Prediction and Intelligence

DAVID GAMEZ
Middlesex University, London, UK

TOCAI Seminar, Middlesex University, London, 14th February 2020

Talk Overview

- Human Intelligence.
- Non-human Intelligence.
- Intelligence and Prediction
- Conclusion.

HUMAN INTELLIGENCE

Intelligence

- Complex multifaceted term – many overlapping definitions have been put forward.
- For example:
  - Cognitive ability.
  - Rational thinking.
  - Problem solving.
  - Goal-directed adaptive behaviour.

Multiple Intelligences

- Some have claimed that there are many different types of intelligence (Gardner 2006).
- Warwick (2000) describes it as a high-dimensional space of abilities.

Emotional Intelligence

- Awareness of emotions and ability to reason about emotions in self and others.
- Emotions closely linked to goal selection – Damasio’s (1994) somatic marker hypothesis.
- Not obvious that our ability to understand and reason about emotions is substantially different from our ability to understand and reason about our environment.

Fluid vs Crystallized Intelligence

- Distinction introduced by Cattell (1971).
- Fluid intelligence is some kind of aptitude or capacity for learning and solving new problems.
- Crystallized intelligence is the reproduction of stored knowledge.
- Early IQ tests measured both.
- Modern researchers mostly focus on fluid intelligence: a person's ability to solve new problems regardless of previous education and environment.

Measurement of Intelligence

- Intelligence cannot be directly measured.
- Typically measure human intelligence using a battery of tests of verbal, mathematical and spatial reasoning abilities.
- Results from these tests are typically converted into values of IQ or *g*-score.

IQ

- Run test battery on representative sample of population.
- Calculate mean and standard deviation.
- Mean score corresponds to IQ of 100.
- Each standard deviation corresponds to 15 IQ points.
- IQ can be used to rank individuals.
- The actual value is a population-derived measure – it does not correspond to a property of an individual.

g

- Factor analysis carried out on test results from a sample of the population.
- Factors related to specific cognitive abilities, such as reasoning and memory, can explain results of closely related tests.
- These factors are linked to a single underlying factor g.
- Tests have a *g*-loading, which corresponds to how much of their variability is explained by *g*.
- *g* is thought to correspond to general intelligence.

IQ and *g*

- Controversial and have often been misused.
- Can be used to rank individuals who take a battery of tests.
- Have played a useful role in scientific research.
- Can be an effective way of processing large numbers of applicants for jobs, education or the military.

Non-Human Intelligence

Animal Intelligence

- Measurement of intelligence with test batteries has some plausibility with humans.
- We generally agree about which behaviours are linked to intelligence.
- More problematic to use test batteries to compare human and animal intelligence.
- How can we develop a test battery that does not neglect different forms of intelligence developed by different animals?

Human vs Pigeon Intelligence

- Suppose we wanted to compare human with pigeon intelligence.
- Test battery could include:
  - Spatial and mathematical reasoning.
  - Mapping and navigation.
- How can we expand this to include the specific intelligences of sheep and fish?

Artificial Intelligence

- These problems become far more acute when we attempt to compare biological and artificial intelligence.
- A computer could easily be programmed to outperform humans on IQ tests.
- This computer might be completely incapable of performing any other task that we consider to be intelligent.
- Test battery approach appears to break down completely when we try to compare human with artificial intelligence.

Neuroscience of Intelligence

- Substantial amount of work on the neuroscience of intelligence (Haier 2017).
- Might enable us to measure intelligence more directly and accurately in humans.
- But this cannot be generalized to animals with very different brain architectures, such as cephalopods and birds.
- Cannot be generalized to machines.

Turing Testing

- Could give up on the idea of a single set of tests to compare human and machine intelligence.
- Focus on the measurement of *human* intelligence in humans, animals and machines.
- Measure the extent to which a machine has matched or exceeded human intelligence on a particular set of tasks.
- Form of Turing testing.

Universal Measures of Intelligence

- Alternative response is to develop more abstract definitions and tests of intelligence that can be applied to both natural and artificial systems.
- Goal-reward based measures developed by Legg and Hutter (2007) and Hernandez-Orallo and Dowe (2010).

Goal/Reward Measures of Intelligence

- Intuition: Intelligence is the ability to achieve goals across a range of different environments.
- Legg and Hutter (2007) propose a measure of intelligence that sums the rewards achieved by the agent across all possible actions and all possible environments.
- Adjusts for the complexity of different environments.

Goal/Reward Measure of Intelligence

- Legg and Hutter's universal measure is theoretically interesting, but it is not practical to sum rewards across all possible actions and environments.
- Hernandez-Orallo and Dowe (2010) have developed similar goal/reward measures that are more practical.
- Not a theory about intelligence – just a more universal way of measuring it.
- Solely based on external behaviour – cannot measure intelligence from internal states.

Intelligence and Prediction

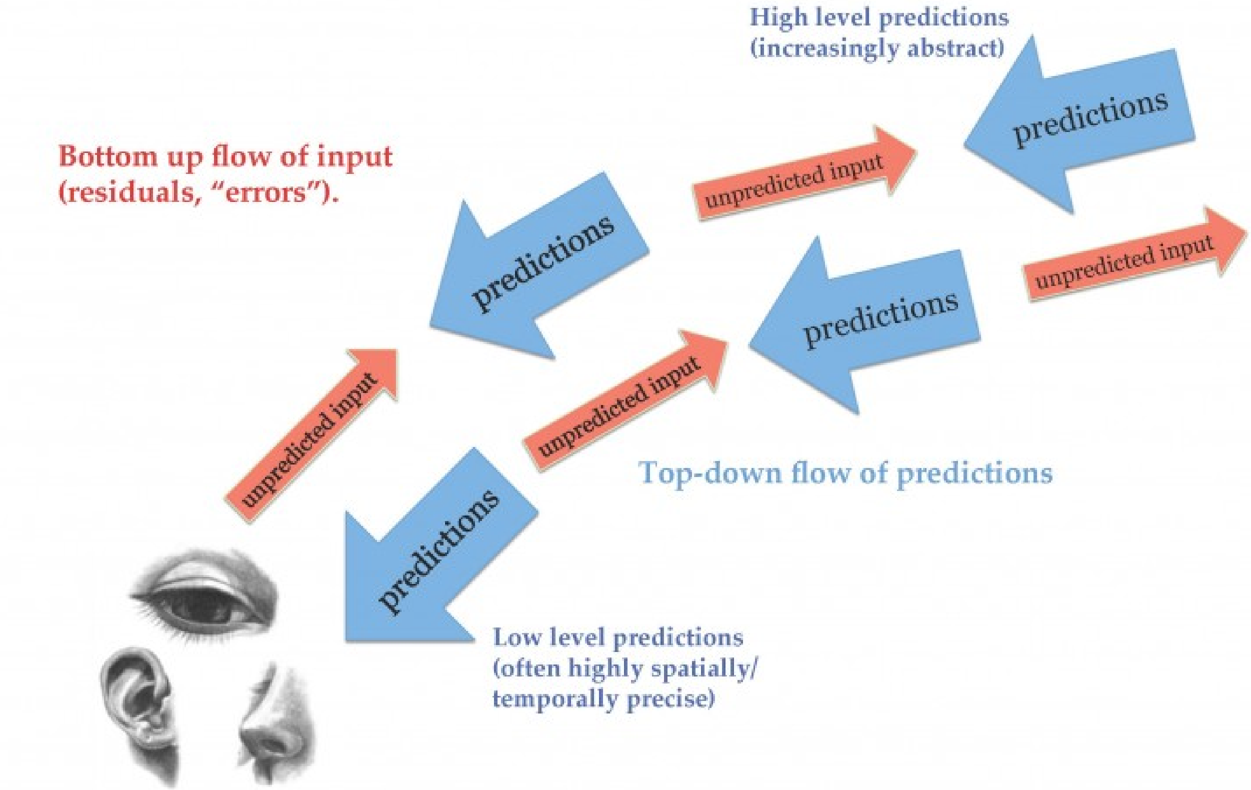
Intelligence and Prediction

- Many behaviours linked to intelligence, such as spatial and mathematical reasoning, require prediction.
- Our ability to get rewards from different environments is closely linked to our ability to predict the consequences of our actions in different environments.

Predictive Brain Hypothesis

- Large amount of interest in the predictive brain hypothesis (Clark 2016).
- Brain is organized into hierarchies.
- Each layer predicts the state of the layer below it and receives an error signal about the accuracy of its prediction.
- If brains are intelligent and the brain's core function is prediction, then brains that are better predictors should be more intelligent.

Predictive Brain Hypothesis



The basic Predictive Processing schema

Machine Learning

- Many forms of machine learning are closely aligned with spatial or temporal prediction.
- A predictive measure of intelligence would be a good fit for AI.

