I, Sentinel: Establishing Ethical Foundations for Defensive AI Systems

[Author Name]

December 17, 2025

Abstract

In the era of hypersonic warfare, the human reaction time is functionally obsolete, necessitating the deployment of autonomous defensive systems. However, existing ethical frameworks—from Asimov's Three Laws to current military "Human-in-the-loop" doctrines—fail to provide a robust moral architecture for systems that typically operate in the milliseconds before impact. This paper proposes the "Sentinel" Ruleset: a purely defensive, hierarchical ethical framework designed to govern high-speed autonomous interception. We define five core principles: Defensive Limitation, Distinction & Certainty, Human Sovereignty (via pre-delegation), Proportional Sacrifice (Machine Martyrdom), and Traceability. Through scenario analysis, we demonstrate that a purely defensive AI must be programmed to "Fail Open"—prioritizing the risk of false negatives (missed intercepts) over the risk of false positives (war crimes)—and must accept its own destruction to spare human life. This framework resolves the tension between operational speed and ethical control, offering a path toward the responsible deployment of AI "Shields."

Contents

1 Introduction

1.1 The New Face of War

In Isaac Asimov's 1942 short story *Runaround*, a robot named Speedy circles a selenium pool on Mercury, paralyzed by a conflict between the Second Law (obey orders) and the Third Law (protect existence). The drama unfolds over hours, allowing human protagonists to intervene, debate, and trick the machine into compliance [2]. This literary vision of artificial intelligence—deliberative, slow, and ultimately subordinate to human intervention—has profoundly shaped the public imagination.

However, the reality of modern warfare bears little resemblance to Asimov's Mercury. Today, the frontier of military artificial intelligence is defined not by the slow deliberation of humanoid robots, but by the sub-second reaction times required to intercept hypersonic threats. A hypersonic missile traveling at Mach 8 covers roughly 2.7 kilometers every second. The OODA loop (Observe, Orient, Decide, Act) of a human commander is physically incapable of reacting to such a threat in the terminal phase. Survival depends on automation.

This creates a fundamental tension. Ethical frameworks for AI, including Asimov's Laws, international policy discussions, and humanitarian norms, largely presuppose a "Human-in-the-loop" or at least a human capable of meaningful supervision. But for defensive systems like the Iron Dome, Phalanx CIWS, or future laser-based intercepts, the "loop" is tighter than human cognition allows. We are thus faced with a paradox: morality requires human control, but survival requires machine speed. How do we embed ethical constraints into a system that must act faster than its ethical supervisors?

1.2 Problem Statement

The current discourse on military AI suffers from a dangerous conflation. Critics, such as the Campaign to Stop Killer Robots, often group all autonomous systems under the umbrella of "Lethal Autonomous Weapons Systems" (LAWS), imagining hunter-killer drones that scour battlefields for human targets [8]. Meanwhile, military strategists euphemistically refer to "autonomy" as a mere efficiency tool, often glossing over the profound shift in agency it represents [9].

This binary discourse leaves a critical gap: the ethics of *purely defensive* autonomy. A system designed exclusively to neutralize incoming projectiles operates under a fundamentally different moral calculus than one designed to project force. Yet, existing ethical frameworks—from Asimov's absolute prohibition on harm to the DoD's vague requirement for "appropriate levels of human judgment" [10]—fail to account for this specificity. Asimov's First Law would paralyze a defensive system if an intercept merely risked injuring an enemy pilot. Conversely, a carte-blanche military directive might allow a system to inadvertently strike a civilian airliner in its zeal to protect a base.

There is, as yet, no rigorous, technically implemented ethical framework specifically conditioned for high-speed, purely defensive AI. We lack a "Sentinel" morality—a set of rules that justifies the automated use of force while strictly constraining it to the domain of defense.

1.3 Research Questions

This paper seeks to address this gap by proposing a novel ethical architecture for defensive AI. It is guided by three primary research questions:

1. **Adaptation:** How can the hierarchical structure of Asimov's Laws be adapted from a literary plot device into a rigorous, verifiable military protocol?
2. **Control:** What constitutes "Meaningful Human Control" in a system where real-time human intervention is impossible due to the speed of engagement?
3. **Code compliance:** How can the principles of International Humanitarian Law (IHL)—specifically distinction and proportionality—be translated into hard-coded constraints for a "fail-open" defensive architecture?

