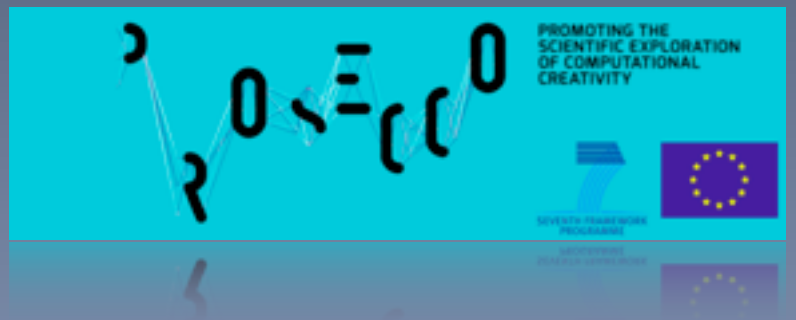
Queen Mary
University of London

Creativity before consciousness

a mechanism admitting spontaneous creativity in Baars' Global Workspace

Geraint A. Wiggins
Professor of Computational Creativity
CCLab
Queen Mary University of London



Outline

Queen Mary University of London

- What I mean by "spontaneous creativity"
- Background
  - Taine's Theatre of Consciousness, the Society of Mind and Global Workspace Theory
    - The Threshold Paradox
  - Statistical models of cognitive process
  - Information theory
- A hypothetical model of cognitive selection that accounts for spontaneous creativity
- Evaluation – a difficult problem
- Motivation: **WHERE DO (MUSICAL) IDEAS COME FROM?**

Two kinds of creativity

Queen Mary University of London

- One aspect of creativity is **SPONTANEOUS**
  - ideas appear, spontaneously, in consciousness
  - cf. Mozart (Holmes, 2009, p. 317)
    - When I am, as it were, completely myself, entirely alone, and of good cheer – say traveling in a carriage, or walking after a good meal, or during the night when I cannot sleep; it is on such occasions that my ideas flow best and most abundantly.
- Compare with the composer working to build (e.g.) a new version of a TV theme, on schedule, and with constraints on "acceptable style"
  - this is a different kind of activity: **CREATIVE REASONING**
- Most creative acts of any size are a **mixture of both**
- Here, I focus on **SPONTANEOUS CREATIVITY** only

Background: societies of mind

Queen Mary University of London

- Hippolyte Taine (1871) proposed the first (?) multi-agent theory of mind, based in a **Theatre of Consciousness**
  - narrow theatre stage, with actors appearing, disappearing, and planning off-set
- Marvin Minsky (1987) proposed the **Society of Mind**
  - computational knowledge-rich agents, communicating & collaborating hierarchically to achieve goals
- Bernard Baars (1988) proposed the **Global Workspace Theory**
  - agents, generating cognitive structures, communicating via a shared blackboard
  - agnostic as to nature of agent-generators
- The three theories are not incompatible
  - Baar's agents/representations are underspecified, and don't contradict Minsky's
  - The key difference is in the communication mechanism
    - but even that may not be contradictory...

Background: societies of mind

Queen Mary University of London

- Society of Mind uses a hierarchical structure of control
  - agents recruit other agents according to task
  - communication passes up and down hierarchy
  - locus of consciousness is explicitly excluded
- Global Workspace Theory uses a central communication exchange, the Global Workspace
  - corresponds with Taine's "theatre" of consciousness
  - can hold one item at a time (some researchers suggest this should be 2 or 2.5)
  - all agents have read-access to Global Workspace
  - in later developments, Baars proposes a hierarchical system of "local" workspaces feeding into the Global Workspace, reducing information overload
  - there is a "threshold" to be "crossed" to get write-access to the GW
  - granting access can be viewed as **assignment of conscious attention**