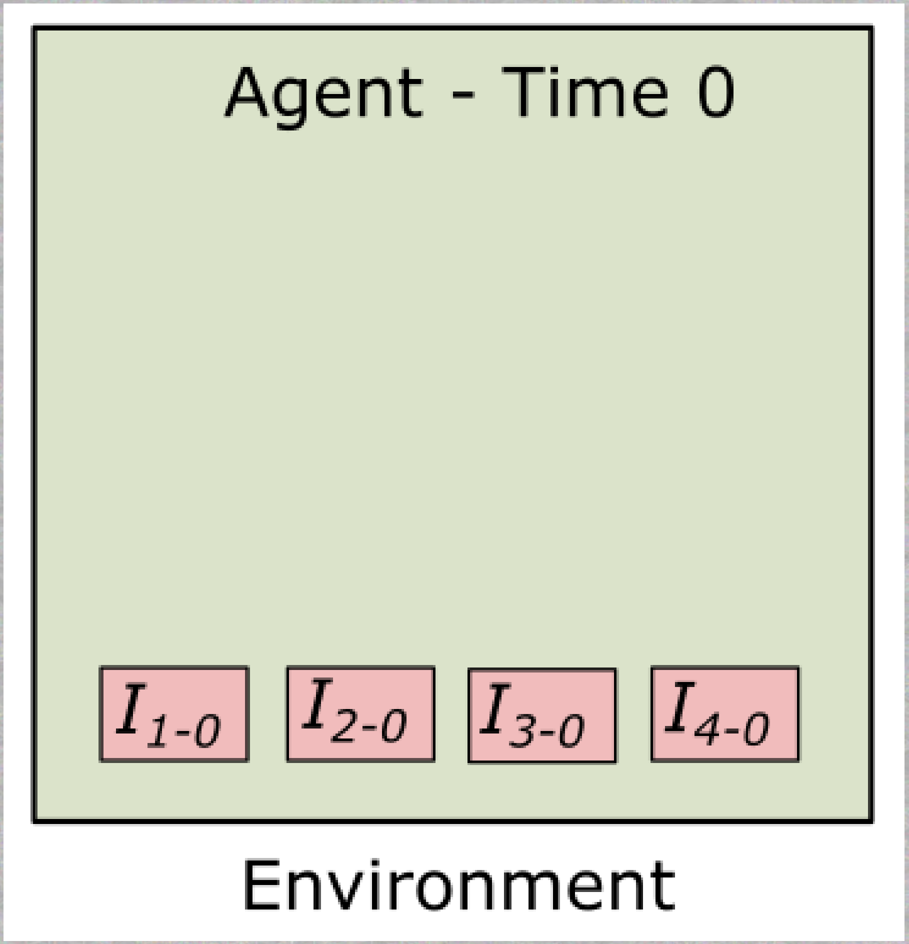
Measure of Predictive Intelligence

- In 2019 I started working on an algorithm for measuring predictive intelligence.
- This talk outlines the current version.
- Work in progress.
- Suggestions and offers of collaboration welcome!

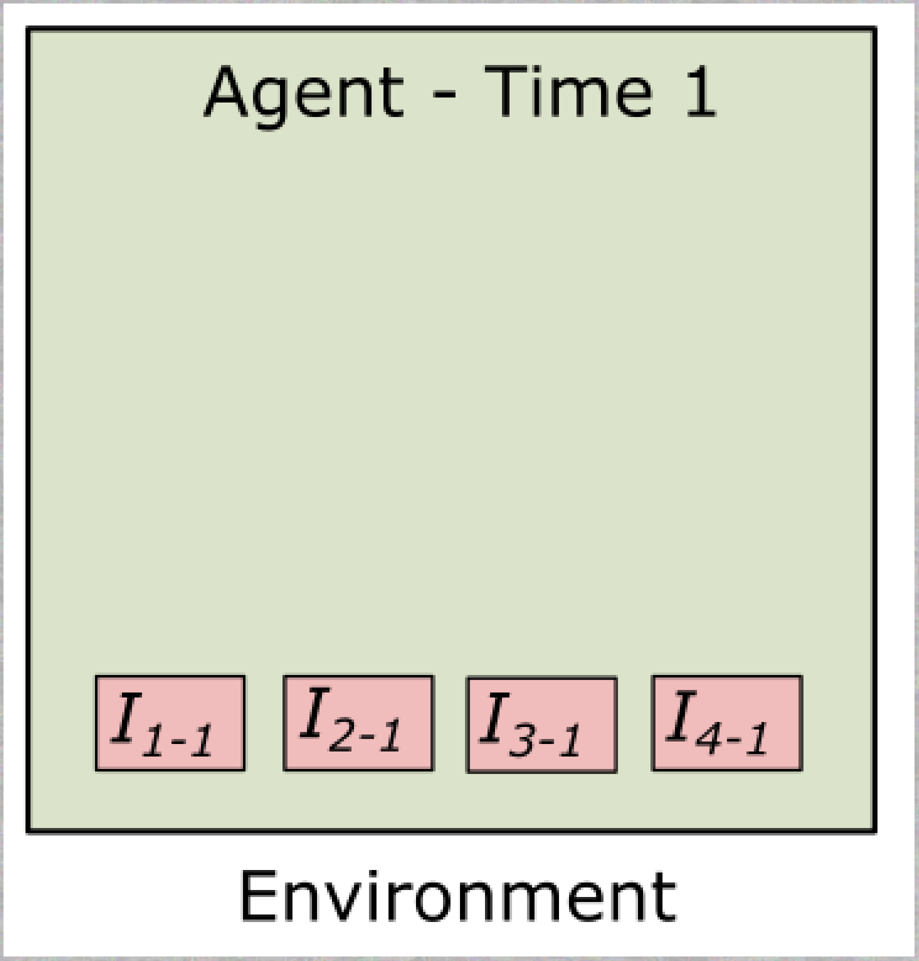
Agent

- Will consider a simple agent with internal states that are connected to the environment.
- Input states $\boldsymbol{I_1} \ldots \boldsymbol{I_n}$ change in time as environment changes and agent interacts with environment.

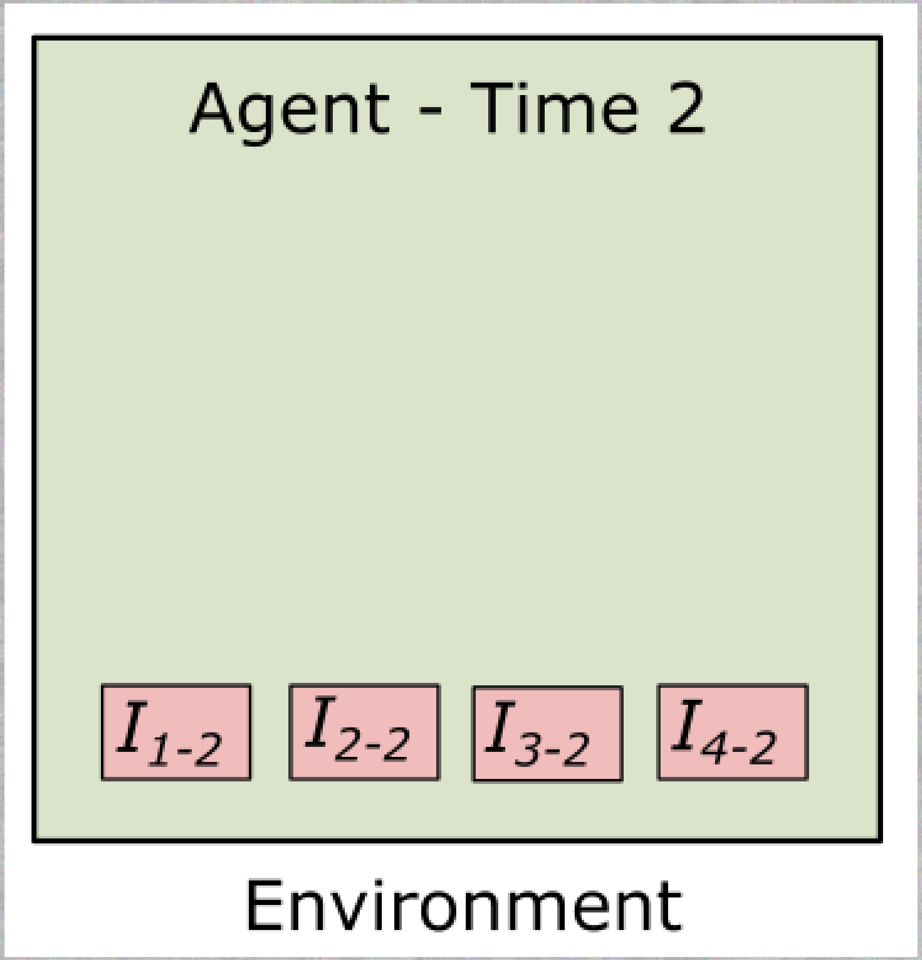
Agent



Agent



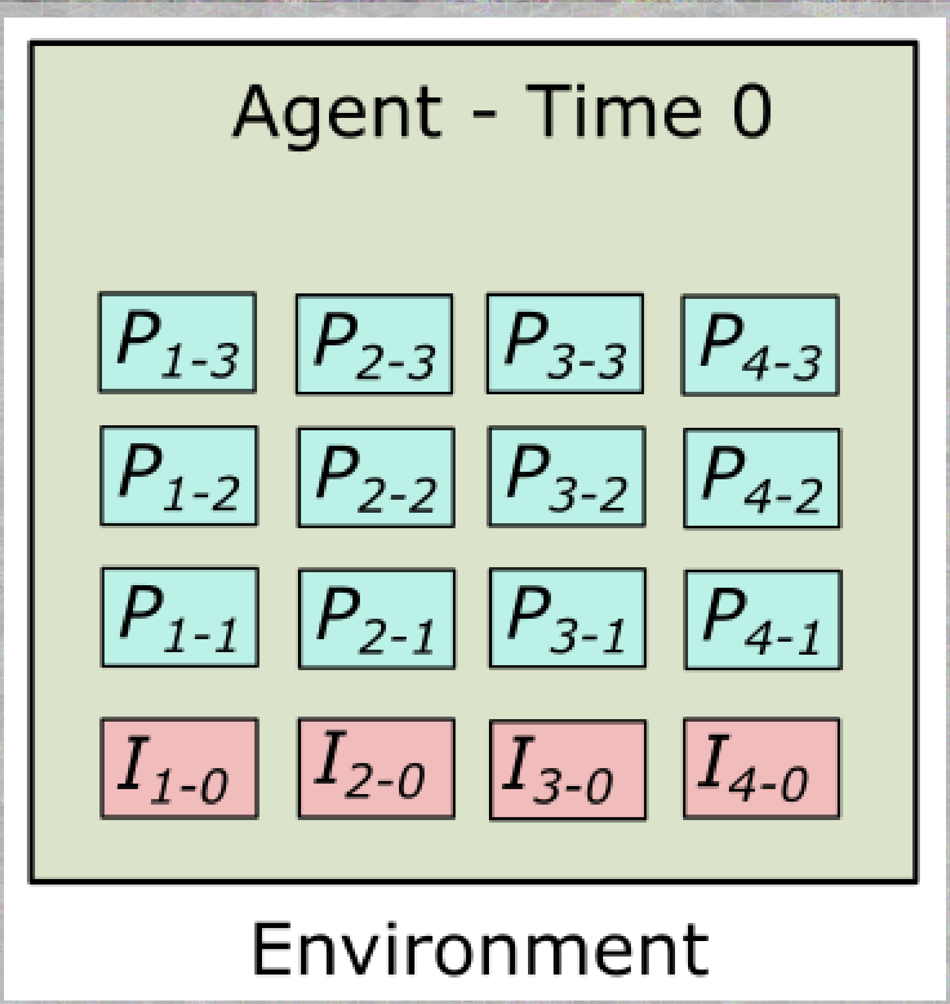
Agent



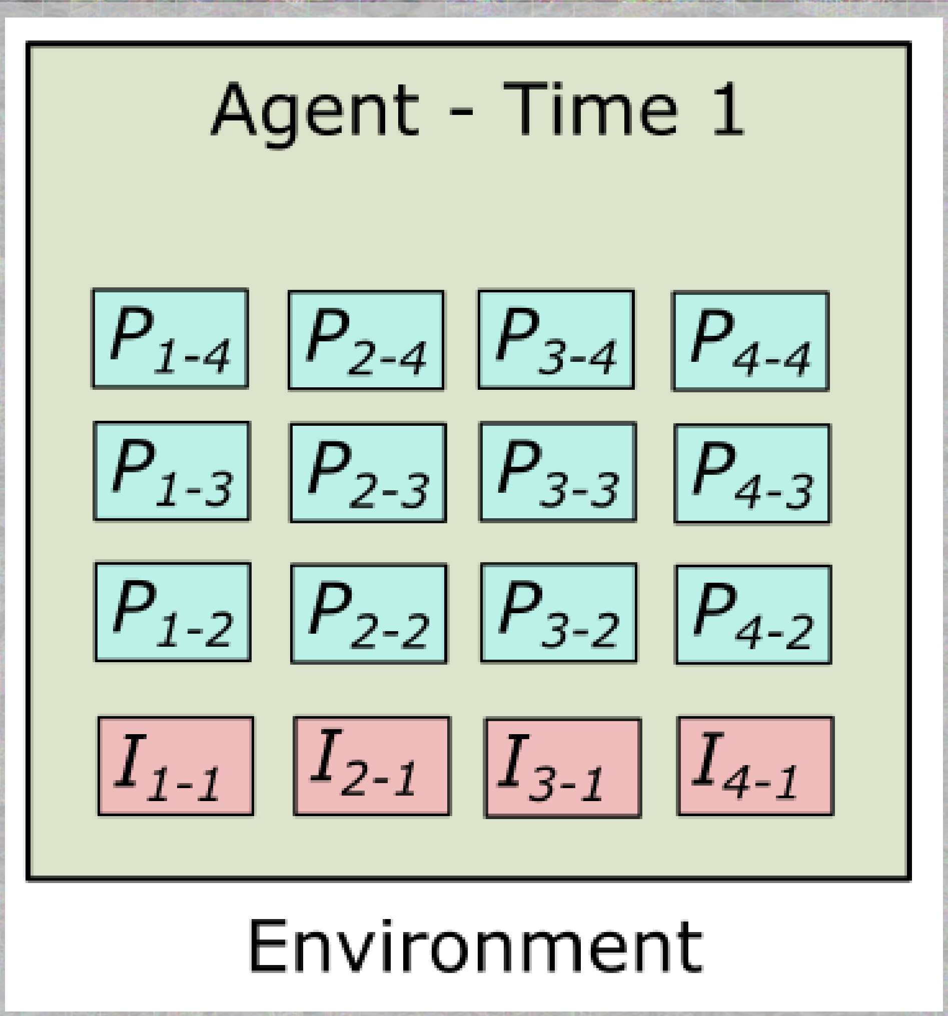
Agent Predictions

- Agent generates time indexed predictions $\boldsymbol{P_1} \ldots \boldsymbol{P_n}$ about its future internal states.

