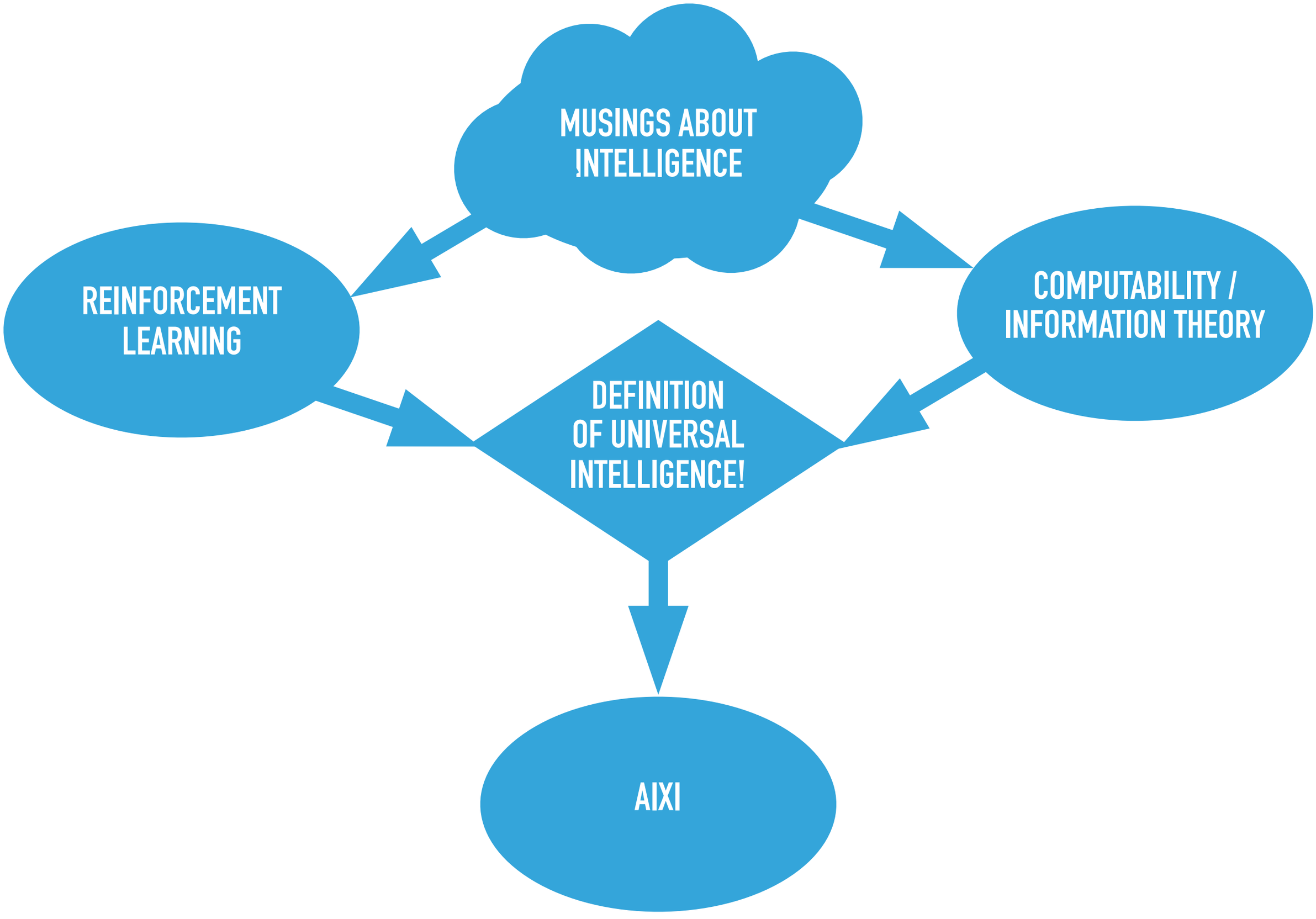
UNIVERSAL INTELLIGENCE

# A DEFINITION

RESEARCH BY HUTTER AND LEGG
PRESENTED BY COLIN MCDONNELL

# FORMAT OF TALK



property of an agent

that interacts with its environment

to successfully achieve goals

across a wide range of environments

# 2, 4, 6, 8, . . .

$$F(N) = 2N-1 + (N-1)(N-2)(N-3)(N-4)$$

property of an agent

that interacts with its environment

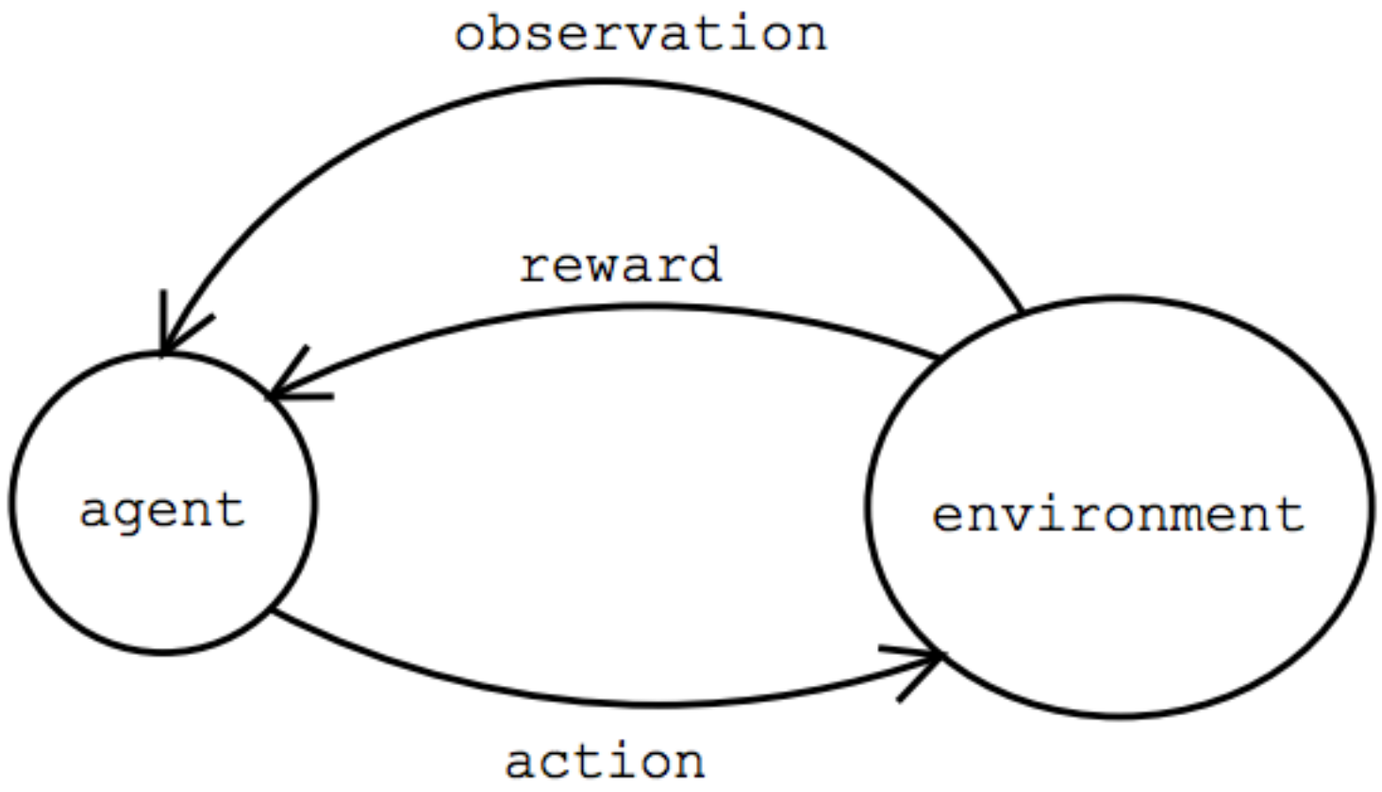
to successfully achieve goals

across a wide range of environments

with a preference for simpler solutions

# REINFORCEMENT LEARNING

# REINFORCEMENT LEARNING




A - action space

P - perception space

perception := observation-reward pair

# THE AGENT

A - action space
P - perception space

modeled as a probability distribution

$$\pi(a_k|o_1 r_1 a_1 \ldots o_k r_k)$$

very general

# THE ENVIRONMENT/REWARD

A - action space
P - perception space

modeled as a probability distribution

$$\mu(o_k r_k | o_1 r_1 a_1 o_2 r_2 a_2 \ldots o_{k-1} r_{k-1} a_{k-1})$$

very general

example: guessing coin toss

# SUCCESS VS REWARD

A - action space
P - perception space

time scale dependence

discounting methods

answer: summability constraint

$$V_\mu^\pi := \mathbf{E}\left(\sum_{i=1}^{\infty} r_i\right) \leq 1$$

problems?