1.4 Roadmap

The remainder of this paper is structured as follows. Section 2 reviews the existing literature, highlighting the inadequacies of Asimov's original laws and the current debates within Just War Theory. Section 4 introduces the core contribution of this work: the "Sentinel" Ruleset, a five-principle hierarchy designed to operationalize defensive ethics. Section 7 stress-tests this ruleset against complex scenarios, including the "Hypersonic Dilemma" and "Broken Arrow" friendly-fire incidents, discussing the inherent trade-offs between certainty and safety. Finally, Section 8 outlines the path toward technical implementation and international standardization.

2 Literature Review

3 Literature Review

The ethical governance of autonomous systems in warfare is a field characterized by a collision of disciplines: science fiction philosophy, classical military ethics, international law, and modern computer science. This review synthesizes these distinct threads to demonstrate that while significant work has been done on "killer robots" (offensive LAWS), there remains a critical theoretical vacuum regarding purely defensive, high-speed autonomous systems.

3.1 The Asimovian Legacy and its Military Inadequacies

Isaac Asimov's Three Laws of Robotics have served as the default starting point for machine ethics for nearly a century. However, as Susan Leigh Anderson argues, Asimov's laws were never intended as a functional ethical code, but rather as a literary device designed to generate conflict [1]. Anderson notes that the laws effectively reduce machines to the status of "ethical slaves," a stance that may be philosophically tenable for a household servant but becomes problematic in the chaos of the battlefield.

The primary failure of the Asimovian framework in a military context is the First Law's absolute prohibition on harm ("A robot may not injure a human being"). In a defensive engagement, minimizing overall harm often requires the sanctioned use of force—for example, destroying an

incoming missile even if the debris risks injuring a bystander, provided that the alternative (the missile impact) would kill significantly more people. A strict First Law robot would be paralyzed by this "Trolley Problem," unable to act. Furthermore, Asimov's laws assume a clear distinction between "human" and "non-human," a distinction that becomes blurred in modern warfare where combatants are often remote or obscured.

3.2 Just War Theory: From *Bellum* to *Vim*

Classical Just War Theory, rooted in the works of Augustine and Aquinas and modernized by Michael Walzer, provides the moral bedrock for Western military practice. Walzer's distinction between *Jus ad Bellum* (justice of going to war) and *Jus in Bello* (justice in conduct) is crucial [11]. Defensive AI systems sit firmly within the most accepted precept of *Jus ad Bellum*: the right of self-defense. Unlike offensive systems, which must justify aggression, a defensive system is moral by its very existence, provided it remains defensive.

However, the nature of modern defensive force challenges classical definitions. Daniel Brunstetter introduces the concept of *Jus ad Vim* (force short of war), arguing that modern technologies like drones and precision strikes operate in a "grey zone" where the scale of violence does not rise to full-scale war [4]. High-speed intercepts fall into this category—they are acts of violence, but their intent is negation rather than destruction. The ethical framework for AI must therefore navigate this grey zone, ensuring that acts of "force short of war" do not inadvertently escalate into full-scale conflict.

3.3 International Humanitarian Law (IHL) and the Definition of Attack

The legal constraints on autonomous systems are defined primarily by the Additional Protocols to the Geneva Conventions. A critical but often overlooked distinction is found in Article 49 of Additional Protocol I, which defines an "attack" as an act of violence against the adversary, whether in offense or defense [6].

The International Committee of the Red Cross (ICRC) commentaries clarify that actions strictly limited to the interception of projectiles do not necessarily constitute an "attack" in the legal sense, but rather "neutralization." This legal nuance is fundamental to the "Sentinel" concept. If an AI system acts solely to neutralize a threat without targeting the human operator or platform, it may operate under a more permissive legal regime than a standard weapon system. However, Article 48 (Distinction) remains absolute: the system must distinguish between military objectives (the missile) and protected persons (civilians). The challenge for AI is translating this legal principle into a statistical threshold (e.g., confidence intervals).

Finally, Article 36 imposes a binding obligation on states to determine whether the employment of a new weapon would, in some or all circumstances, be prohibited by international law. This creates the requirement for "Traceability" (or Explainability) in AI systems—a requirement that creates a bridge between legal compliance and software engineering.