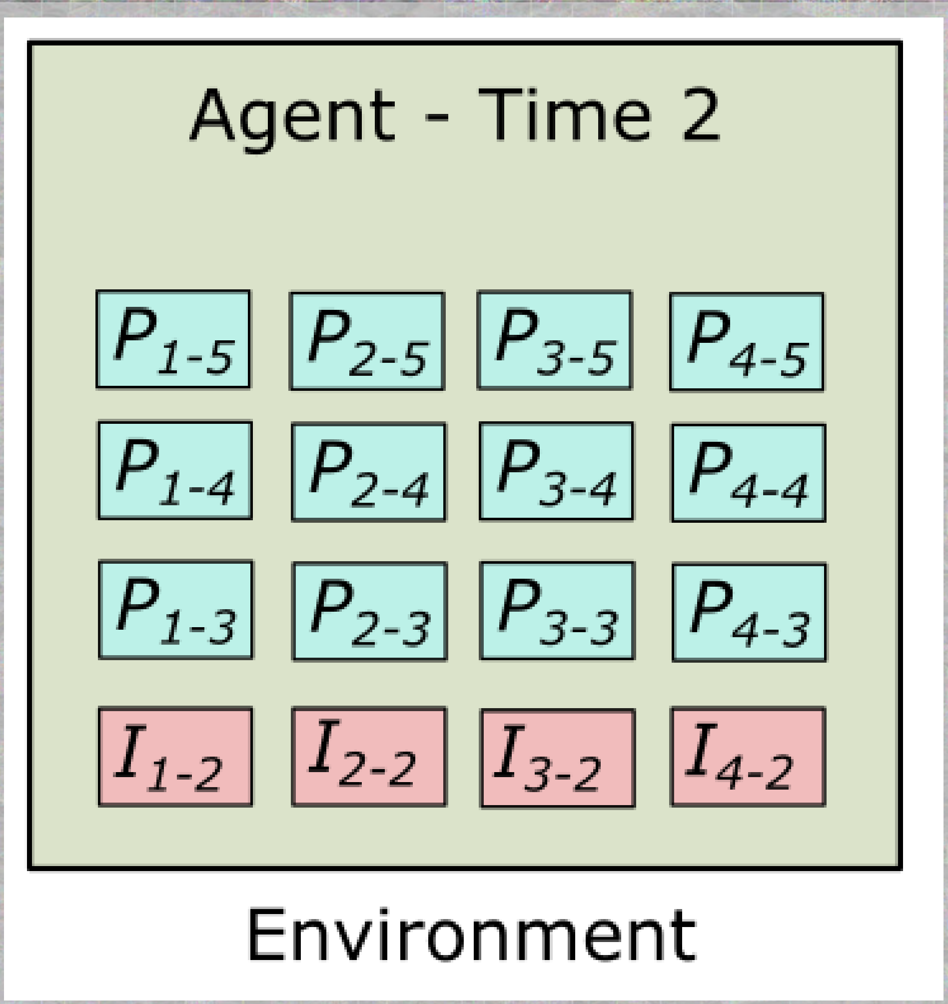
Agent Predictions



Agent Predictions



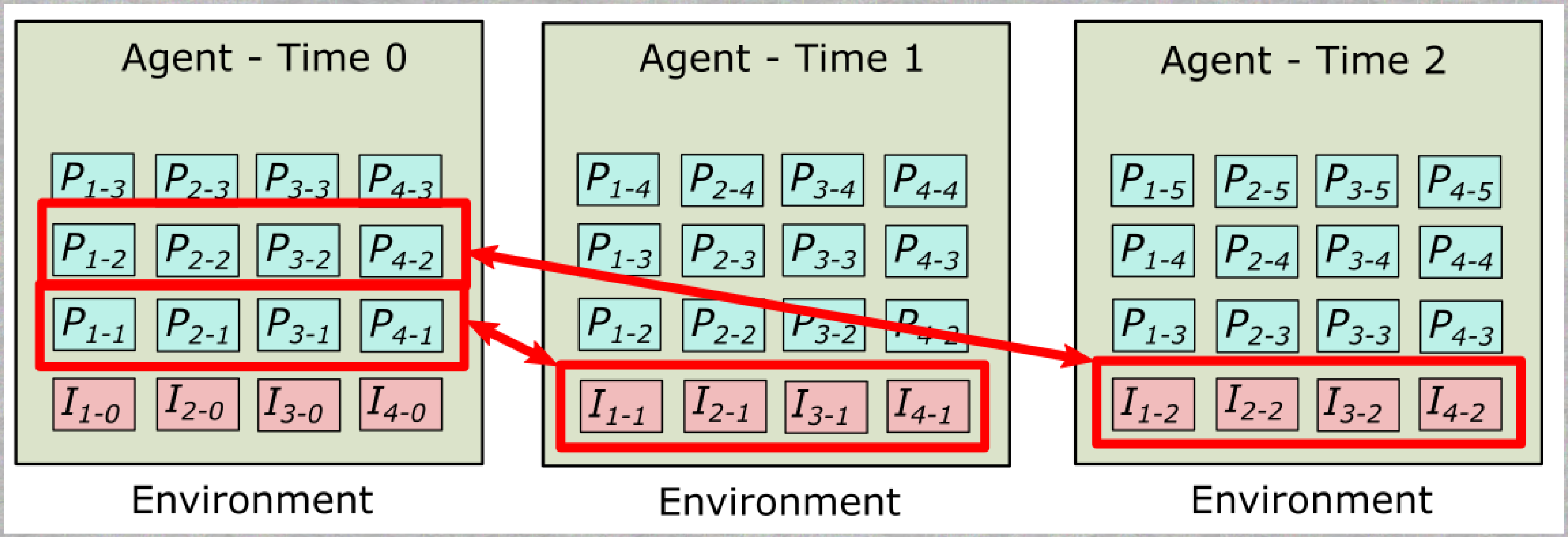
Agent Predictions



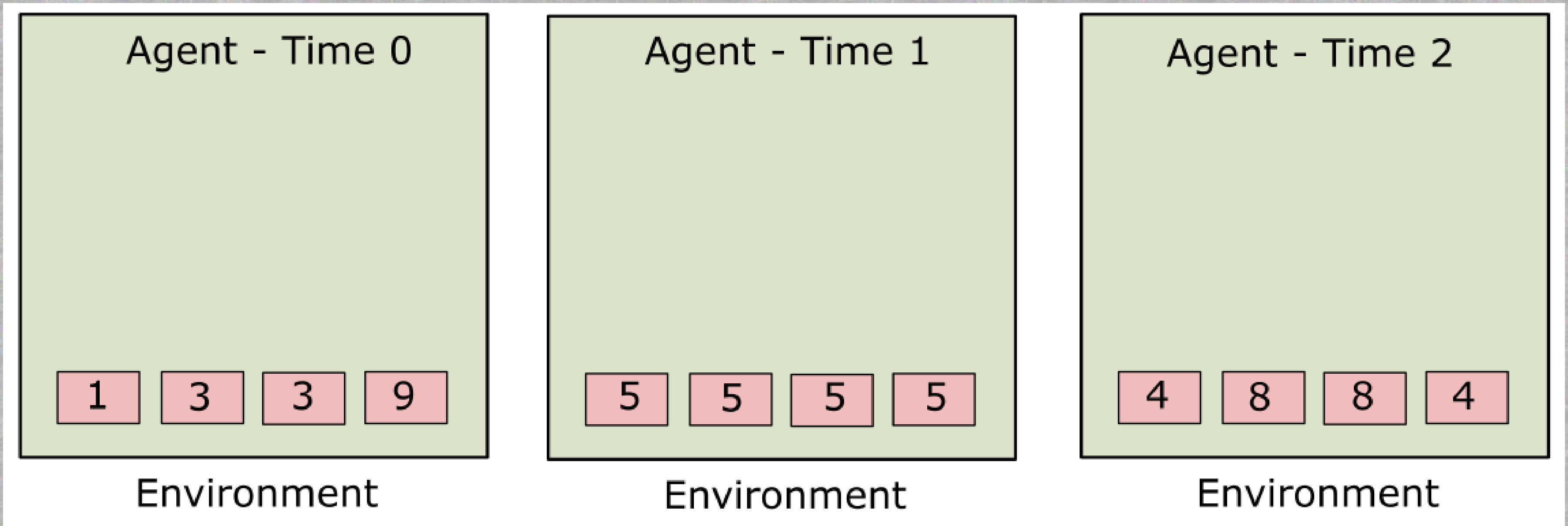
Internalist Approach

- Different agents have different sensors and process data from their environment in different ways.
- Cockroaches and humans experience music very differently.
- Cannot directly compare predictions with states of the environment
- Compare agent's predictions about its internal states with its future internal states.

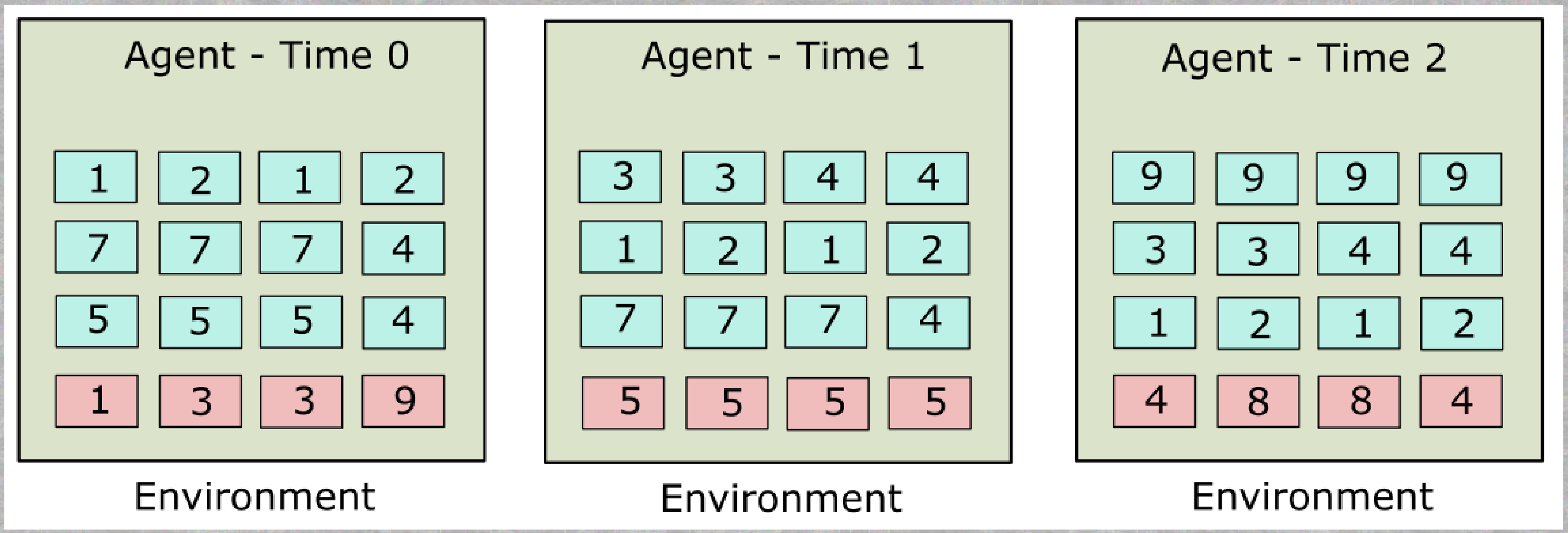
Agent Predictions



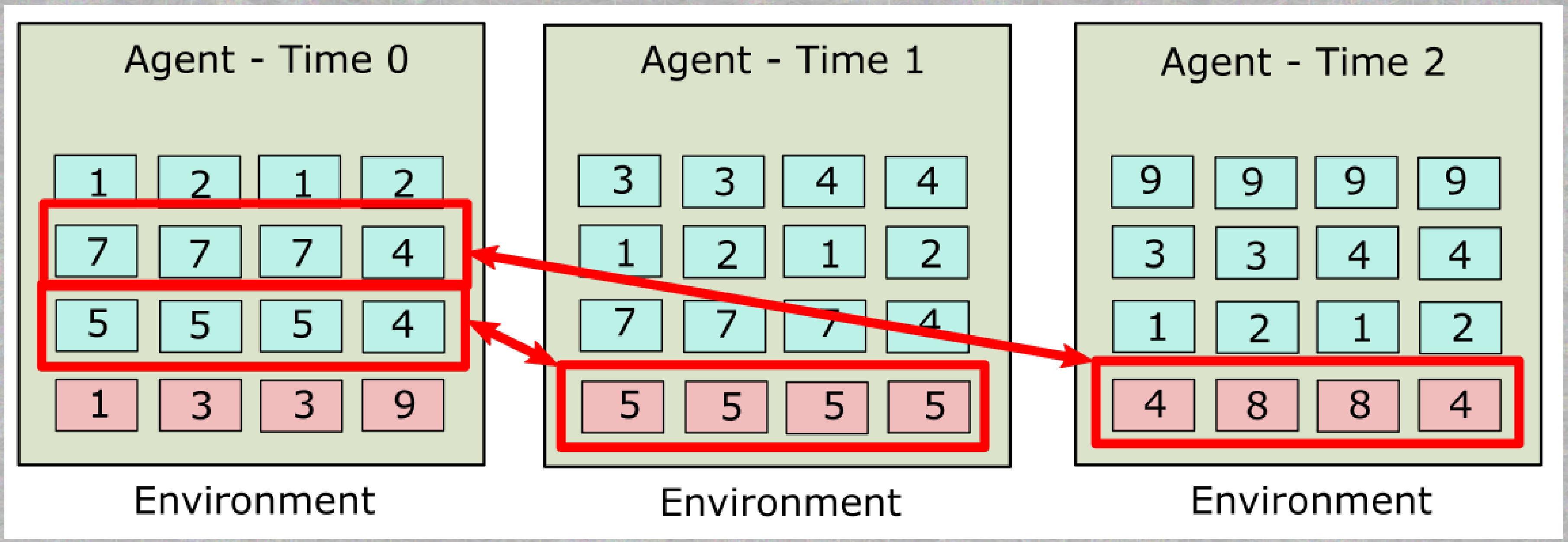
Simple Example



Simple Example



Simple Example



Predictions are Random Variables

- Simple example illustrates the general idea.
- In practice predictions are random variables.
- Future states have a probability distribution.
- This is subjectively true and true of some machine learning algorithms.

