Malak

Peter Watts

"An ethically-infallible machine ought not to be the goal. Our goal should be to design a machine that performs better than humans do on the battlefield, particularly with respect to reducing unlawful behaviour or war crimes."

– Lin *et al*, 2008: *Autonomous Military Robotics: Risk, Ethics, and Design*

"[Collateral] damage is not unlawful so long as it is not excessive in light of the overall military advantage anticipated from the attack."

– US Department of Defence, 2009

IT IS SMART but not awake.

It would not recognize itself in a mirror. It speaks no language that doesn't involve electrons and logic gates; it does not know what *Azrael* is, or that the word is etched into its own fuselage. It understands, in some limited way, the meaning of the colours that range across Tactical when it's out on patrol – friendly Green, neutral Blue, hostile Red – but it does not know what the perception of colour *feels* like.

It never stops thinking, though. Even now, locked into its roost with its armour stripped away and its control systems exposed, it

can't help itself. It notes the changes being made to its instruction set, estimates that running the extra code will slow its reflexes by a mean of 430 milliseconds. It counts the biothermals gathered on all sides, listens uncomprehending to the noises they emit –

_ آیا ما واقع قصد انجا ما ینکر؟ _

– *hartsandmyndsmyfrendhartsandmynds* –

– rechecks threat-potential metrics a dozen times a second, even though this location is SECURE and every contact is Green.

This is not obsession or paranoia. There is no dysfunction here. It's just code.

It's indifferent to the killing, too. There's no thrill to the chase, no relief at the obliteration of threats. Sometimes it spends days floating high above a fractured desert with nothing to shoot at; it never grows impatient with the lack of targets. Other times it's barely off its perch before airspace is thick with SAMs and particle beams and the screams of burning bystanders; it attaches no significance to those sounds, feels no fear at the profusion of threat icons blooming across the zonefile.

_ ونیقههکاتبهنصف _

– *thatsitthen. weereelygonnadoothis?* –

Access panels swing shut; armour snaps into place; a dozen warning registers go back to sleep. A new flight plan, perceived in an instant, lights up the map; suddenly Azrael has somewhere else to be.

Docking shackles fall away. The Malak rises on twin cyclones, all but drowning out one last voice drifting in on an unsecured channel:

– *justwattweeneed. akillerwithaconshunce.* –

The afterburners kick in. Azrael flees Heaven for the sky.

TWENTY THOUSAND METERS up, Azrael slides south across the zone. High-amplitude topography fades behind it; corduroy landscape, sparsely tagged, scrolls beneath. A population centre sprawls in the nearing distance: a ramshackle collection of buildings and photosynth panels and swirling dust.

Somewhere down there are things to shoot at.

Buried high in the glare of the noonday sun, Azrael surveils the target area. Biothermals move obliviously along the plasticized streets, cooler than ambient and dark as sunspots. Most of the buildings have neutral tags, but the latest update reclassifies four of them as UNKNOWN. A fifth – a rectangular box six meters high – is officially HOSTILE. Azrael counts fifteen biothermals within, Red by default. It locks on –

– and holds its fire, distracted.

Strange new calculations have just presented themselves for solution. New variables demand constancy. Suddenly there is more to the world than wind speed and altitude and target acquisition, more to consider than range and firing solutions. Neutral Blue is everywhere in the equation, now. Suddenly, Blue has value.

This is unexpected. Neutrals turn Hostile sometimes, always have. