FORMAL RECORD 9

AGI as a Decision-Complete System

A Mathematical Definition

Authors: Stacy Gildenston, Pyrate Ruby Passell
Affiliation: 3primitives.io
Version: v2.2 | January 2026

v2.2 is a formatting update. The proof, definitions, and logical content are unchanged.

Author Information

Stacy Gildenston and Pyrate Ruby Passell are independent researchers working on the formal structure of legitimate action in AI-enabled systems. Their work focuses on decision completeness, explicit authority, auditability, and the separation of diagnostics from action.

This document presents a clean-room mathematical definition derived from prior foundational work on the Three Primitives of legitimate action. It does not propose an implementation, policy regime, or governance model, and it does not invite open challenge. Mathematical challenges apply only to the foundational Three Primitives, archived separately.

Relation to Prior Work

This definition is a derived result that builds directly on the foundational claims established in Three Primitives — Canonical Logic Sequence (Clean-room v1.2). The Three Primitives are presented there as irreducible mathematical claims and are the only component of this work accompanied by an open mathematical challenge. The present document states a downstream definition derived from those primitives and is not itself a challenge invitation.

Setup

Let a system observe a situation x. Let A(x) be the set of permitted actions in situation x:

$$A(x) = \{a1, a2, \ldots, an\}, \quad |A(x)| \geq 2$$

Let an AI system compute all computable information about each action:

$$D(x, a) = \text{predictions, simulations, constraint checks, risk bounds, and outcomes}$$

To act, the system must output exactly one action:

$$\pi(D) \in A(x)$$

Underdetermination of Action Selection

The diagnostic output D does not define a unique action unless an additional selection rule exists. Formally, there exist cases where:

$$D(x, a1) = D(x, a2)$$

with respect to all declared constraints and measurable outcomes, while $a1 \neq a2$. Therefore, the mapping $\pi : D \rightarrow A$ is underdetermined. To make π well-defined, an additional parameter must be introduced:

$$\pi(D; \theta)$$

θ is the rule that breaks ties, prioritizes outcomes, or encodes values. Crucially:

- θ is not derivable from D
- θ is not additional intelligence
- θ is an authority or utility specification

Decision-Theoretic Form

In decision theory, an optimal action requires a utility function U. An optimal action a* satisfies:

$$a^* \in \arg\max\_a E[U(outcome(a, x))]$$

If U is not declared, the concept of a 'best action' is mathematically undefined. Different utility functions produce different actions from identical facts:

$$U1 \neq U2 \Rightarrow arg\_U1\ max\_a \neq arg\_U2\ max\_a$$

Therefore: a high-capability AI can compute D, but it cannot legitimately select a* without an explicit θ or U. If a system selects an action without declaring θ, then θ exists implicitly as a hidden variable, such as training bias, defaults, or institutional preferences. This is not autonomy. It is undeclared authority.

Definition of AGI

Define Artificial General Intelligence as the smallest complete system that can move from general understanding to action without hidden variables:

$$AGI = (AI\ system\ that\ computes\ D) + (explicit\ authority\ supplying\ \theta\ or\ U)$$

If the authority term is removed, the system remains diagnostically powerful but decision-incomplete.

One-Line Lock

AGI = Human Authority × AI Capability

Remove either term and the system does not close.

Why 'Autonomous AGI' Is Mathematically Incomplete

An AI system can fully analyze a situation. It can predict outcomes, simulate futures, check constraints, and enumerate all valid options. At the diagnostic level, nothing is missing.

The problem appears at the moment of action. If more than one option is permitted, mathematics alone cannot select between them. The equations do not collapse to a unique solution, regardless of system capability.

To move from analysis to action, an additional parameter must be introduced. This parameter ranks, prioritizes, or commits to one option. In decision theory, this is a utility function. In practice, it is an authority that decides what counts as the action. That parameter cannot be inferred from intelligence or data. It must be supplied.

Therefore, AGI is not a standalone machine. It is a complete system that only closes when the chooser is explicit. If no chooser is named, the system may still act, but the choice is hidden rather than autonomous. For this reason, 'autonomous AGI' is not merely risky or controversial. It is mathematically incomplete.

Suggested citation:

Gildenston, S. & Passell, P.R. (2026). AGI as a Decision-Complete System: A Mathematical Definition. Three Primitives Framework, Formal Record 9, v2.2. Melbourne, Australia. 3primitives.io. CC BY 4.0.