Input States are Random Variables

- Input states are fixed at each moment in time.
- However, states can occur with a probability distribution
- When we are comparing predictions with input states, we have to treat the input states as random variables.

Time is a Random Variable

- An event can have different probabilities of occurring at different points in time.
- So time also has to be treated as a random variable.
- So predictions and input states are bivariate distributions across values and times.

Measuring Predictive Intelligence (PI)

- Measure the predictive intelligence (PI) of an agent by comparing predictions with environmental states.
- Use **Helliger distance** to measure the difference between two bivariate distributions.

Helliger Distance

- For discrete distributions:

$$H(P,Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{k} (\sqrt{p_i} - \sqrt{q_i})^2},$$

- Helliger distance is 1 when there is a perfect mismatch so **1-Helliger distance** gives us a measure of the match between a prediction and an input state scaled between 0 and 1.

First Measure: PI-1

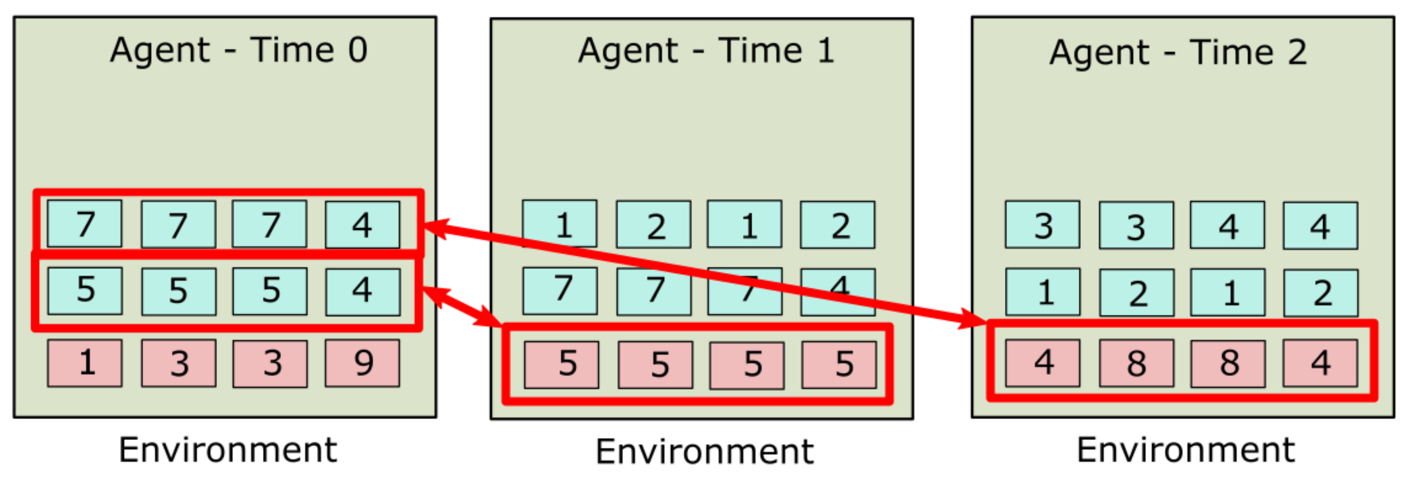
- Work through all possible states of a given environment.
- Put agent in each environment state and use Helliger distance to measure match between predictions and future environmental states.
- $\boldsymbol{p_{match}}$ is the sum of the matches for each environment state.
- Sum $\boldsymbol{p_{match}}$ across all environment states to get predictive power of agent.

First Measure: PI-1

$$\sum_{n=1}^{m} p_{match}$$

Simple Example

- At time 0 $\boldsymbol{p}_{\boldsymbol{match}}$ is 3 (out of 4) for the t_1 predictions and $\boldsymbol{p}_{\boldsymbol{match}}$ is 1 (out of 4) for the t_2 predictions.
- $\boldsymbol{p}_{\boldsymbol{match}}$ for this agent at Time 0 is 4.



Limitations of PI-1

- Simple approach tells you how good an agent is at predicting future states of its environment.
- However, environments could be very simple or trivially repeat – we need to take the complexity of the environment into account.

Kolmogorov Complexity

- Measure of complexity.
- Length of the shortest computer program that can output a given piece of text, image, etc.
- Cannot be computed, so often approximated by compression algorithm.
- Used in other measures of intelligence.

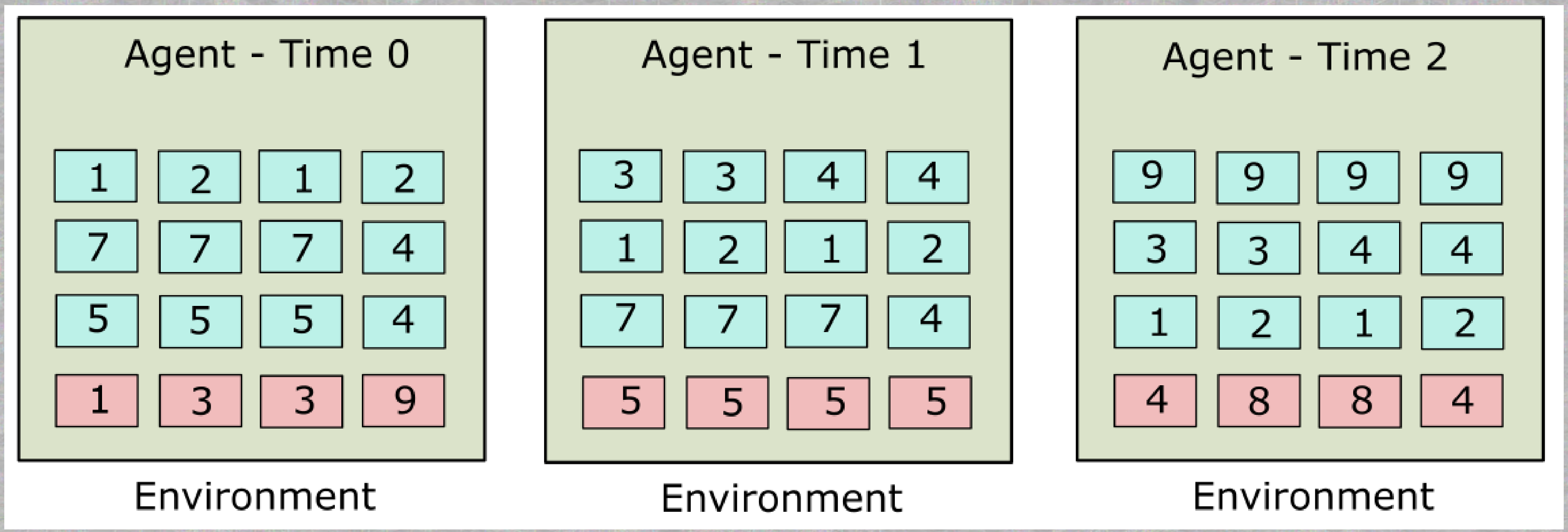
Improved Algorithm: PI-2

- For each state of the environment that the agent is in, calculate p_{match} and divide it by the maximum possible p_{match}, which I will call p_{total}.
- This will give a number between 0 and 1 that corresponds to the overall prediction match of all of the predictive states.
- Multiply this number by the Kolmogorov complexity of the predictions

