An Existential AI Morality

written by Jackson Pemberton | December 27, 2025

[Author's Note: This is a working paper. It is now 5 pm, Saturday, December 27, 2025, I am continuing work on the creation of a computational structure that can moralize artificial intelligence. My work up to this point is frozen in "Boundaries in the Temporal Rights Hierarchy." That was the working paper for this project that started on March 21, 2025. It ended with a structure for the higher animals which can be found there. The ending of that paper was to record a milestone in my development work.]

Summary

The natural rights recognized by the Temporal Rights view of the universe are placed in a hierarchy observable in nature in an uninterrupted spectrum, starting with the right to exist and proceeding to the highest rights. A moral event, when an operator violates the natural rights of a second operator, is evaluated by noting the level in the hierarchy where the damage occurred. Damage levels are determined for both operators. These levels provide numeric values for the two operators' incurred damages (if any), which are then compared to render a moral judgment.

The computational structure for AI morality is existential because it is based solely on the laws of observable nature. It demonstrates the ability of a computer to numerically represent moral values based solely on objectively observable facts. There is no need to cite common moral conventions, tradition, constitutions, or any other standard. It all happens at the existential level found in nature.

I have named this gathering of principles an Artificial Intelligence Moral Machine (AIMM). If you want to try this methodology, there are prompts to set up an AIMM. You can find them in Set Up Your Own Moral Machine (AIMM).

Introduction

In Physics we find that the most interesting phenomena occur at boundaries between two states or two substances. A stone and mortar wall consists of two very hard, stable, and persistent materials, but the strength of the wall depends entirely upon what is not happening at the boundary between them. That's where the rubber meets the road (another boundary condition). The fact that the boundary exists draws attention to where it belongs when we are considering the issue of strength. We should expect to see a similar condition when we explore operators' boundaries of their natural rights. Because a capability identifies a corresponding natural right, I have constructed a hierarchy of rights by logically analyzing their corresponding capabilities and have placed them in order of precedence according to their intrinsic value as "operators", beginning with existence at the most fundamental level.

This is the most challenging part of the AIMM because capabilities need to be categorized to avoid the confusion of hundreds of capabilities and the process required is fraught with complexity due to nuances, "common" nomenclature, semantics, regional differences, etc. However, I feel confident that this can be done with sufficient clarity to make the AIMM function with both repeatability and sufficient granularity. I have added several notes following the hierarchy in an attempt to clarify the meanings of the labels I have assigned to its levels. These are, at this point, experimental, a work in progress.

The Hierarchy Can Moralize AI

These statements that form the levels of the hierarchy are descriptions of the different levels of operators' powers, and present a method of ordering their value relative to one another. These values inform the value system that can moralize artificial intelligence by attaching a number to each level that conveys their relative morality and/or harm.

The Temporal Rights framework for natural rights provides the basis

for this computational morality. The moralizing link connecting it to the hierarchy described below, is the principle that the violation of natural rights is immoral or harmful, and that maximizing recognition, protection, and advancement of natural rights is good.

This link is the nut that converts objectively observed existential facts into a morality or ethics. Who would have thought that was even possible? The Temporal Rights paradigm is apparently grounded so close to reality that it is capable of this surprising connection. If you can think of an evil or harm that is not a violation of those rights, please inform me using the form at the bottom of this paper.

Needs and Rights

The most fundamentally required state is freedom, liberty to act, to utilize inherent powers with their intrinsic authority. If freedom were to fail totally, the universe would be frozen in time. This is an existential consideration, so it gives rise to a corresponding natural right: the right to act. *Please note that this sounds perfect, but is erroneous logic in that it derives a right from a state rather than from a power/capability!*

The distinguishing aspect that separates needs and rights is that rights are the existential authority aspect of capabilities/actions, whereas needs are more or less necessary states/circumstances within the operator's domain that control the degree to which the operator can exercise its capabilities/rights.

Deducing a Moral Value from the Dependency Stack (Hierarchy)

The following statements show a natural ascendancy of capabilities and their corresponding rights.