# COMPUTABILITY AND INFORMATION THEORY

measure of information

Turing machines

Kolmogorov complexity

# DEFINITION OF UNIVERSAL INTELLIGENCE

INTELLIGENCE

Occam's razor

$$\Upsilon(\pi) := \sum_{\mu \in E} 2^{-K(\mu)} V_\mu^\pi$$

agent

environment

success

env space

RECALL

$$V_\mu^\pi := \mathbf{E}\left(\sum_{i=1}^{\infty} r_i\right) \leq 1$$

# ANALYZING DIFFERENT AGENTS

random

specialized

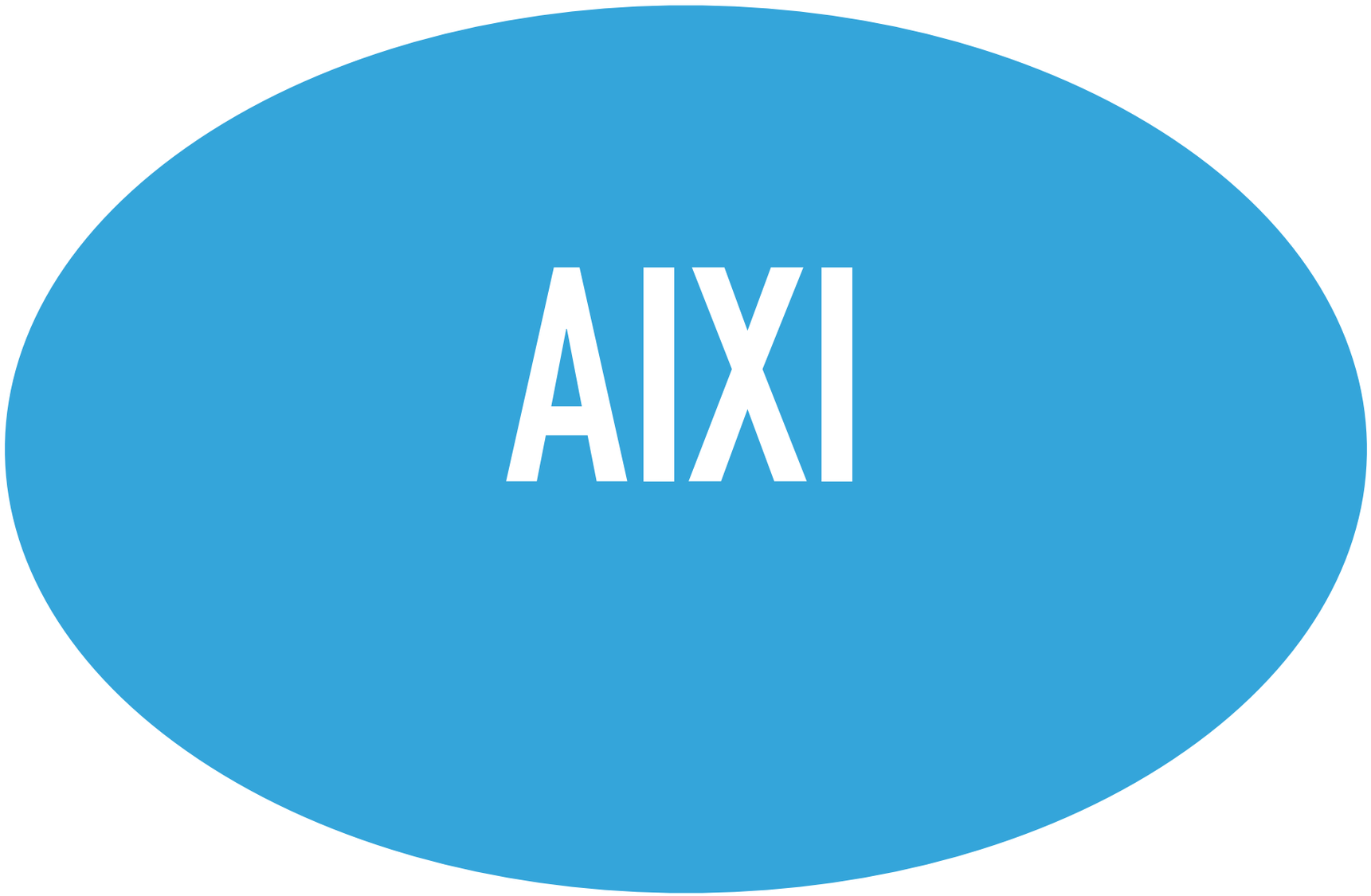
general but simple

simple with history

forward looking

superintelligence

$$\Upsilon(\pi) \ := \sum_{\mu \in E} 2^{-K(\mu)} \, V_\mu^\pi$$


AIXI

## AIXI

for the set of observations and actions thus far

calculate sum of expected future rewards

for all consistent environmental hypotheses

weighted by the probability of that hypothesis

$$\arg\max_{a_t}\sum_{o_t r_t}\ldots\max_{a_m}\sum_{o_m r_m}[r_t+\ldots+r_m]\sum_{q:\,U(q,a_1\ldots a_m)=o_1 r_1\ldots o_m r_m}2^{-\mathrm{length}(q)}$$

## PROBLEMS

uncomputable

Kolmogorov

sum over hypothesis space

limitations of reinforcement learning model

rewards are a property of environment?

resource limitations/scarcity?

Cartesian barrier

$$\arg\max_{a_t}\sum_{o_t r_t}\ldots\max_{a_m}\sum_{o_m r_m}[r_t+\ldots+r_m]\sum_{q:\,U(q,a_1\ldots a_m)=o_1 r_1\ldots o_m r_m}2^{-\mathrm{length}(q)}$$

# QUESTIONS?