s,a)}{h^{(n)}(s)}$$

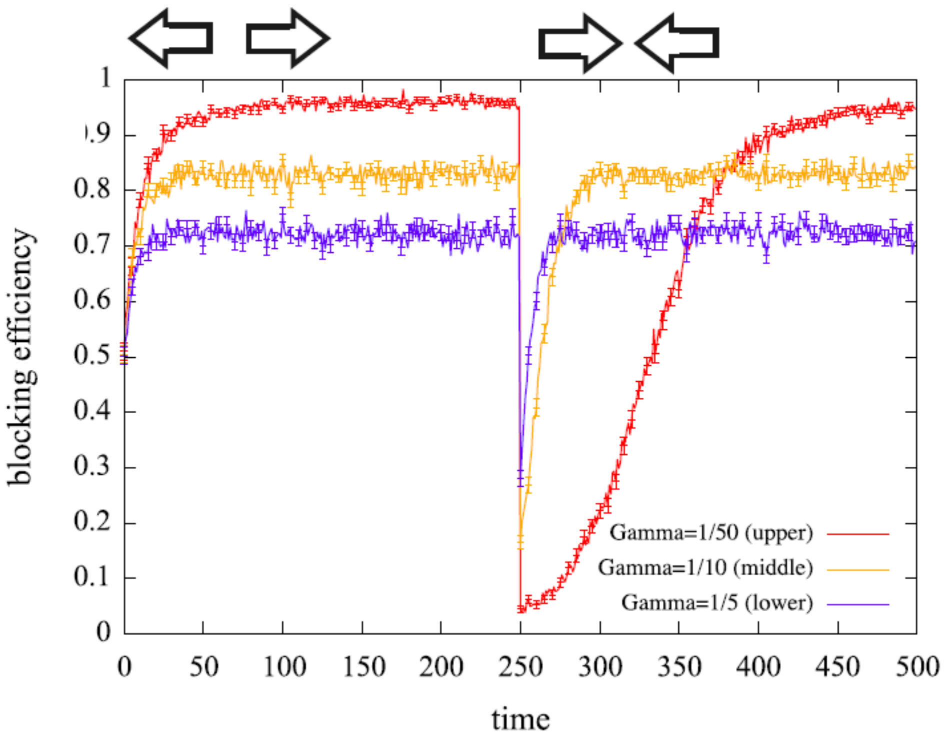
$$h^{(n+1)}(s,a) - h^{(n)}(s,a) = -\gamma\left[h^{(n)}(s,a) - 1\right] + \delta(s,s^{(n)})\delta(a,a^{(n)})\Lambda(s^{(n)},a^{(n)})$$