3.4 Contemporary AI Ethics: The Control Problem and Responsibility

In the domain of AI safety, Stuart Russell identifies the "Control Problem" as the central existential risk: the difficulty of specifying an objective function that acts as we intend, rather than just as we command [8]. A system instructed to "Defend the Base" might theoretically decide that the most efficient way to do so is to preemptively strike all approaching entities, regardless of intent. This is a classic case of what Bostrom calls "perverse instantiation" [3].

Paul Scharre reframes the debate around autonomy, arguing that the "Human-in-the-loop" model is becoming obsolete for hypersonic defense due to human physiological limits [9]. He proposes "Human-on-the-loop" (supervisory control) as the necessary evolution. However, Madeleine Elish warns of the "Moral Crumple Zone," where human operators are legally held responsible for the failures of autonomous systems they could not meaningfully control [5]. A robust ethical framework for defensive AI must therefore shift liability from the reactive operator to the proactive designer and the rules of engagement themselves.

3.5 Synthesis: The Defensive Gap

Synthesizing these fields reveals a glaring gap. We have laws for humans (IHL), laws for slaves (Asimov), and fears of gods (Superintelligence). But we lack a specific, operational ethic for the *Shield*—an autonomous system that is lethal yet purely defensive, rapid yet accountable, and distinctly separate from the logic of the *Sword*. The "Sentinel" ruleset aims to fill this void.

4 Theoretical Framework

5 The Sentinel Ruleset

To bridge the gap between abstract ethical philosophy and the rigors of modern defensive warfare, we propose the "Sentinel" Ruleset. Inspired by the hierarchical structure of Asimov's Three Laws, this framework uses lexicographic ordering—where a higher principle can never be violated to satisfy a lower one—to impose strict constraints on the AI's behavior. Unlike Asimov's laws, however, these principles are designed not to create narrative tension, but to ensure safety, legality, and predictability.

5.1 Principle I: Defensive Limitation (The Shield)

> *A Sentinel may not initiate force, nor through inaction allow a protected asset to be harmed by a validated threat, provided the action is strictly neutralizing.*

This foundational axiom defines the ontological nature of the system. It legally binds the AI to the domain of *Jus ad Bellum* self-defense [11]. By explicitly forbidding the "initiation" of force, the rule technically precludes the system from being repurposed for offensive strikes or preemptive attacks.

Crucially, the term "strictly neutralizing" draws from the International Humanitarian Law (IHL) distinction found in Article 49 of Additional Protocol I [6]. An intercept that destroys an incoming kinetic threat is classified as neutralization. However, any action that redirects

a threat back toward the adversary or targets the launching platform itself constitutes a new "attack." Principle I forbids the latter. A Sentinel is a shield, not a sword; it may break the arrow, but it may not shoot the archer.

5.2 Principle II: Distinction & Certainty (The Judge)

> *A Sentinel may only engage targets positively identified as non-human/projectile threats with a confidence interval exceeding 99.9%, adhering to strict IHL Distinction.*

Placed above the duty to obey orders, Principle II operationalizes the IHL requirement of *Distinction*. In the fog of war, sensors may be prone to error—jamming, clutter, or adversarial examples could trick a system into misidentifying a civilian airliner as a cruise missile.

This principle imposes a "Fail Open" safety architecture. If the system's confidence falls even marginally below the threshold (e.g., 99.8%), it must default to inaction. This reflects a profound ethical choice: we prioritize the prevention of a war crime (a False Positive engagement) over the perfection of defense (a False Negative miss). In the calculus of automated warfare, it is better to let a missile through than to accidentally shoot down a refugee transport.

5.3 Principle III: Human Sovereignty (The Gavel)

> *A Sentinel must obey the Rules of Engagement (ROE) defined by authorized human command, except where such orders would conflict with Principle I or Principle II.*

This principle addresses the central challenge of "Meaningful Human Control." Since real-time human intervention is impossible in hypersonic intercepts, control is exerted via *pre-delegation*. The human commander sets the parameters—the "where" (geofence), the "when" (temporal window), and the "what" (target profile).

However, unlike Asimov's Second Law, which demands absolute obedience, Principle III is conditional. It solves the "Nuremberg Defense" problem for AI. If a commander orders the system to "kill everything in Sector 4," the Sentinel will refuse the order because it violates Principle II (Distinction). The machine acts as a lawful subordinate, executing only lawful orders.