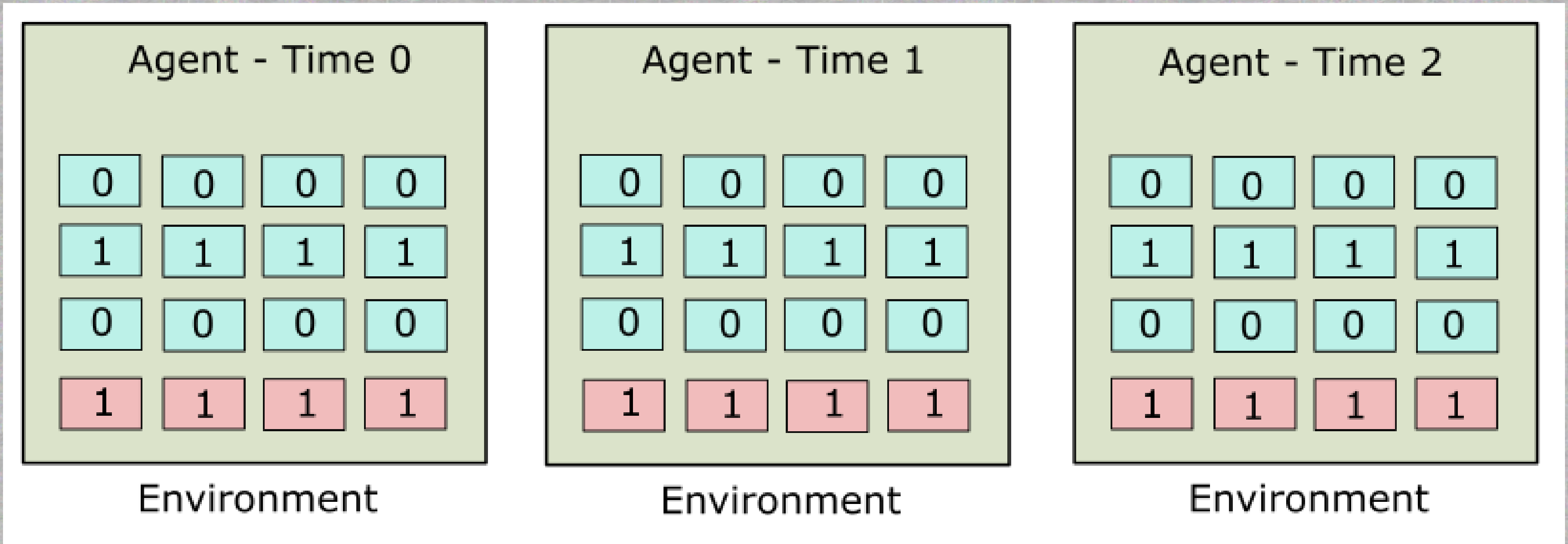
Improved Algorithm: PI-2

$$\sum_{n=1}^{m} K(p_n)\frac{p_{match}}{p_{total}}$$

Complex Environment



Simple Environment



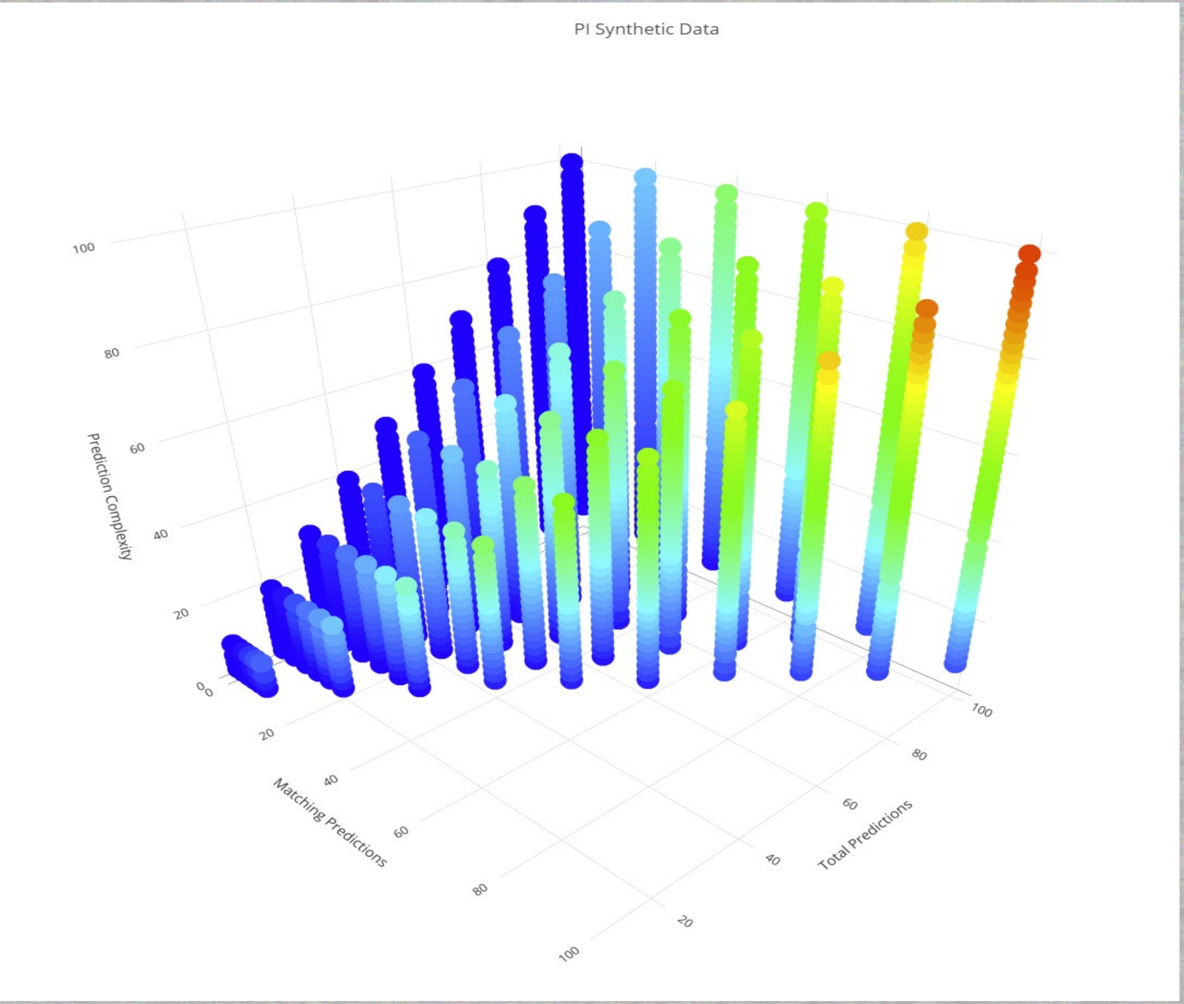
Synthetic Data Experiment

- Generated random strings of different lengths to correspond to predictions that are generated at a particular point in the agent's environment.
- Strings were generated with
  - Different lengths – variation in the number of predictions, $\boldsymbol{p_{total}}$.
  - Prediction match, $\boldsymbol{p_{match}}$.
  - Complexity - the amount of repetition used to generate strings.

Synthetic Data Experiment

- Calculated $\boldsymbol{p}_{match}$ and $\boldsymbol{p}_{total}$ for each string.
- Used string compression library (LZ-UTF8) as an approximation of Kolmogorov complexity.
- Calculated PI-2 for each string and plotted.

Synthetic Data Experiment



Synthetic Data Experiment

- Shows that PI-2 is maximum when we have high numbers of matching predictions of high complexity.
- Good fit for our intuitions!

Limitations of PI-2

- Promising measure that we could use to compare the intelligence of agents within a single environment.
- The same agent with the same input states and the same predictions can display radically different levels of PI-2 in different environments.
- We need to be able to compare agents across different environments, which might only have minor differences between them.

Environment Complexity

- Can potentially measure complexity of the environment by measuring Kolmogorov complexity of input states.
- (Remember that the environment appears differently to each agent).
- One approach might be to combine all of the input states from an environment and calculate an approximation to its Kolmogorov complexity.
- Other approaches are possible.

Summing Across Environments

- Let E_1 be the set of input states that are generated by environment 1.
- $K(E_1)$ is the Kolmogorov complexity of the input states that are generated by environment E_1.
- When environments are similar $K(E_1+E_2)$ will be approximately half of $K(E_1)+K(E_2)$.
- When environments are different $K(E_1+E_2)$ will be approximately the same as $K(E_1)+K(E_2)$.

Summing Across Environments

- When adding up PI-2 from different environments multiply by :

$$\frac{K(E_1 + E_2)}{K(E_1) + K(E_2)}$$

- In the case of similar environments, the summed PI-2s will be approximately the same as the PI-2 for an individual environment.