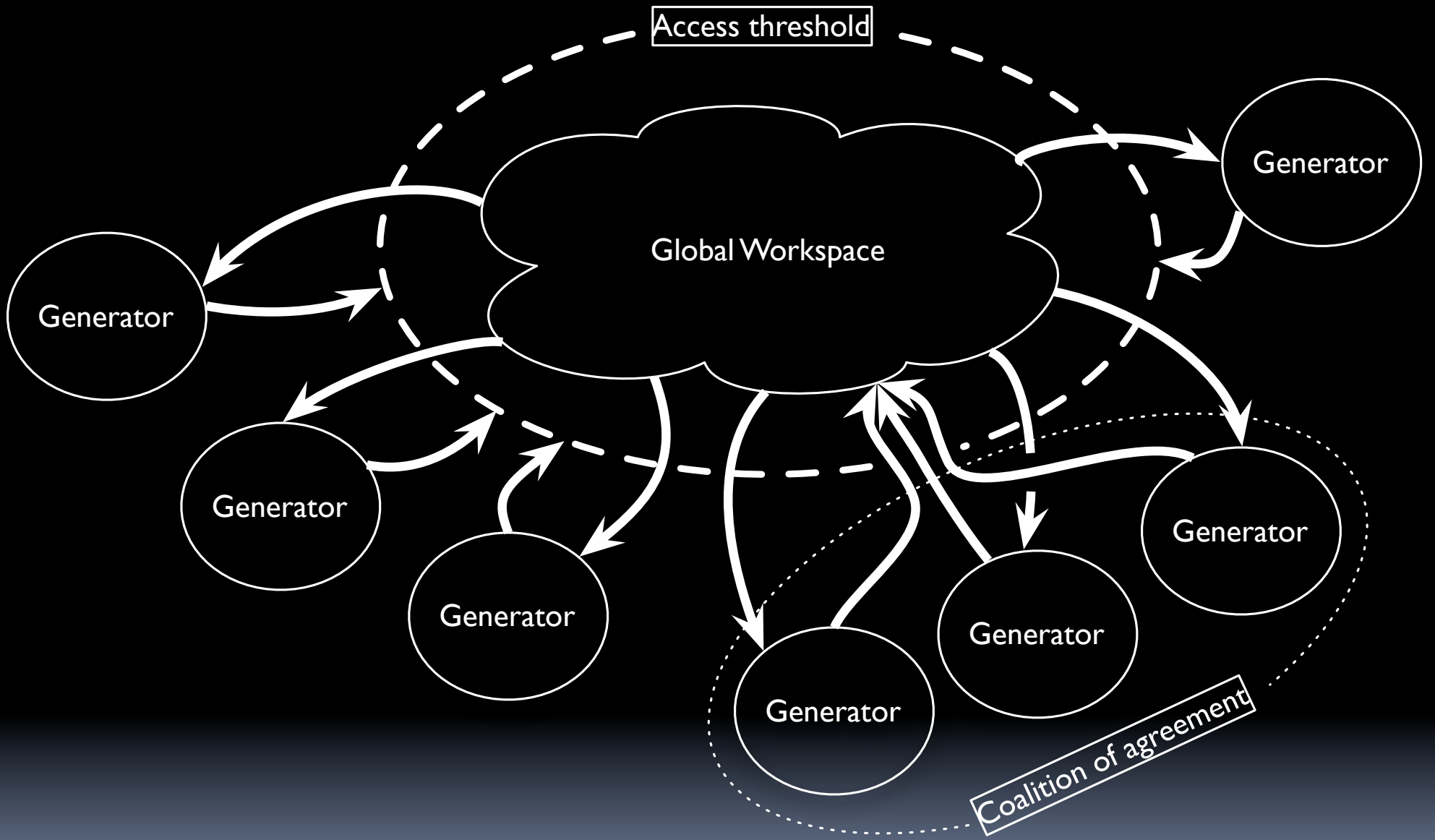
Background: the threshold paradox

Queen Mary University of London

- Baars writes (somewhat metaphorically) about agents "recruiting" others to support a given cognitive structure
  - when enough agents support the structure it is "loud" enough to pass the threshold and enter consciousness
    - I'll use this analogy of "volume" later; Baars proposes synchrony as the implementing mechanism and Shanahan (2010) identifies the necessary neural substrate
- However, there is a problem: **The Threshold Paradox**
  - To communicate in the global workspace, an agent needs to recruit supporters
  - To recruit supporters, an agent must communicate in the global workspace
    - because that is the only medium of communication
- This talk is about an alternative view of access to the Global Workspace

Background: the threshold paradox

Queen Mary University of London



Background: information theory

Queen Mary University of London

- I use two versions of Shannon's *entropy* measure (MacKay, 2003)
  - the number of bits required to transmit data between a hearer and a listener given a shared data model
  - *information content*: estimated number of bits required to transmit a given symbol as it is received:

$$h = -\log_2 p_s$$

    - models **unexpectedness**
  - *entropy*: expected value of the number of bits required to transmit a symbol from a given distribution, prior to sending/receipt:

$$H = -\sum_i p_i \log_2 p_i$$

    - models **uncertainty**
  - p_s, p_i are probabilities of symbols; i ranges over all symbols in the alphabet