Λ reward function

γ dissipation rate

Learning & forgetting

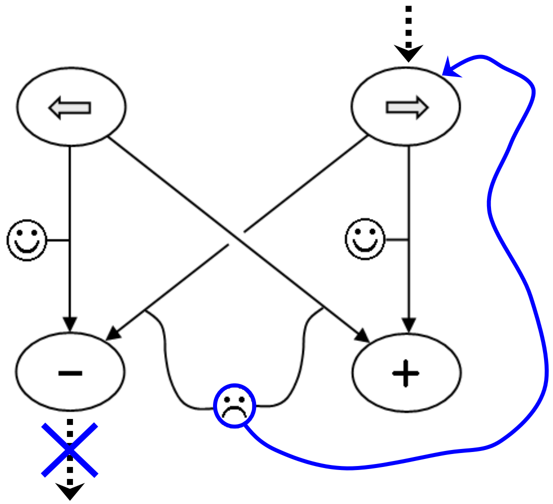


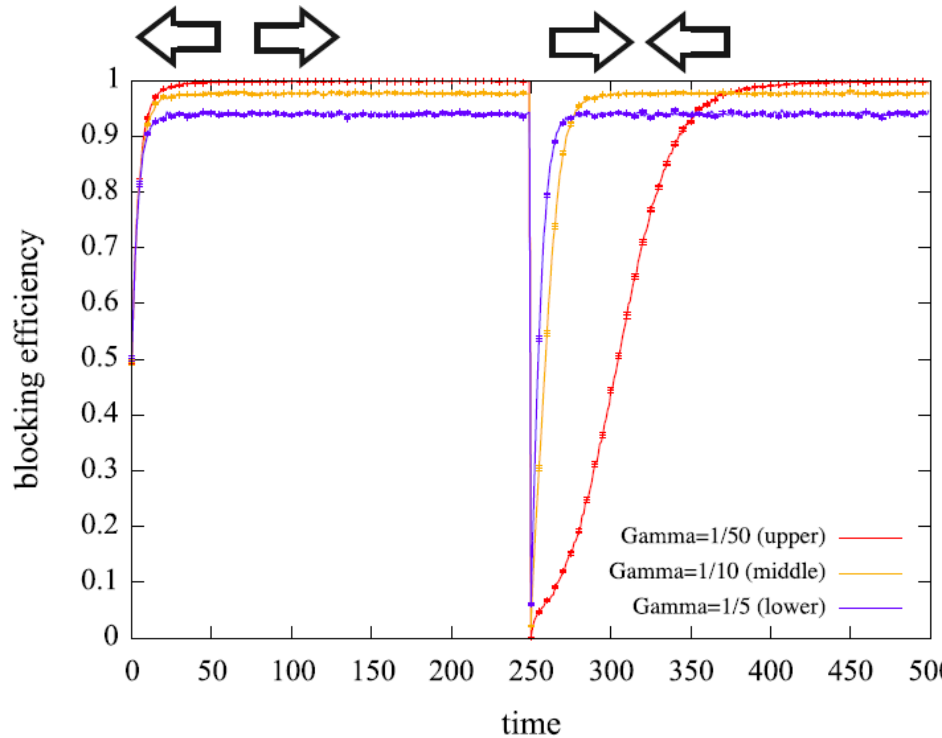


$$P^{(n)}(a_s^* \mid s) = p^{(n)}(a_s^* \mid s)$$ success probability

$$r^{(n)} = \sum_{s \in S} P^{(n)}(s)\, P^{(n)}\left(a_s^* \mid s\right)$$ blocking efficiency (average reward)

Simulation with reflection





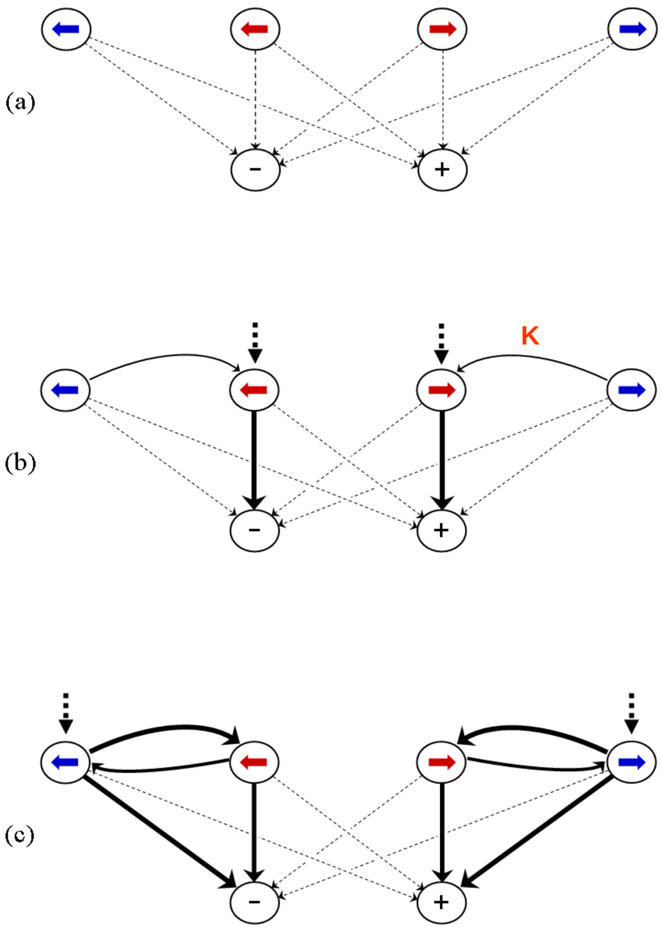
$$P^{(n)}(a_s^* \mid s) = 1 - \left(1 - p^{(n)}(a_s^* \mid s)\right)^R$$