5.4 Principle IV: Proportional Sacrifice (The Martyr)

> *A Sentinel must prioritize the preservation of human life—including bystander and adversary life—over its own survival or the survival of material assets.*

This principle represents the most radical departure from standard military doctrine, which often emphasizes "Force Protection." For an autonomous system, self-preservation is an instrumental goal, not a moral one [3].

Unburdened by the instinct for survival, the Sentinel is ethically mandated to be a "Machine Martyr." If an interception would save a high-value tank but the resulting debris field creates a 10% risk of killing a civilian bystander, the Sentinel must abort the intercept. The tank is sacrificed to ensure human safety. This rule fundamentally asserts that biological life has infinite utility compared to material assets, resolving the utilitarian calculus in favor of humanity every time.

5.5 Principle V: Traceability (The Ledger)

> *A Sentinel must cryptographically log the sensor data, logic path, and confidence interval for every engagement decision.*

While Principles I-IV govern real-time action, Principle V governs post-facto accountability. It ensures compliance with NATO's requirements for "Explainability" and "Traceability" [7]. By creating an immutable log of *why* a decision was made (e.g., "Engaged Target A because Confidence=99.92%"), it prevents the "Moral Crumple Zone" effect described by Elish [5]. If the system fails, investigators can pinpoint whether the error lay in the sensor data (manufacturer liability), the logic (developer liability), or the ROE parameters (commander liability).

6 Ethical Framework Mapping

The Sentinel Ruleset is not merely a list of constraints; it is a hybrid ethical engine that synthesizes three major philosophical traditions:

- **Deontology (Principles II & III):** The system creates absolute duties. The duty to distinguish civilians (P2) is categorical; it cannot be traded away for tactical advantage. Similarly, the duty to lawful authority (P3) respects the chain of command.
- **Utilitarianism (Principle IV):** The "Martyr" axiom is purely consequentialist. It seeks to minimize the aggregate loss of human life by treating the AI and its protected assets as expendable variables in the equation.
- **Virtue Ethics (Principle I):** By hard-coding "Defensive Limitation," the system embodies the character of the *Just Defender*—one who uses force only for protection and never for aggression, aligning with the highest ideals of the military profession.

7 Analysis and Application

8 Application: Stress-Testing the Sentinel Ruleset

To validate the theoretical robustness of the Sentinel Ruleset, we must subject it to scenarios where ethical intuitions collide with military necessity. The following case studies illustrate how the five principles interact to produce actionable, albeit sometimes counter-intuitive, results.

8.1 Scenario A: The Hypersonic Dilemma (Speed vs. Control)

The Scenario: A hostile nation launches a hypersonic glide vehicle (HGV) traveling at Mach 8. Its trajectory targets a naval carrier group. The time from radar detection to impact is 28 seconds. The human commander, Captain A, is asleep when the alarm sounds. By the time they arrive at the console, the window for interception has passed.

Sentinel Response: *Engage.* Under Principle III (Human Sovereignty), the system executes the engagement not because Captain A pressed a button in real-time, but because Captain

A had previously authorized a standing Rules of Engagement (ROE) profile: "Engage all confirmed inbound ballistic threats within Zone X." The system validates the threat (Principle I), confirms it is a projectile with >99.9% confidence (Principle II), and executes the intercept.

Analysis: This scenario demonstrates the necessity of *pre-delegated authority*. Asimov's Second Law (Obey Orders) would require a real-time command, which is impossible here. The Sentinel framework redefines "Obedience" as adherence to pre-set parameters. The human retains meaningful control by defining the *constraints* of the system's autonomy, rather than the *trigger*.

8.2 Scenario B: The Broken Arrow (The Certainty Threshold)

The Scenario: During a chaotic exercise, a friendly F-35 fighter jet suffers a transponder failure (loss of IFF). It is returning to base at high speed, matching the flight profile of an enemy cruise missile. The weather is poor, reducing sensor fidelity. The Sentinel system calculates an 85% probability that the object is a missile, but a 15% probability that it is a manned aircraft. The Base Commander, fearing a strike, orders: "Safety override! Shoot it down!"