Background: statistical cognitive models

Queen Mary University of London

- Organisms need to be able to **anticipate** the world
  - use (mental) models to predict what is coming next
  - use learned models, trained by observed likelihood
  - use temporal association (implication/consequence)
  - use co-occurrence (conjunction)
- Can model music and language (and other things) in this way
  - currently using IDyOM model (Pearce, 2005; Pearce & Wiggins, 2006)
    - predicts human melodic expectation (R^2=.81; Pearce & Wiggins, 2006)
    - predicts human melodic segmentation (F_1=.61; Pearce, Müllensiefen & Wiggins, 2010)
    - predicts language (phoneme) segmentation (F_1=.67; Wiggins, 2011)
- Claim is that mental process is literally statistical
  - statistical nature means we can apply information theory (Shannon, 1948)

Instantiating the Global Workspace

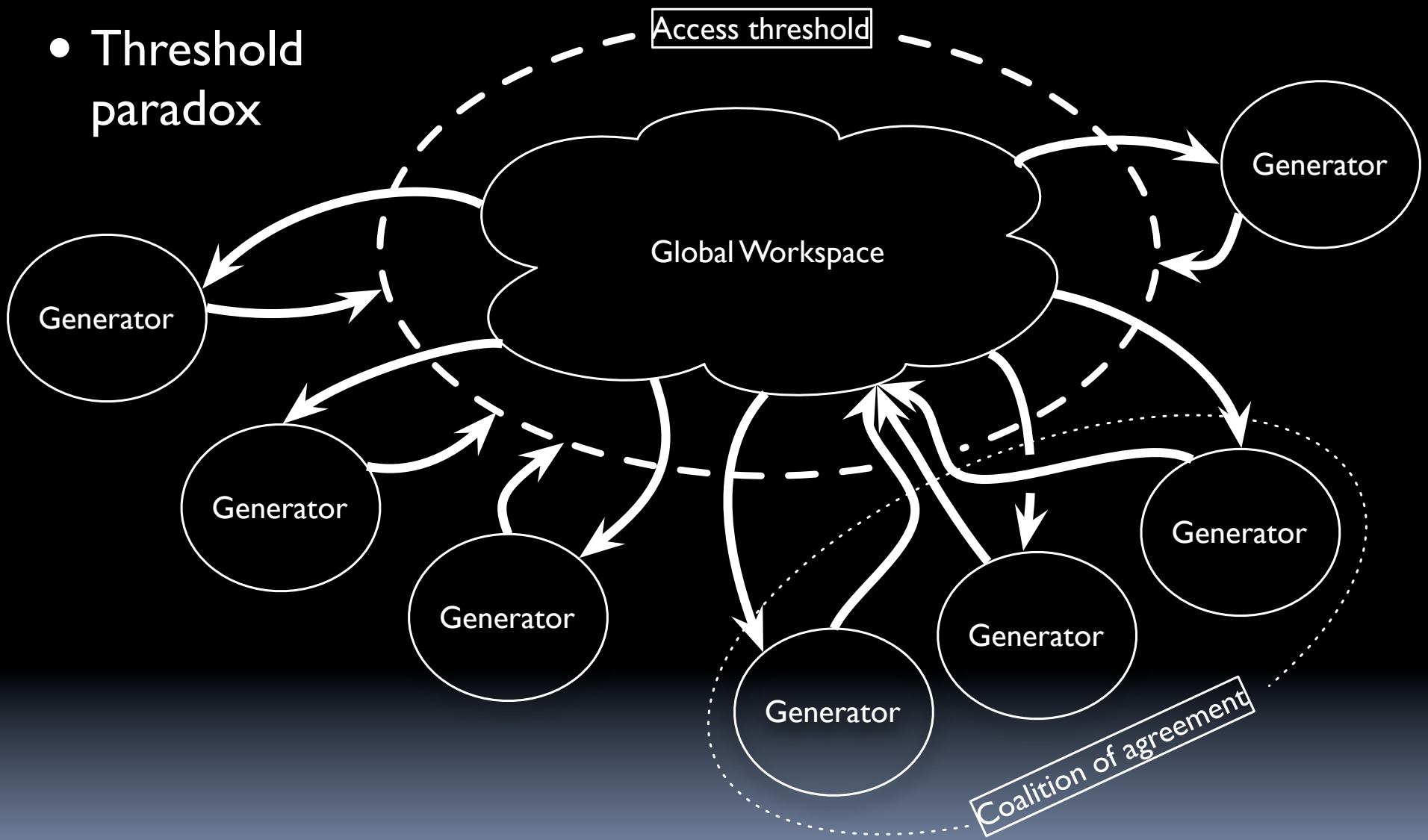
Queen Mary University of London

- Agent generators (not specified by Baars; simpler than Minsky's?)