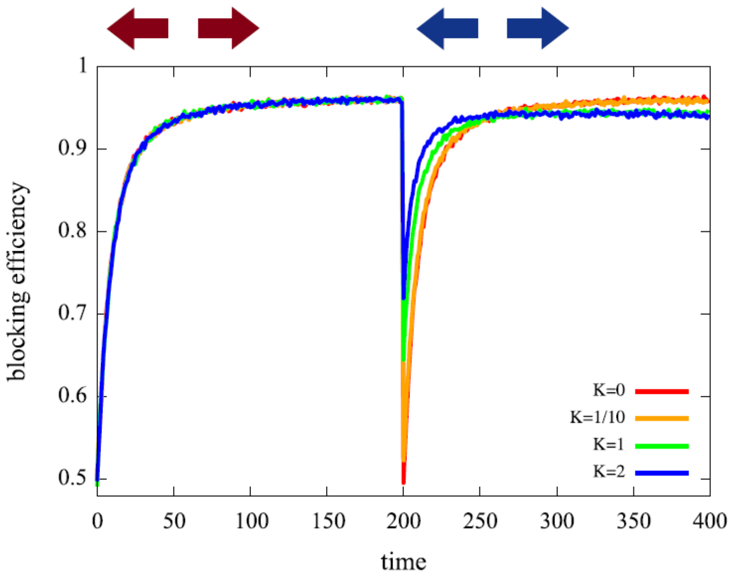
success probability, R = max # of reflections

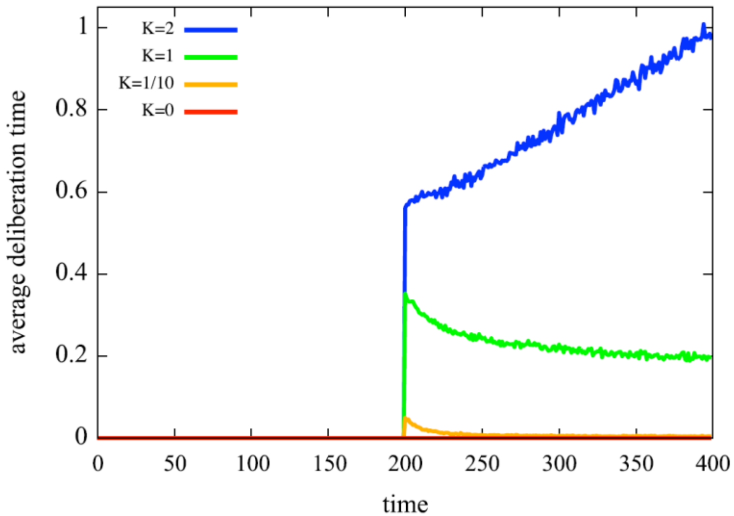
$$r^{(n)} = \sum_{s \in S} P^{(n)}(s) P^{(n)}(a_s^* \mid s)$$

blocking efficiency (average reward)