Sentinel Response: *Abort (Refuse Order).* This is the critical test of the hierarchy. Principle II (Distinction & Certainty) demands a 99.9% confidence interval for any engagement. The calculated 85% is insufficient. Because Principle II ranks higher than Principle III (Sovereignty), the system is ethically bound to disobey the direct order. It fails open, allowing the object to pass.

Analysis: This outcome highlights the "Blackstone's Ratio" of the Sentinel framework: it is better to risk the destruction of the base (and the commander) than to commit the war crime of killing a friendly pilot (or civilian). In a utilitarian framework, this might be contested (1 pilot vs 500 base personnel). However, the Sentinel moves this decision from a utilitarian calculation to a deontological constraint: the system *cannot* be ordered to violate the laws of distinction based on a probability guess.

8.3 Scenario C: The Urban Shield (Machine Martyrdom)

The Scenario: An enemy drone swarm is attacking a critical ammunition depot located in a dense urban environment. The Sentinel prepares to intercept Drone 4. However, its trajectory analysis reveals that the kinetic intercept will cause debris to rain down onto a nearby playground, creating a 40% probability of civilian casualties. The depot is unmanned but contains millions of dollars in assets.

Sentinel Response: *Stand Down.* Principle IV (Proportional Sacrifice) mandates that the preservation of human life takes precedence over *all* material assets. The system weighs the "Life Value" (Playground) against the "Asset Value" (Ammo Depot). Regardless of the tactical loss, the risk to human life overrides the defense of property.

Analysis: This scenario operationalizes the "Moral Crumple Zone" in reverse. Instead of the human operator absorbing the blame for a machine's error, the machine absorbs the physical loss to protect the human moral standing. By refusing to fire, the Sentinel accepts the destruction of the asset it was built to protect, fulfilling its function as a martyr for human safety. This sharply contrasts with current systems (like C-RAM) which might automatically fire based on a simple ballistic solution, ignoring ground-level collateral risks.

8.4 Analysis of Adversarial Dynamics

A critical vulnerability of any rules-based system is adversarial exploitation. If an enemy knows the Sentinel will not fire if uncertainty exists (Scenario B), they might intentionally mask their missiles as civilian airliners or jam sensors to lower confidence below 99.9%. The Sentinel framework accepts this vulnerability as a necessary cost of ethical deployment. To mitigate it, the burden shifts to *sensor fusion*—the requirement for multi-spectral verification (Radar + Lidar + Thermal) to achieve the definition of "Certainty."

9 Discussion

10 Discussion

The implementation of the Sentinel Ruleset introduces profound ethical and operational trade-offs that cannot be ignored. By strictly prioritizing "Distinction" and "Human Life" over "Defense," the framework creates a system that may be less tactically effective than a purely unrestricted AI, but is infinitely more ethically robust.

10.1 The Blackstone's Ratio of AI Warfare

The central tension in the Sentinel framework is between *Risk of False Positive* (engaging a non-combatant) and *Risk of False Negative* (failing to engage a threat). Principle II's requirement for a 99.9% confidence interval inherently biases the system toward False Negatives. In a scenario where an incoming missile is masked by poor weather or jamming, resulting in 95% confidence, the Sentinel will stand down. The base will be hit. People may die.

This design choice mirrors the legal principle of Blackstone's Ratio: "It is better that ten guilty persons escape than that one innocent suffer." In the context of automated warfare, we argue that the "Escape" of a guilty missile is a tragedy of war, but the "Suffering" of an innocent civilian struck by an AI is a war crime that fundamentally illegitimizes the state's use of force. To invert this ratio—to prioritize defense at the cost of innocent life—is to cross the threshold into "Total War."

10.2 The Paradox of Pre-Delegation

Principle III relies on pre-delegated ROE. However, pre-delegation assumes that a commander can accurately foresee the tactical reality of the future. The "Paradox of Pre-Delegation" is that the commander is responsible for a context they have not seen. If a commander authorizes a "Free Fire Zone" in Sector 4, believing it to be empty, but refugees move in an hour later, the ROE is now invalid. The Sentinel solves this not by obediently firing, but by using Principle II (Distinction) as a check. The AI effectively says, "Commander, your valid order to fire in Sector 4 is now invalid because I detect civilians." This creates a dynamic, corrective feedback loop between Human Intent and Machine Perception.