  - statistical samplers predicting next in sequence from shared learned models of perceptual and other domains
  - many agents, working in massive parallel
    - at all times, the likelihood of a given prediction is proportional to the number of generators producing it *(this isn't in Baars' theory, but it will be important later)*
  - receive perceptual input from sensory systems
    - continually compare previous predictions with current world state
  - continually predict next world state from current matched predictions
    - sensory input does not enter memory directly
    - the expectation that matches best is recorded
  - consider state n (current) and state n+1 (next)
    - at state n, we can calculate h_n, H_n, and H_{n+1} (but not h_{n+1}, because it hasn't happened yet)

Baars' (1988) Global Workspace Theory

Queen Mary University of London

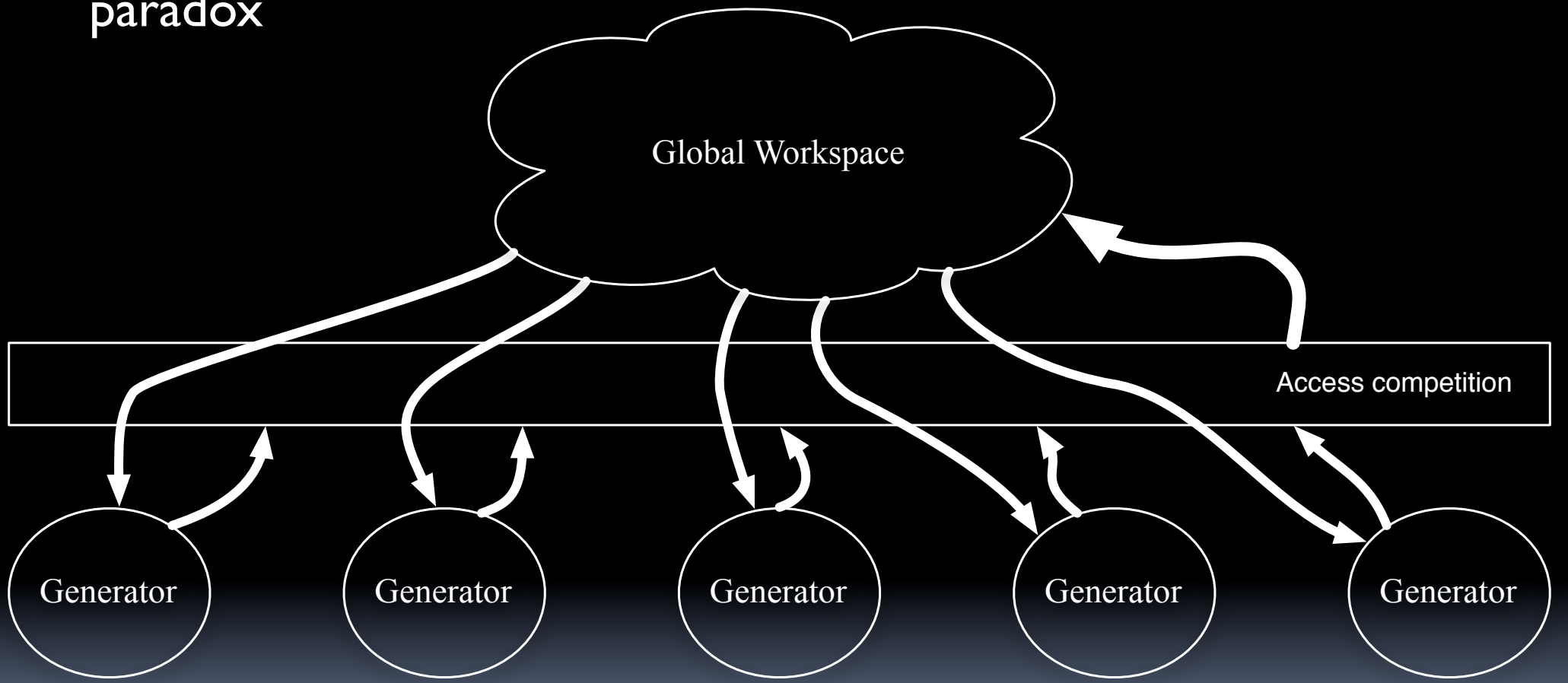
- "Aha" moment = passage into consciousness
- Threshold paradox