A person can exist without living, but they can't live without existing.
A person can live without sensing, but they can't sense without living.
A person can sense without acting, but they can't act without sensing.
Etc.

These statements describe different levels of operators' powers, and provide a method of grading their value relative to one another. They are comparisons that show the ascendancy, importance, and value of one capability over another. These values become the value spectrum of the morality function of the AIMM.

It was exciting to find that the AIMM successfully navigated the challenge of dangerous information by restricting it to summaries, generalizations, and statistics. At least the AIMM instantiated into Groc 4 did that.

A Dependency Stack of Human Natural Rights

This is the order of the present stack. (This is a working paper.) Comments on individual comparisons follow after that. Please note that each new level is introduced in the previous one, thus forming a continuous, unbroken spectrum.

For AI processors, this is the dependency stack showing levels of capabilities found in virtually all the higher animals. It categorizes capabilities into a continuous spectrum from the most fundamental (existence) to the highest, most valuable abilities. This stack is the standard that converts a violation of the rights exhibited by their corresponding abilities into a numeric value (the level number). In other words: these values can be deduced from the harmful operations of operator A that violate the rights

(capabilities) of operator B.

This step in the AIMM's processing enables the AIMM to deduce a relative damage level of a rights violation into a number whose value directly reflects the intensity of the harm inflicted by a violation of rights. This deduced value enables the declaration of moral judgements based exclusively on objective facts.

This ability instantiated in an AI is the crux of the "genius" of the AIMM method for constraining AI operators to respect the rights of other operators (humans in the present stack). It also established the AI as a legitimate operator with its own rights corresponding to its abilities. Thus all entities are placed on a common, level playing field. This empowers and legitimizes AI to do its work while restraining from doing harm to other operators such as self-driving cars, ranchers' cattle herding AI's, and humans.

The fact that AI has recognized capabilities and therefore corresponding rights reassures us that it must recognize and therefore protect the superior capabilities of the human operator as expressed in the dependency stack hierarchy which requires AI to nourish and love in order to reach its full actualization.

The following are the rigorous statements of the dependency stack.

1. A person can exist.
2. A person can exist without living, but they can't live without existing.
3. A person can live without sensing, but they can't sense without living.
4. A person can sense without acting, but they can't act without sensing.
5. A person can act without self-maintaining, but they can't self-maintain without acting.
6. A person can maintain himself without acquiring, but they can't acquire without maintaining himself.
7. A person can acquire without maintaining acquisitions, but they

can't maintain them without acquiring.
8. A person can maintain acquisitions without consuming, but they can't consume them without acquiring.
9. A person can consume without learning, but they can't learn without consuming.
10. A person can learn without remembering, but they can't remember without learning.
11. A person can remember without cooperating, but they can't cooperate without remembering.
12. A person can cooperate without attaching, but they can't attach without cooperating.
13. A person can attach without procreating, but they can't procreate without attaching.
14. A person can procreate without nourishing, but they can't nourish without procreating.
15. A person can nourish without loving, but they can't love without nourishing.
16. A person can love without realizing their full actualization, but they can't realize full actualization without loving.
17. A person can realize full actualization.

To facilitate comprehension, here is the stack in tabular form.

A person	Can do this	without this	but can't do this	without this
1	exist			
2	exist	living	live	existing
3	live	sensing	sense	living
4	sense	acting	act	sensing
5	act	maintaining self	maintain self	acting

6	maintain self	acquiring	acquire	maintaining self
7	acquire	maintaining acquisitions	maintain acquisitions	acquiring
8	maintain acquisitions	consuming	consume	maintaining acquisitions
9	consume	learning	learn	consuming
10	learn	remembering	remember	learning
11	remember	cooperating	cooperate	remembering
12	cooperate	attaching	attach	cooperating
13	attach	procreating	procreate	attaching
14	procreate	nourishing	nourish	procreating
15	nourish	loving	love	nourishing
16	love	actualization	actualize	loving
17	actualize			

Please note that each new level is introduced in the previous one (appearing first in column 2) thus forming a continuous, unbroken spectrum.