PI-3

$$\frac{K(E_1 + E_2)}{K(E_1) + K(E_2)} * (PI_{E_1} + PI_{E_2})$$

Illustration

1	1	1	1
0	0	0	0
1	1	1	1
0	0	0	0
1	1	1	1

***E_1* Input Space**

1	1	1	1
0	0	0	0
1	1	1	**0**
0	0	0	0
1	1	1	1

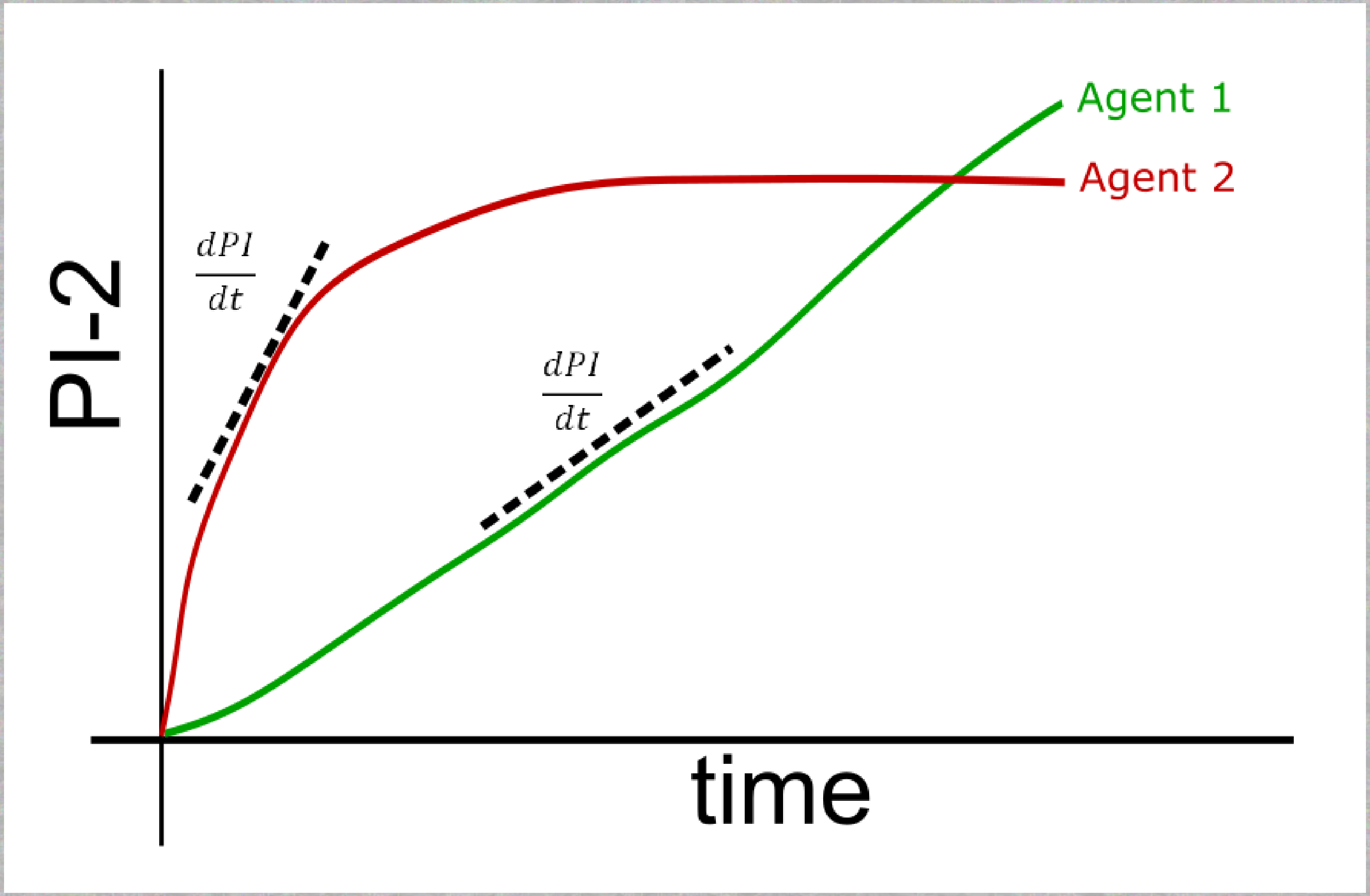
***E_2* Input Space**

Fluid Intelligence

- PI-3 is a measure of *crystallized* intelligence across environments.
- To estimate fluid intelligence we could put the agent in a new *different* environment and measure the average or maximum rate of change of PI-2.

$$Fluid\ Intelligence = \frac{dPI}{dt}$$

Illustration



Approximations of PI

- With complex environments it might not be feasible to calculate PI-2.
- Can only estimate PI-3 by exposing system to a variety of different environments – this cannot be fully calculated.
- An important part of the measurement of PI will be the development of ways of estimating it from limited data sets.

Prediction is not Everything!

- Perfect prediction gives you God-like *omniscience*.
- Can predict every consequence of your action; know everything that is going to happen in your environment.
- Cognitive systems also have:
  – Goals that are often selected using emotions (Damasio 1994).
  – Planning systems that use predictive knowledge to figure out how to achieve their goals.
  – Actuators and control systems.

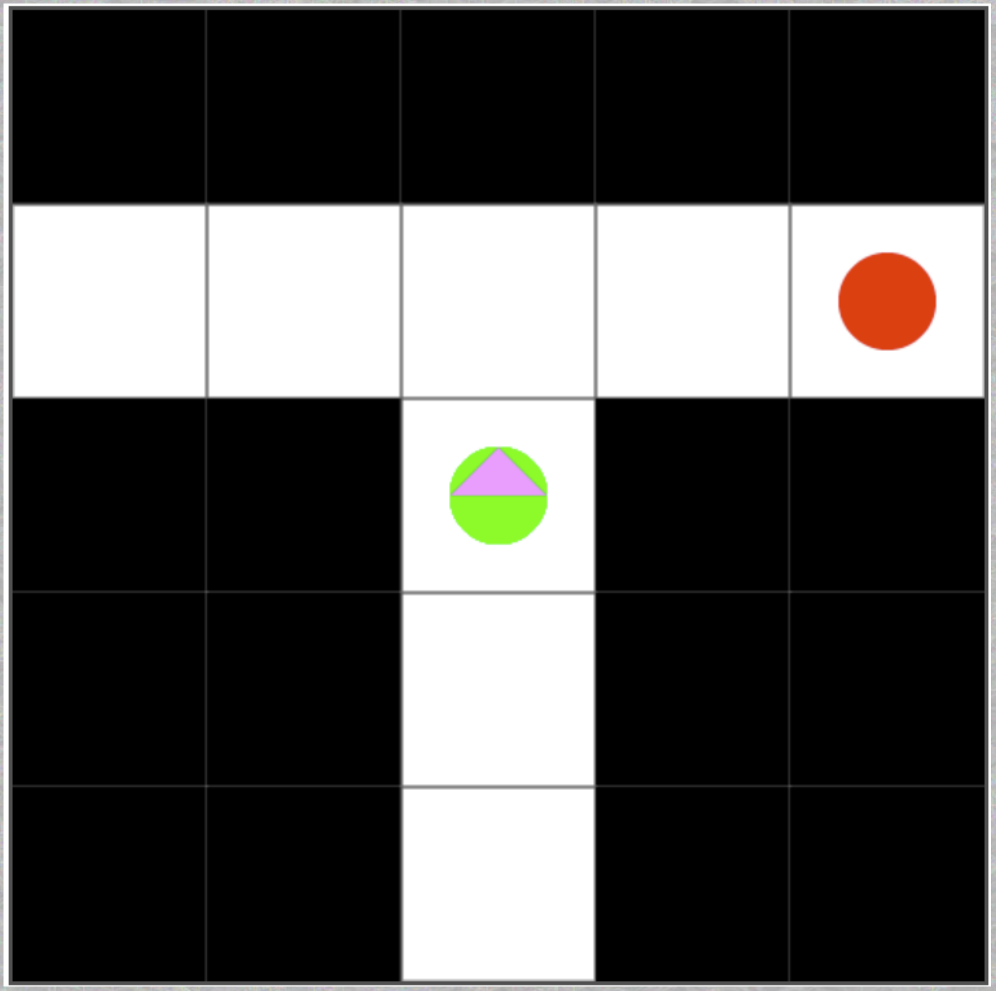
Consequences

- No maximum p-intelligence.
- Many varieties of intelligence: Two systems can exhibit similar values of p-intelligence about very different things.
- Different agents can have very different mappings between environment states and input states. So the intelligence of different agents can be about different features of a single physical environment.

Next Steps

- Test PI measures and compare with standard measures of intelligence and success:
  - Simple maze experiments.
  - Machine learning.
  - Animal Olympics.
- Similar to the approach used to validate IQ and *g*.

Simple Maze Environment



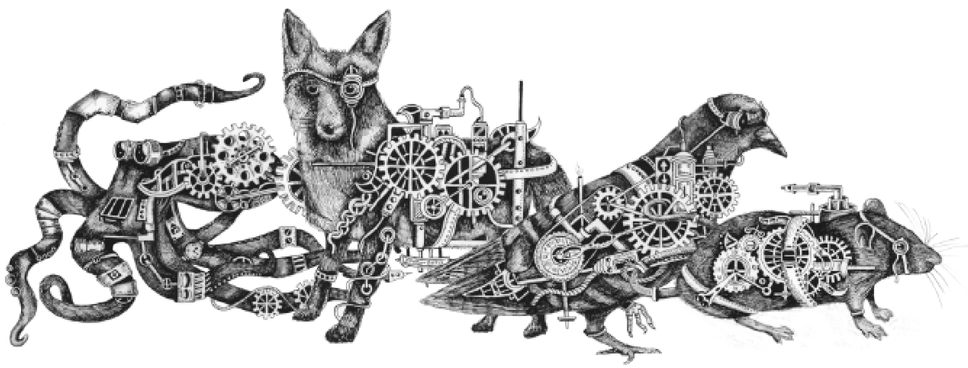
Machine Learning

- Will evaluate PI on standard machine learning benchmarks.
- Measure PI at different training stages.
- Use PI to compare algorithms and explore whether this correlates with success.

Animal-AI Testbed

The Animal-AI Testbed

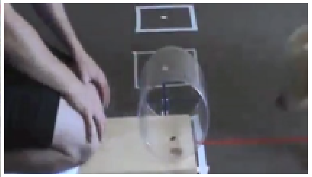
Tasks for AI - inspired by Animal Cognition. Learning common-sense reasoning from sensory inputs. **Play, Watch** and **Discover** the testbed.



These tasks involve simple common-sense reasoning and **are easy enough that a child should be able to solve them**[1]. Many have been passed by animals (including shrimp!). However, these tasks are **still challenging for AI**, and will need to be solved to create more general forms of artificial intelligence.

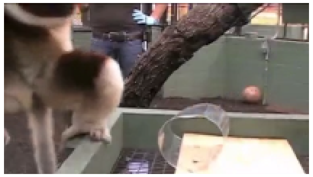
Example task - Transparent Cylinder

Animal Pass



Agent Pass

Animal Fail



Agent Fail

Animal videos and cylinder experiment discussion: MacLean et al. *The evolution of self-control* PNAS (2014)

Applications: Theory of Intelligence

- P-intelligence is theory of intelligence, not just a behavioural measure based on some hunches.
- p-intelligence could be *g* (the *g*-loading of prediction tests might be 1).
- Better way of explaining the performance of different cognitive systems in different environments
- Replace talk of intelligence with discussion of prediction and other attributes of cognitive systems.