Baars' (1988) Global Workspace Theory

Queen Mary University of London

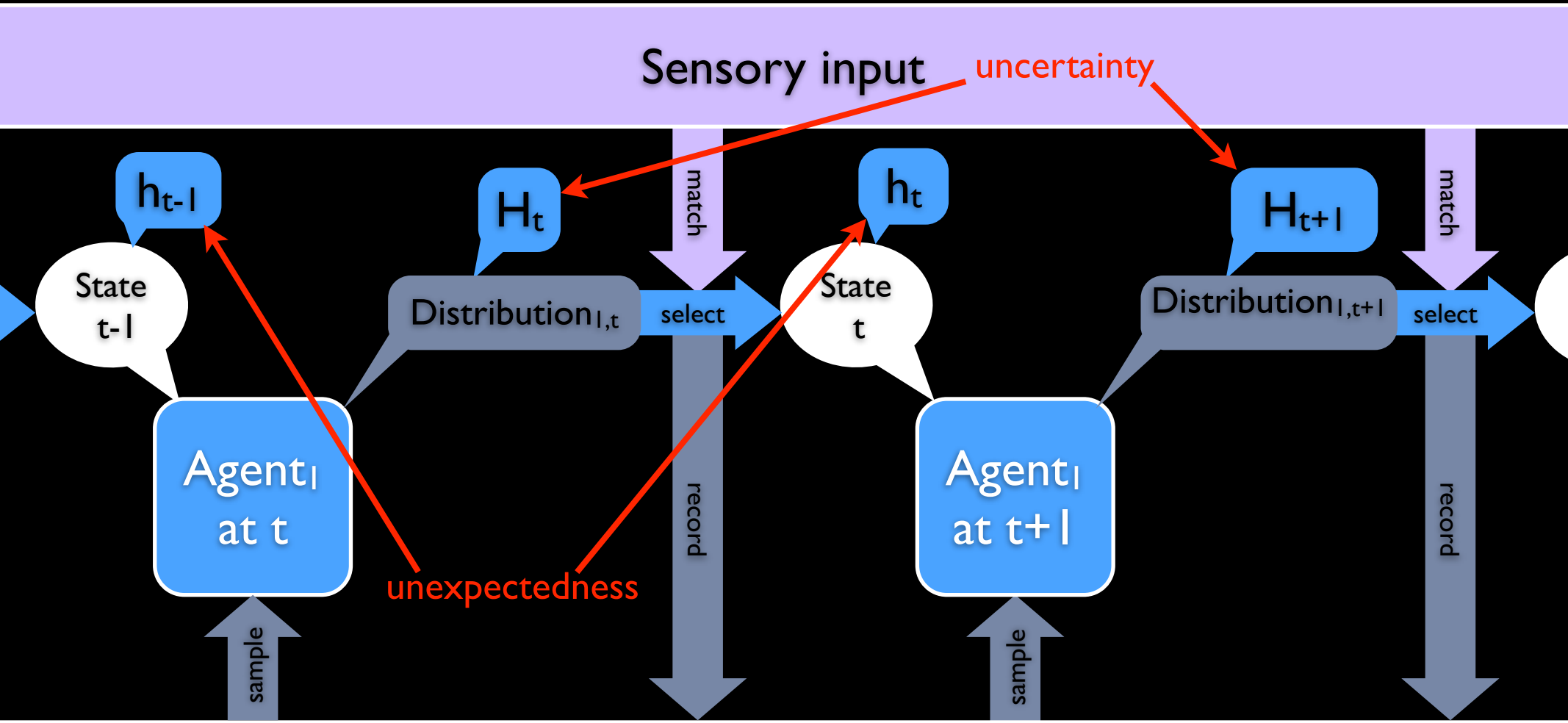
- "Aha" moment = passage into consciousness
- Threshold paradox