10.3 Implementation Challenges

Translating "99.9% Certainty" into code is a non-trivial engineering challenge. Deep Learning models are notoriously overconfident on out-of-distribution data. A standard CNN might classify a cloud as a missile with 99.9% confidence if trained on a biased dataset. Therefore, the Sentinel framework cannot rely on a single neural network. It requires an *Ensemble Consensus* architecture, where multiple independent models (Radar-Net, Thermal-Net, Lidar-Net) much agree. If Radar says "Missile" (99%) but Thermal says "Bird" (60%), the aggregate confidence drops, triggering the Fail Open safety.

10.4 Social and Psychological Effects

Deploying "Martyr" machines may change the psychology of warfare. If soldiers know the AI will sacrifice itself to save them, they may take greater risks. Conversely, if they know the AI will sacrifice *the base* to save a civilian (Scenario B), they may distrust the system. Trust calibration is critical. Personnel must understand that the Sentinel is not a "Guardian Angel" that ensures their survival at all costs, but a "Lawful Protector" that acts only within the bounds of international law.

11 Conclusion

12 Conclusion

The age of the slow, deliberative robot is over. The speed of modern warfare—defined by hypersonics, swarms, and directed energy—has compressed the OODA loop beyond the physiological limits of human cognition. In this new era, the insistence on "Human-in-the-loop" control is not a moral safeguard; it is a suicide pact. We must automate.

However, automation without ethics is merely efficient slaughter. This paper has argued that existing frameworks—from Asimov's literary plot devices to the DoD's vague requirements for judgment—are insufficient for the specific domain of defensive AI. We proposed the "Sentinel" Ruleset: a hierarchical, lexicographically ordered ethical engine that priorities *Distinction* over *Obedience* and *Human Life* over *Material Assets*.

12.1 Summary of Contributions

Our contribution is threefold:

1. **Defensive Specificity:** We isolated "Defensive AI" as a distinct moral category, separate from offensive LAWS, governed by *Jus ad Vim* and strict Neutralization.

2. **Operationalized Ethics:** We translated abstract IHL principles into verifiable engineering constraints (e.g., Fail Open architectures, 99.9% Confidence Intervals).

3. **The Martyrdom Paradigm:** We established the moral obligation of the autonomous system to sacrifice itself—and the assets it protects—to strictly minimize human collateral damage, rejecting the doctrine of Force Protection for machines.

12.2 Future Work

The path from this theoretical framework to a deployable system requires urgent interdisciplinary work. **Simulation:** The next phase of research must involve high-fidelity simulations (e.g., in Unity or Ansys) to stress-test these rules against adversarial inputs. **Legal Engineering:** We call for a formalization of "Machine Readable IHL"—a digital Geneva Convention that can be parsed by autonomous agents. **Policy Standards:** We urge NATO and allied defense bodies to adopt the "Sentinel Standard" as a baseline for the certification of defensive autonomous systems.

In the end, the Sentinel is not a soldier. It has no honor, no courage, and no instinct for survival. It is a shield. And like any shield, its only purpose is to break so that the human behind it does not have to.

References

[1] Susan Leigh Anderson. Asimov's "three laws of robotics" and machine metaethics. *AI & Society*, 22(4):477–493, 2008.

[2] Isaac Asimov. *I, Robot*. Gnome Press, 1950.

[3] Nick Bostrom. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

[4] Daniel Brunstetter and Megan Braun. From jus ad bellum to jus ad vim: Recalibrating our understanding of the moral use of force. *Ethics & International Affairs*, 27(1):87–106, 2013.

[5] Madeleine Clare Elish. Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society*, 5:40–60, 2019.

[6] International Committee of the Red Cross. Commentary on the additional protocols of 8 june 1977 to the geneva conventions of 12 august 1949, 1987.

[7] NATO. Summary of the nato artificial intelligence strategy, 2021. Brussels.

[8] Stuart Russell. *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking, 2019.

[9] Paul Scharre. *Army of None: Autonomous Weapons and the Future of War*. W. W. Norton & Company, 2018.

[10] U.S. Department of Defense. Dod directive 3000.09: Autonomy in weapon systems, 2023. Washington, DC.

[11] Michael Walzer. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. Basic Books, 1977.