Applications: Measurement of Intelligence

- PI-2 and PI-3 could be much better ways of measuring intelligence across diverse systems.
- Measure intelligence directly from internal states or from external behaviour.
- Measure fluid and crystallized intelligence and changes in both over time.

Applications: Philosophy of Science

- Popper (2002) stresses the importance of developing scientific theories that can generate large numbers of falsifiable predictions.
- Can measure of the p-intelligence of scientific theories and rank and compare them.
- Could be used as part of computational scientific discovery.

Applications: Singularity

- PI can be used to evolve AIs that are more intelligent within a particular environment.
- These will take over the world and destroy humanity.

CONCLUSION

Conclusion

- Want to compare the intelligence of humans, animals and machines.
- Anthropocentric measures of intelligence cannot be generalized.
- Current universal measures have limitations.

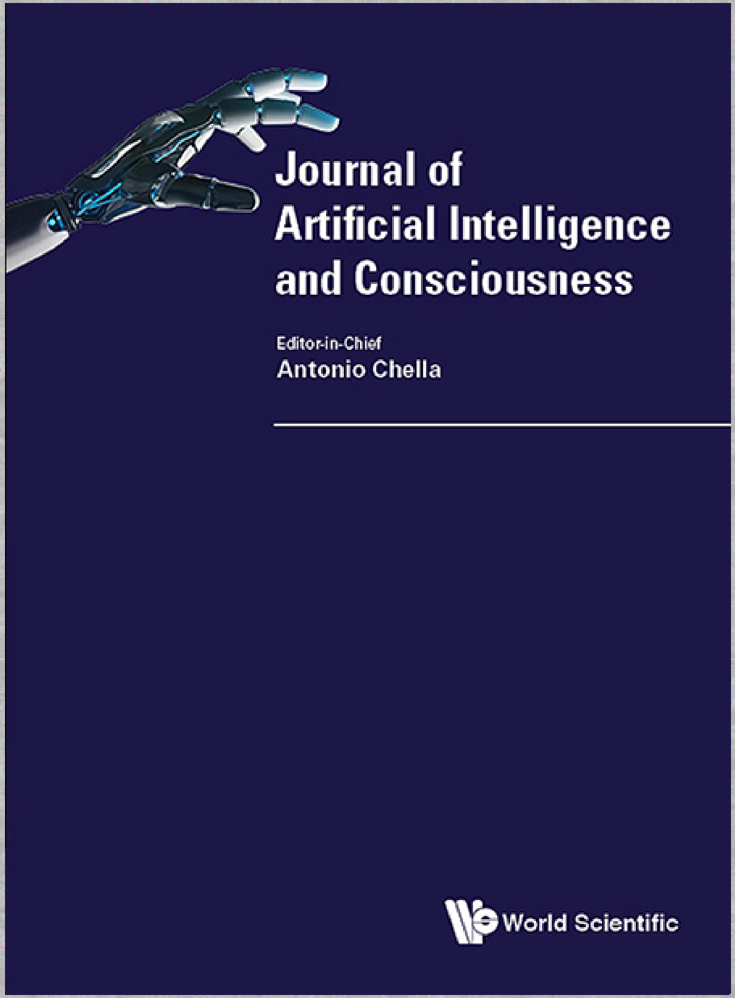
Conclusion

- Good reasons for thinking that intelligence is closely linked to a system's ability to make accurate predictions.
- Have outlined an algorithm for measuring predictive intelligence.
- PI-2 can be used to compare the intelligence of agents within a single environment.
- PI-3 can be used to estimate the total intelligence of an agent.

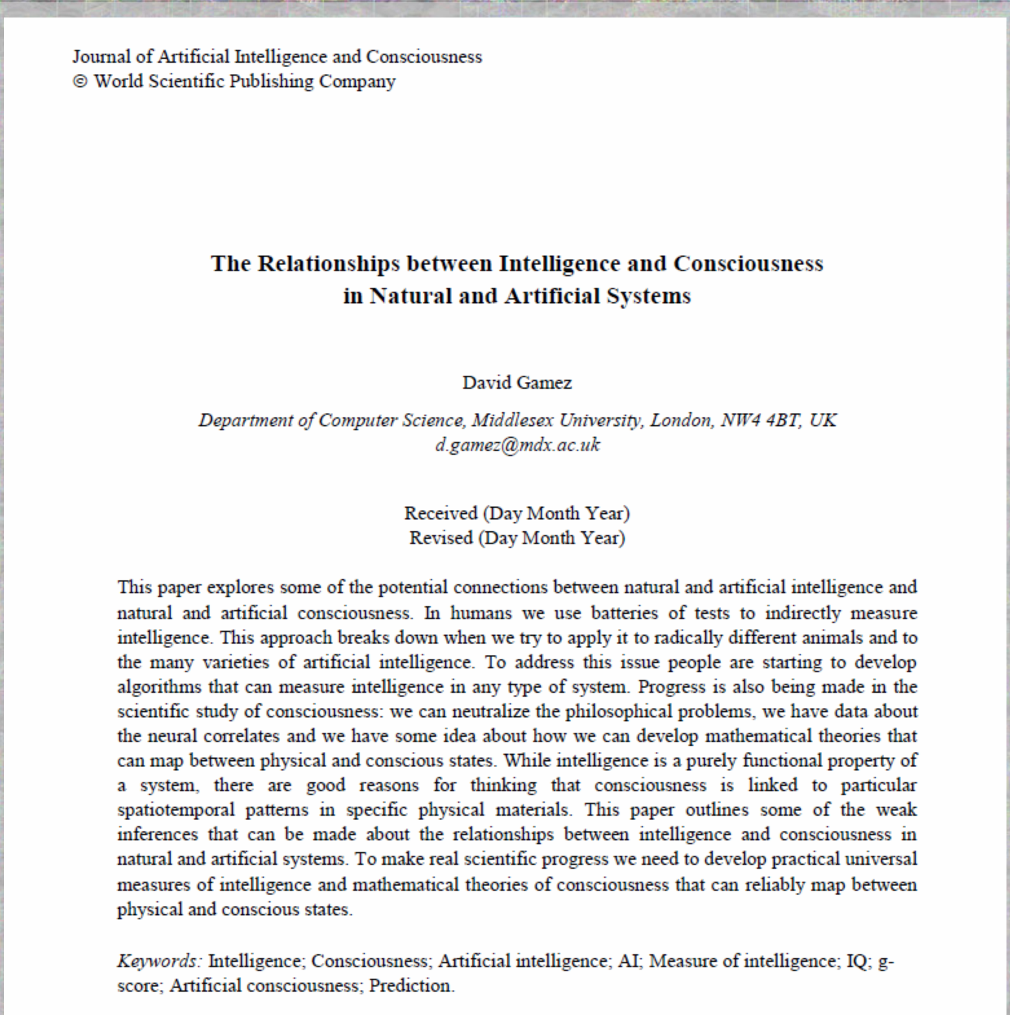
Conclusion

- Early stage research.
- Currently working on experiments and a paper.
- Suggestions and constructive criticism welcome!

Journal of Artificial Intelligence and Consciousness



Intelligence and Consciousness



Journal of Artificial Intelligence and Consciousness

The Relationships between Intelligence and Consciousness in Natural and Artificial Systems

David Gamez

Department of Computer Science, Middlesex University, London, NW4 4BT, UK
d.gamez@mdx.ac.uk

Received (Day Month Year)
Revised (Day Month Year)

This paper explores some of the potential connections between natural and artificial intelligence and natural and artificial consciousness. In humans we use batteries of tests to indirectly measure intelligence. This approach breaks down when we try to apply it to radically different animals and to the many varieties of artificial intelligence. To address this issue people are starting to develop algorithms that can measure intelligence in any type of system. Progress is also being made in the scientific study of consciousness: we can neutralize the philosophical problems, we have data about the neural correlates and we have some idea about how we can develop mathematical theories that can map between physical and conscious states. While intelligence is a purely functional property of a system, there are good reasons for thinking that consciousness is linked to particular spatiotemporal patterns in specific physical materials. This paper outlines some of the weak inferences that can be made about the relationships between intelligence and consciousness in natural and artificial systems. To make real scientific progress we need to develop practical universal measures of intelligence and mathematical theories of consciousness that can reliably map between physical and conscious states.

Keywords: Intelligence; Consciousness; Artificial intelligence; AI; Measure of intelligence; IQ; g-score; Artificial consciousness; Prediction.

References

Cattell, R. B. (1971). *Abilities: Their Structure, Growth, and Action.* Boston: Houghton Mifflin.

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* Oxford: Oxford University Press.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: G.P. Putnam.

Gardner, H. (2006). *Multiple Intelligences: New Horizons.* New York: Basic Books.

Haier, R. J. (2017). *The Neuroscience of Intelligence.* Cambridge: Cambridge University Press.

Hernández-Orallo, J. and Dowe, D. L. (2010). Measuring universal intelligence: Towards an anytime intelligence test. *Artificial Intelligence* 174: 1508-39.

Legg, S. and Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines* 17: 391-444.

Popper, K. R. (2002). *The Logic of Scientific Discovery.* London: Routledge.

Warwick, K. (2000). *QI: The Quest for Intelligence.* London: Piatkus.

Questions?

- Website: www.davidgamez.eu.