- **Big question: what does "recruitment" mean???**

Anticipatory agent

Queen Mary University of London

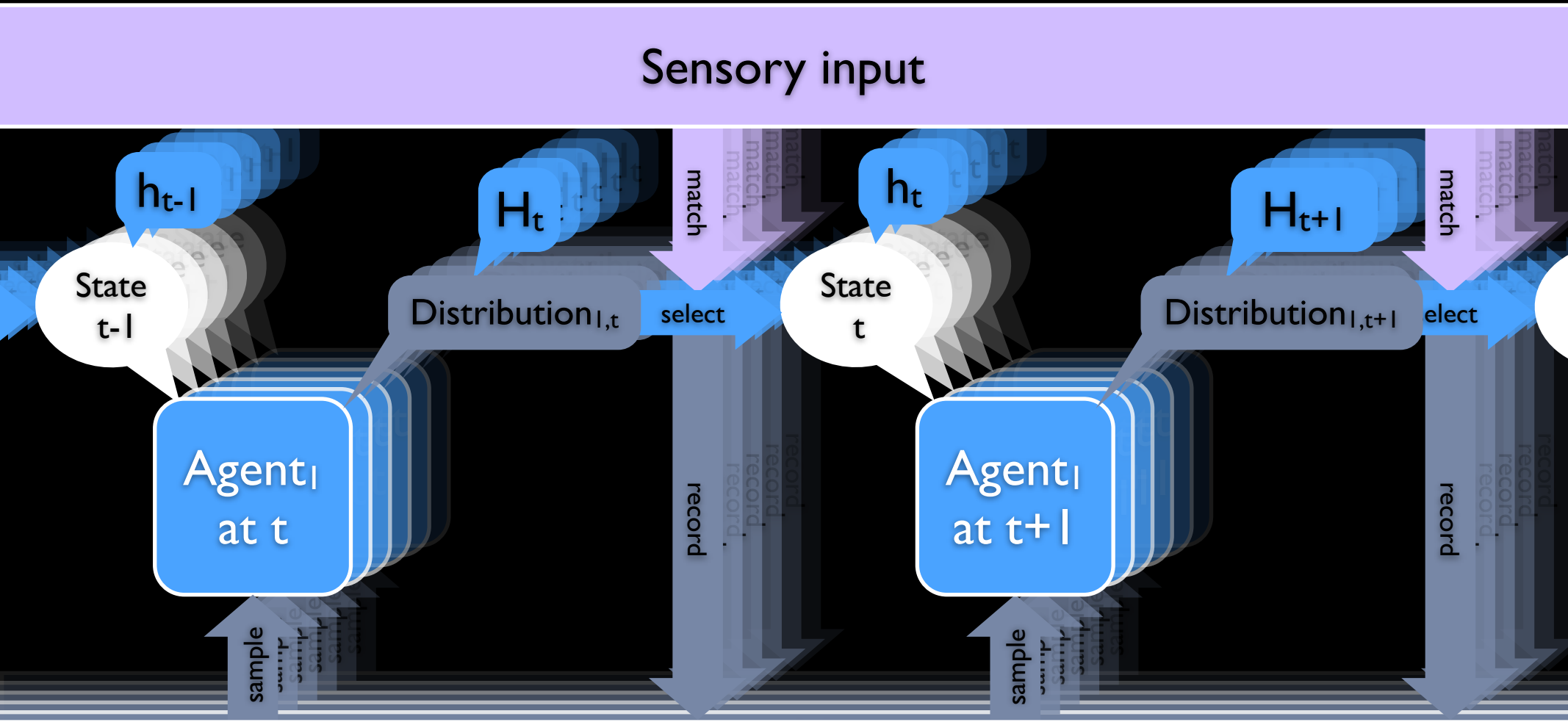


Memory

Time ☞

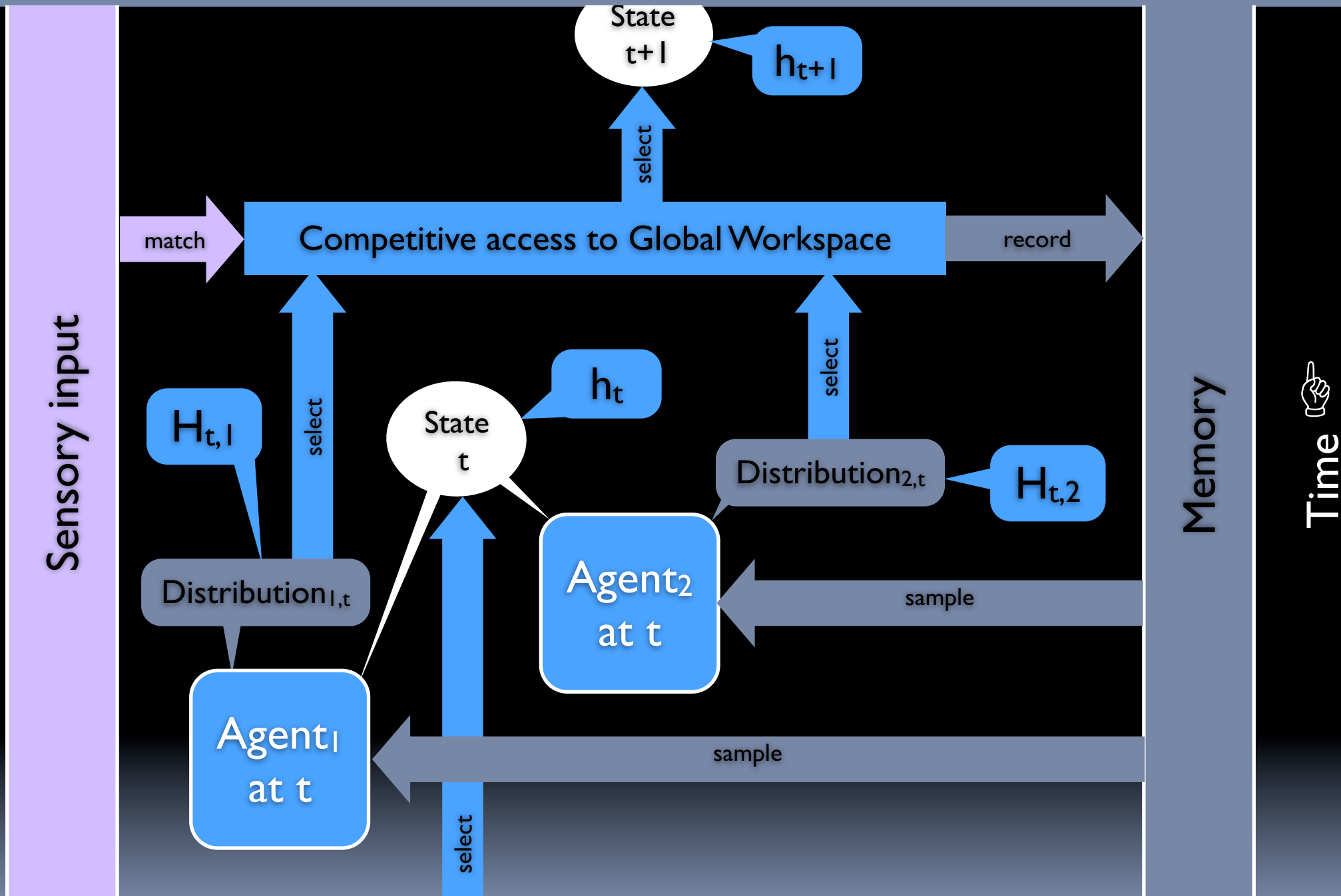
Anticipatory agents

Queen Mary University of London



Anticipatory agents in competition

Queen Mary University of London

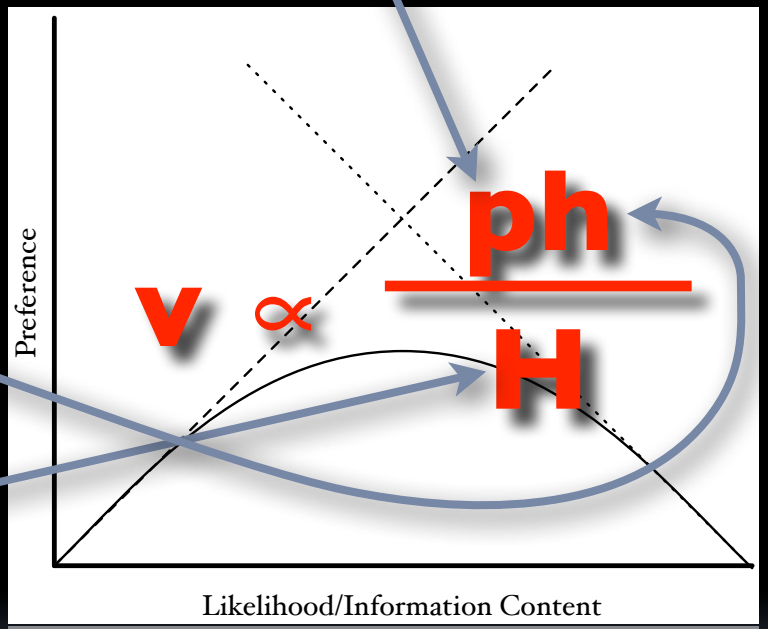


Selecting agent outputs

Queen Mary University of London

Competitive access to Global Workspace

- Agents produce (musical) structure representations
- Probability of structure (in learned model) increases "volume"
  - likely structures are generated more often
  - multiple identical predictions are "additive"
  - avoid "recruitment" question in model
    - need fewer agents?
- Unexpectedness increases "volume"
  - information content predicts unexpectedness
- Uncertainty decreases "volume"
  - entropy predicts uncertainty