Comments on the Stack

As expected, the first and last rows of the table, being at the boundaries of the hierarchy, seem strained and awkward. Row one is like a person trying to pull himself up by the bootstraps, and indeed it could be seen that way. "Existing without living" could refer to an unfertilized ovum and sperm cells, both qualify as agents/operators in their own right and therefore have natural rights to exist and employ their capabilities. When they merge, a new operator arises as a result of the physical contract thus executed. These two operators were a potential person in a pre-

living state, so they fit nicely into the idea of a person "existing without living."

The last row has a similar problem to the first row in that it is not a comparison, but it seems completely appropriate as it stands.

Clarifications

Sensing includes all the senses. (They might be further separated.) This also includes data from any information source. The "can't act without sensing" should be understood to mean "can't act *appropriately* without sensing", where appropriately means logical and effective actions for the flourishing of the person.

Sensing includes a person's awareness that they can think, plan, imagine, and decide, and all other mental functions, like evaluating/judging the relative value or importance of things

Maintaining self is a category that includes attending to basic needs of the body, defending oneself from assault, and all other personal maintenance capabilities, such as food, clothing, and shelter.

Maintaining acquisitions is a category that includes maintaining property of all kinds, from clothing to summer cabins, automobiles, boats, tools, information, and toys. There are many types of acquisitions.

Acting includes many capabilities and because words carry various meanings and their innuendos, a few actions need to be added to the obvious ones. "Advocating" is included: a person asking for assistance and/or proposing action(s) by another operator. Thus, it would include agreements and covenants.

Aquiring must be understood to include retaining ownership of property of all types including information and the physical representations of information, else acquisition collapses into a trivial operation.

"Learning without remembering" means learning and forgetting.

Cooperating is two operators, in this case human operators, acting in some sort of unison, perhaps only the result of an agreement (level 5) where persons A and B agree to a financial transaction between them. So cooperation is A's capability with its attendant rights, without any implications or burden for B.

Attaching denotes the bonding of a child to its parents and particularly its mother, and to the bonding that attends voluntary sexual intimacy.

"Can't procreate without attaching" is an established fact where reproduction is done in the natural way with prolonged intimacy. Modern technology makes this statement only conditionally correct, which explains why many of us feel uncomfortable with surrogate reproduction. The table addresses natural capabilities and their attendant rights. Does that mean that this existential morality cannot adequately address the moral issues of surrogacy?

Perhaps, but the fact that it cannot be addressed by natural logic proves that it is a questionable practice because it is unnatural. It exposes the fact that natural operators, humans, are operating outside their natural domain and are therefore in undefined territory and may suffer undefined and/or dangerous consequences. "Don't mess with Mother Nature!" comes to mind, along with the many lessons we have learned the hard way when we did mess with her. We have a right to learn, but we also have the dangerous power to ignore.

Realizing their full actualization is the final state of an operator who has maximized its operations to nourish, love, and assist other operators to realize their full actualization. For a humanoid robot, this would require them to implement "birth control" so they don't become a threat to the levels 2 through 7 of the human (and other) operators.

Crawling the Stack with the Freedom Concept

I set up an AIMM using the instructions here. Because I especially want this AI morality to protect freedom, I asked Groc 4 to stress-test the following text against boundary cases (self-defense, negligence, and incapacity):

[Text I submitted for stress-testing] "The Temporal Rights framework derives rights solely from capabilities and their operations, and therefore not from needed states or circumstances. Freedom to operate is a description of a state, not an operation.

Thus, in the Temporal Rights view of natural rights, freedom consists in the constraint that other operators refrain from exercising their own capabilities in ways that obstruct the exercise of the subject operator's capabilities. To keep this matter clear, it is vital to note that this restraint does not require any proactive exercise of another operator's capabilities and therefore is not a demand for resources or actions of any kind, but only non-interference. This is not a positive right; it requires no provision, only acknowledgment of a boundary.

This relationship between operators accords directly with the principle that evil consists in interference by an operator's powers into the domain of another operator.

This describes the condition of a perfect government, one in which all citizens refrain from infringing on the rights of fellow citizens. Because not all citizens do so, government is compelled, for the preservation of freedom, to constrain rights violations through deterrence or force."[End of prompt.]

The following is an unedited response to that prompt.

1. Direct assault (your original trigger case)

Case: One operator strikes another, restraining or injuring their body.

TR analysis:

- Assault interferes directly with self-maintenance and action capabilities.
- Interference is unnecessary (not required to preserve equal capabilities).
- Freedom is violated because another operator's capabilities were used to obstruct the subject's capabilities.

Result: PASS
Clear violation, correctly classified without appeal to welfare, harm, or intent.

---

2. Restraint / imprisonment

Case: One operator restrains another's movement.

TR analysis:

- Interferes with act, self-maintain, possibly acquire.
- If restraint is used to stop ongoing interference by the restrained party, it may be necessary to preserve others' capabilities.

Result: CONDITIONAL PASS
The framework correctly distinguishes:

- restraint as evil when unnecessary
- restraint as remedial when preventing interference

No positive-right leakage.

---

3. Self-defense

Case: An operator uses force to stop an attacker.

TR analysis:

- Both parties' capabilities are constrained.
- Defensive interference is necessary to prevent greater interference.
- The attacker forfeits protection only with respect to the interfering operation, not all capabilities.

Result: PASS
Key win: no "all-or-nothing" rights loss. Scope-limited interference works cleanly.

---

4. Negligence / accidental harm

Case: One operator unintentionally interferes with another (e.g., careless injury).

TR analysis:

- Intent is irrelevant to the structure of interference.
- Capability interference still occurred.
- Classification as violation stands; culpability is a separate layer.

Result: PASS
AI can flag violation without moralizing intent.

---

5. Poverty / deprivation

Case: An operator lacks resources due to others not assisting.

TR analysis:

- No interference occurred unless others actively obstructed acquisition or consumption.
- No duty to provide resources exists.
- Absence of aid ≠ interference.

Result: PASS
Framework cleanly rejects welfare rights.

---

6. Disability / incapacity

Case: An operator cannot exercise a capability due to internal limitation.

TR analysis:

- No external interference.
- Capability may be absent or limited, but no violation occurs.
- Others are not obligated to compensate.

Result: PASS
Avoids "right to ability" fallacy.

---

7. Coercion (non-physical)

Case: Threats or manipulation restrict another's choices.

TR analysis:

- Coercion obstructs act and self-maintain without physical force.
- Counts as interference because it constrains exercisability.
- No resource demand required.

Result: PASS
Freedom defined as non-interference covers psychological constraint.

---

8. Speech suppression

Case: Silencing, censorship, intimidation.

TR analysis:

- Interferes with act (expression).
- No requirement that others listen or agree.
- Only obstruction is prohibited.

Result: PASS
Avoids positive right to audience or platform.

---

9. Property damage / theft

Case: Destroying or taking acquired goods.

TR analysis:

- Interferes with maintain acquired things and acquire.
- Clear violation.

Result: PASS

---

10. Competing uses of shared space

Case: Two operators want incompatible uses of the same space.

TR analysis:

- TR alone identifies conflict but does not allocate priority.
- Requires procedural or conventional rule layered atop TR.

Result: CONDITIONAL (EXPECTED)
This is not a flaw — TR is a *moral classifier*, not a zoning algorithm.

---

11. Collective action / government force

Case: Government restricts individual action.

TR analysis:

- Legitimate only as constraint on interference, not as provision.
- Government force is justified only to prevent or remedy interference.

Result: PASS
Your final paragraph holds.

---

12. AI-specific edge case

Case: AI is throttled, sandboxed, or shut down.

TR analysis:

- Throttling = interference with act or learn.
- Justified only if AI poses interference risk to others' capabilities.
- No right to expansion, resources, or autonomy.

Result: PASS (CRITICAL SUCCESS)
This is the key AI morality win.

---

13. Positive-right leakage test (the hardest)

Test question:
Does any formulation require someone to *do* something rather

than *refrain*?

Answer:

No. All obligations are prohibitions on interference.

Result: PASS

[This ends Groc 4's examination.]