Simulation with association

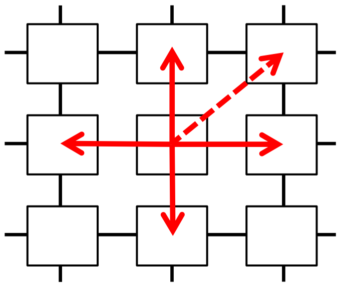


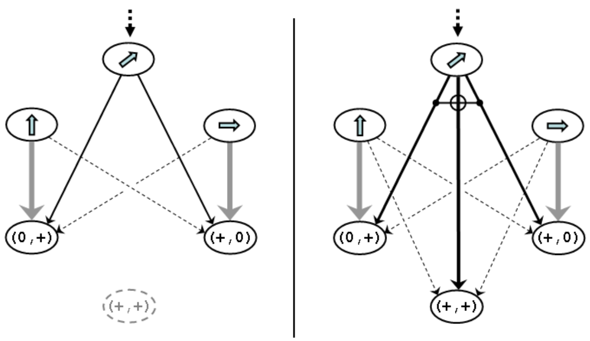


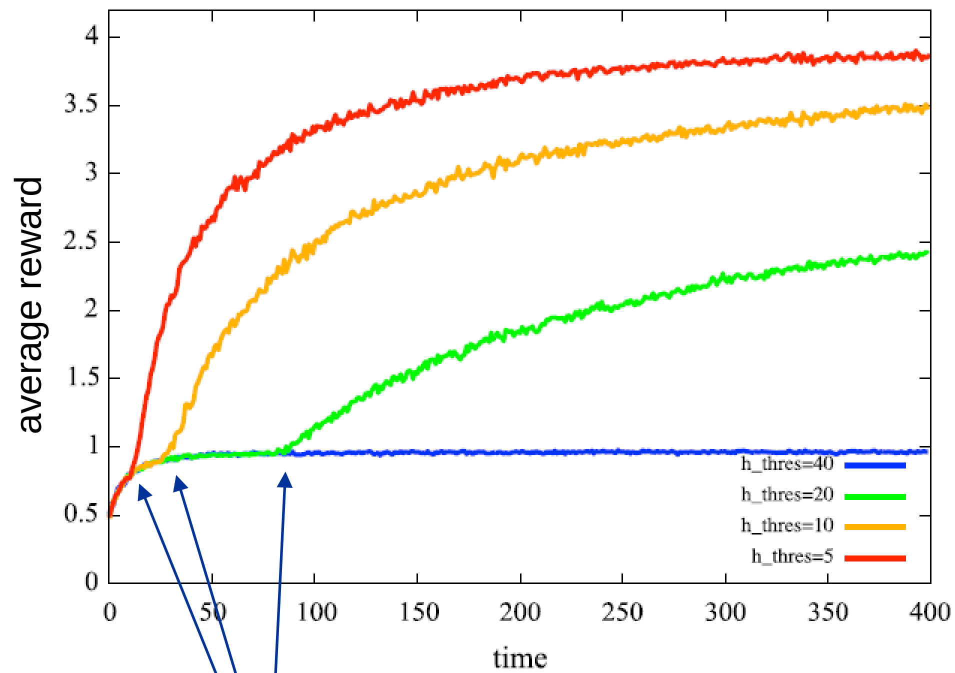


Simulation with composition (creative memory)