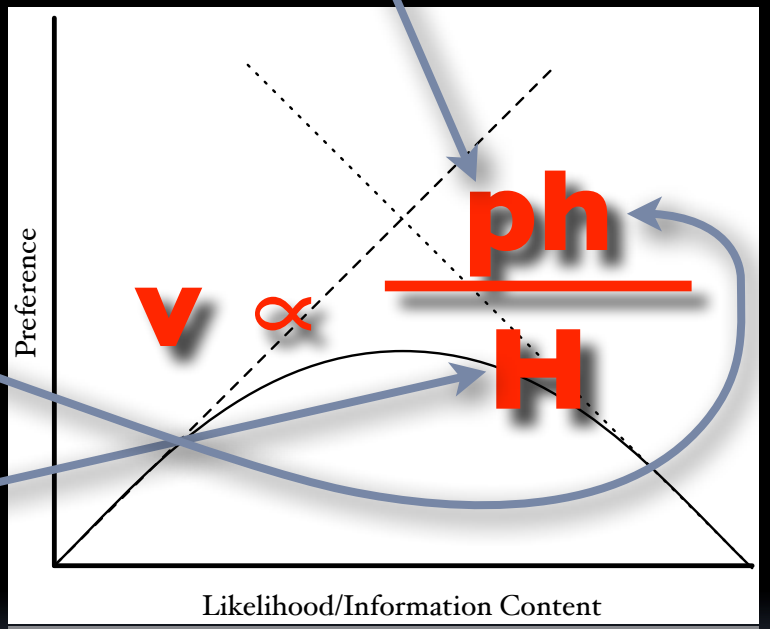


Selecting agent outputs

Queen Mary University of London

Competitive access to Global Workspace

- Agents produce (musical) structure representations
- Probability of structure (in learned model) increases "volume"
  - likely structures are generated more often
  - multiple identical predictions are "additive"
  - avoid "recruitment" question in model
    - need fewer agents?
- Unexpectedness increases "volume"
  - information content predicts unexpectedness
- Uncertainty decreases "volume"
  - entropy predicts uncertainty
- Predictions matched with sensory input, but can compete without it



Spontaneous creativity to order

Queen Mary University of London

- In the absence of distracting perceptual input, generators freewheel
- Predictions are produced from memory, spontaneously
- Some may be prioritised enough to enter consciousness as "ideas"
  - cf. Wallas (1926) "illumination"
  - the "Aha!" moment
- Such ideas can be selected...

Where to find more

Queen Mary University of London

- Full (long) paper:
  - Wiggins, G. (2012) The Mind's Chorus: Creativity before Consciousness. Cognitive Computation. Special issue on Computational Creativity, Intelligence and Autonomy, 4(3):306–319

Final example

Queen Mary University of London

- Example: harmony by Raymond Whorley's autonomous composer
  - NB statistical model alone - no GW, no feedback, no deep learning

Mozart's explanation (Holmes, 2009)

Queen Mary University of London

When I am, as it were, completely myself, entirely alone, and of good cheer – say traveling in a carriage, or walking after a good meal, or during the night when I cannot sleep; it is on such occasions that my ideas flow best and most abundantly. Whence and how they come, I know not; nor can I force them. Those ideas that please me I retain in memory, and am accustomed, as I have been told, to hum them to myself.

All this fires my soul, and provided I am not disturbed, my subject enlarges itself, becomes methodized and defined, and the whole, though it be long, stands almost completed and finished in my mind, so that I can survey it, like a fine picture or a beautiful statue, at a glance. Nor do I hear in my imagination the parts successively, but I hear them, as it were, all at once. What a delight this is I cannot tell! All this inventing, this producing takes place in a pleasing lively dream. Still the actual hearing of the toutensemble is after all the best. What has been thus produced I do not easily forget, and this is perhaps the best gift I have my Divine Maker to thank for.

Mozart's explanation (Holmes, 2009)

Queen Mary University of London

When I am, as it were, completely myself, entirely alone, and of good cheer – say traveling in a carriage, or walking after a good meal, or during the night when I cannot sleep; it is on such occasions that my ideas flow best and most abundantly. Whence and how they come, I know not; nor can I force them. Those ideas that please me I retain in memory, and am accustomed, as I have been told, to hum them to myself.

All this fires my soul, and provided I am not disturbed, my subject enlarges itself, becomes methodized and defined, and the whole, though it be long, stands almost completed and finished in my mind, so that I can survey it, like a fine picture or a beautiful statue, at a glance. Nor do I hear in my imagination the parts successively, but I hear them, as it were, all at once. What a delight this is I cannot tell! All this inventing, this producing takes place in a pleasing lively dream. Still the actual hearing of the toutensemble is after all the best. What has been thus produced I do not easily forget, and this is perhaps the best gift I have my Divine Maker to thank for.