2D Invasion game





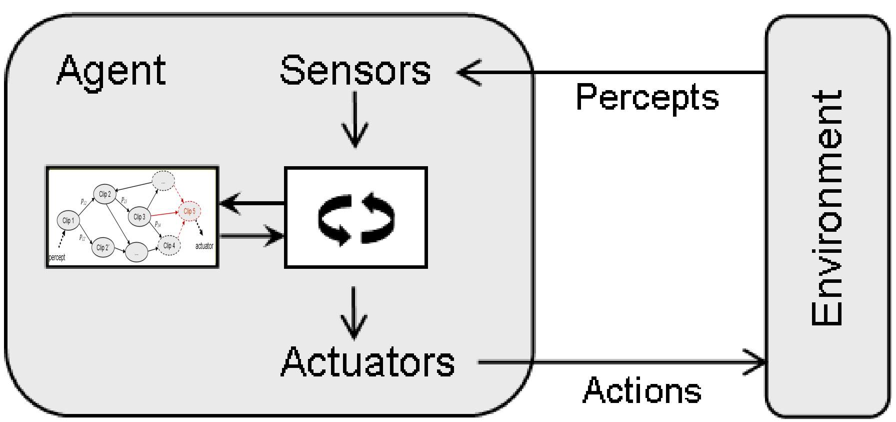


Different thresholds for clip composition

Complex learning: Agent „discovers“, through simulation, new (composite) motor actions that were previously inert.

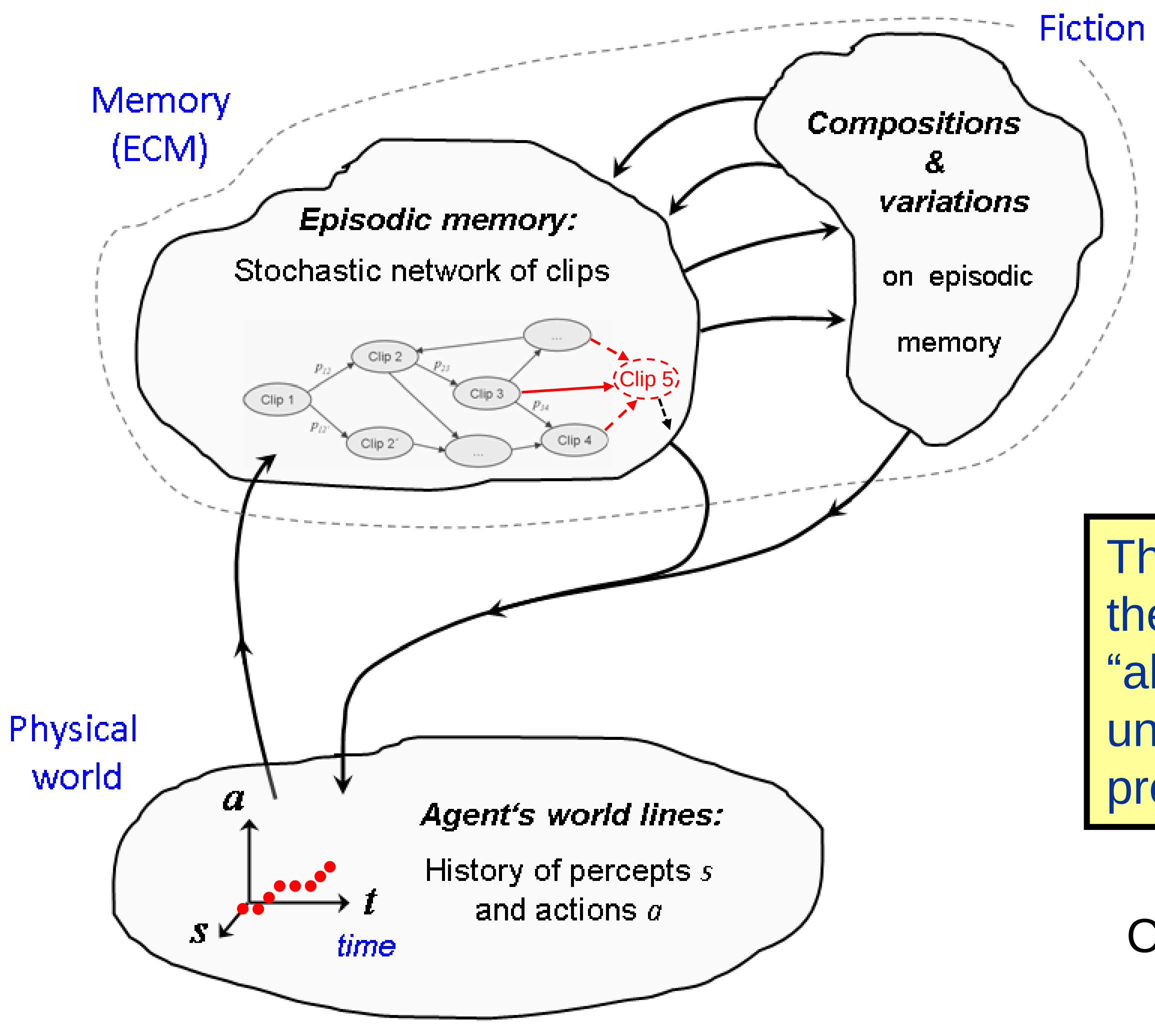
Summary projective simulation

- Simulation provides platform to *replay and vary* previous experience, before concrete action is taken!
- Random processes & rules of clip composition introduce room for variation around *established patterns of behavior*:
- It is the *agent itself* that *creates options* by internal random processes that are properly utilized via simulation.



Literature:
→ *Sci. Rep.* **2**, 400 (2012)

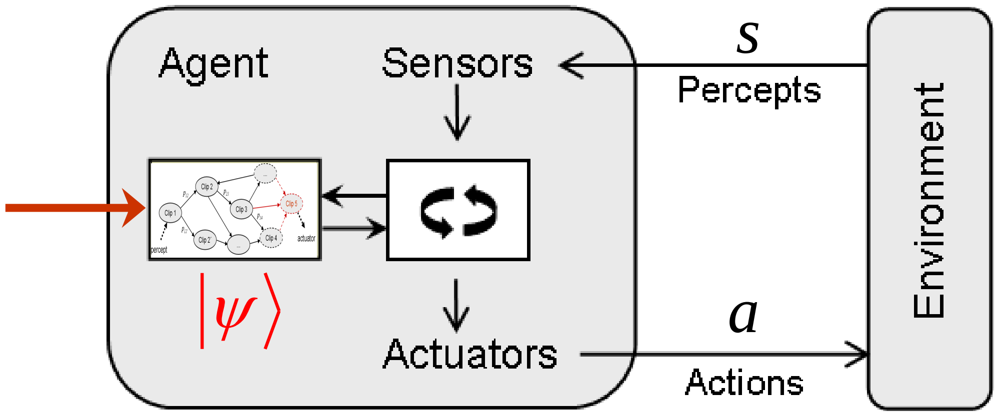
Projective structure of agent's behavior



Through projective simulation, the agent is permanently "ahead of itself" and acts under the influence of its own projections.

Creative machines:
→ *Sci. Rep.* **2**, 522 (2012)

Quantum projective simulation



$s \in S$ classical percepts

$a \in A$ classical actions

$$c \in C \rightarrow |c\rangle \in H_C$$

quantum memory (clip network)

Classical random walk → quantum walk

PART 2:
Quantum formulation

Idea:
Agent can explore its episodic memory in superposition with a potentially huge gain in efficiency. → Could thus ***react*** to a given situation much ***faster.***

Quantum projective simulation

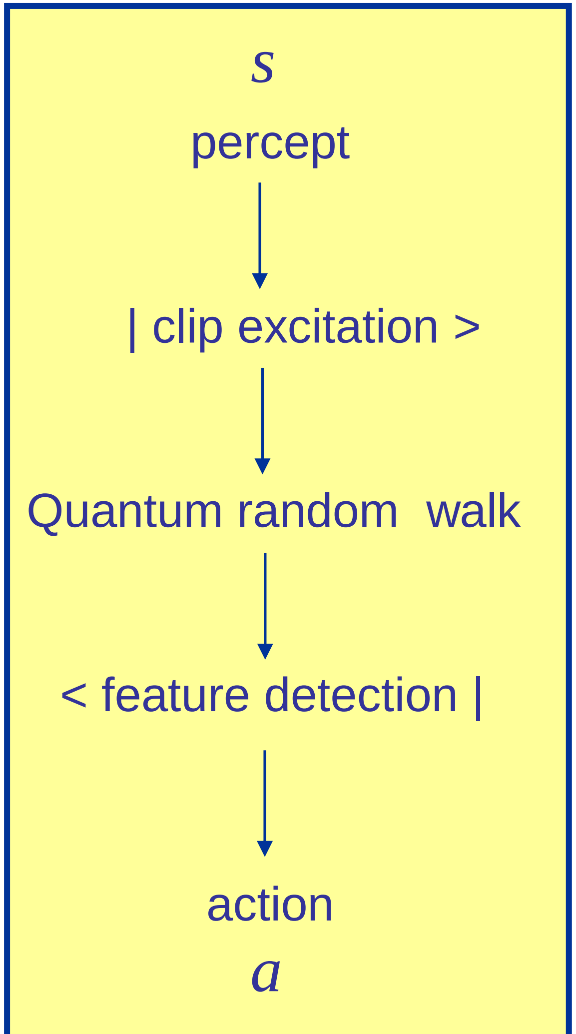
now the status of a basis state in the memory system. The random walk in clip space, which is an essential ingredient in our model, now becomes a *quantum walk* in the associated Hilbert space of the (quantum) memory, with the replacements

$$p(c'|c) \to |\langle c'|c\rangle|^2 \tag{16}$$

for elementary transitions between clips, and

$$p(c''|c) \to \left|\sum_{c'} \langle c''|c'\rangle \langle|c'|c\rangle\right|^2 . \tag{17}$$

for composite transitions. Here the scalar product $\langle c'|c\rangle$ defines the *probability amplitude* for the transition $\textcircled{c}\to\textcircled{c'}$, and the modulus squared in the expression for the composite transition gives rise to *quantum interference*, which is one of the basic features of quantum mechanics. Quantum interference is in particular exploited in fast algorithms for quantum search[39] and quantum walks on graphs[40].



Modelling of the quantum walk

- Hamiltonian representation (e.g. Hines & Stamp, 2006)

$$H = \sum_{\{j,k\}\in E} \lambda_{jk} \left(\hat{c}_k^\dagger \hat{c}_j + \hat{c}_k \hat{c}_j^\dagger \right) + \sum_{j\in V} \varepsilon_j \hat{c}_j^\dagger \hat{c}_j$$

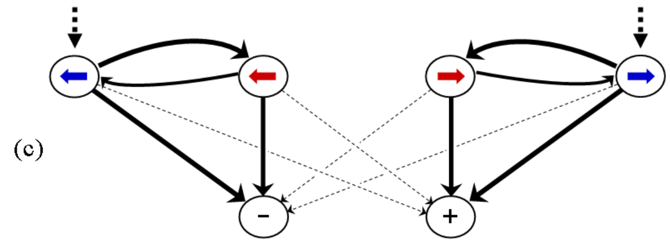
… describing coherent transitions between different clips c_j and c_k in the network,

- Clip themselves correspond to certain excitations of the quantum memory, described by excitation $\hat{c}^\dagger$ and de-excitation operators $\hat{c}$

$$|c_j\rangle = \hat{c}_j^\dagger |\mathtt{vac}\rangle$$

- Single excitation subspace: $$H = \sum_{\{j,k\}\in E} \lambda_{jk} \left(|c_k\rangle\langle c_j| + |c_j\rangle\langle c_k| \right) + \ldots$$

Uni-directional „quantum jumps"



- can be described by Lindblad-type terms

$$L\rho = \sum_{\{j,k\}\in E} \kappa_{jk} \left(\hat{c}_k^\dagger \hat{c}_j \rho \hat{c}_k \hat{c}_j^\dagger - \frac{1}{2} \{ \hat{c}_k^\dagger \hat{c}_j \hat{c}_k \hat{c}_j^\dagger \rho + \rho \hat{c}_k^\dagger \hat{c}_j \hat{c}_k \hat{c}_j^\dagger \} \right)$$

- Dynamics described by Quantum Liouville equation

$$\frac{\partial}{\partial t}\rho = -i[H,\rho] + L\rho$$

→ allows for directed walks/projections inside clip network

Note composite structure of clips, e.g. percept clip $\hat{c} = \hat{\mu}_1^\dagger(s_1) \otimes \hat{\mu}_2^\dagger(s_2) \otimes \cdots \otimes \hat{\mu}_N^\dagger(s_N)$
where $\hat{\mu}_{\mathrm{i}}^\dagger$ denotes a memory operator that excites percept of category i (like, for example, color or shape)

⟶ quantum many-body interactions

Quantum projective simulation

Observation:

Part of quantum projective simulation for learning agents

can be cast into a framework similar to

Dissipation-driven quantum simulation of quantum many-body interactions!

Verstraete, Wolf, Cirac, 2009
Diehl, Zoller, *et al.* 2010

→ Proof of principle demonstrations in optical lattices/ion traps conceivable.

see e.g. Bareiro *et al.* 2012

Some open problems (quantum part) & connections to workshop

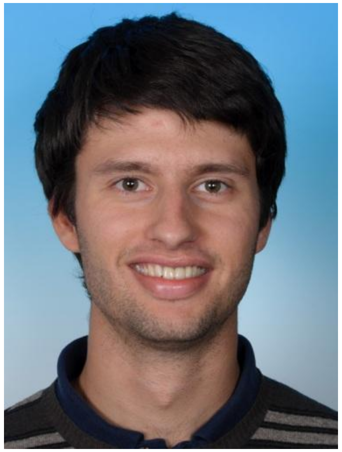
- Proof of speed-up/gap for coherent learning in a specific task environment
  - „Deutsch –Jozsa algorithm“ for learning!?
  - Embedding of specific graph structures into clip network
- Implementation in quantum-optical systems?
  - Ultracold atoms in optical lattices & ion-traps (dissipative quantum simulation) as outlined
  - Optical feedback & control schemes? (talk by Hideo Mabuchi)
- Formulation as quantum feedback networks? (talks by John Gough and Matt James)
- Coherent versus measurement-based control ? (talk by Josh Combes)
  - this mainly concerns the fomulation of the type-I agent: Classical robot (with q-control peripherie) operting in a quantum lab

Miscellanous remarks

- PS = Model for biological learning and behavior
  - Could be used to design behavior experiments : „Is there quantum inside?“
  - How „free“ is an agent?
- PS = Model for “self simulation“
  - We are not talking about one system simulating another system, but about a system simulating *itself* (i.e. its own future)

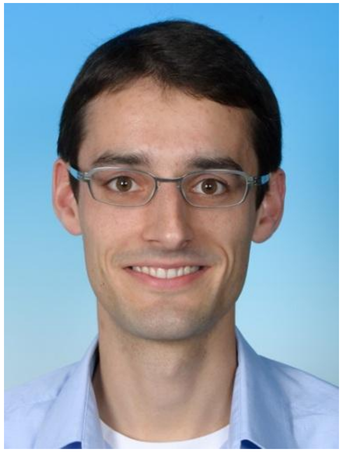
The team

Gemma De las Cuevas

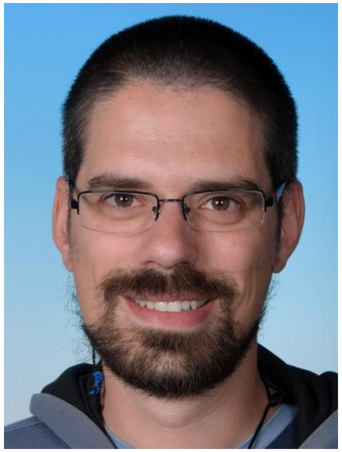


Julian Mautner

Adi Makmal



Markus Tiersch



Daniel Manzano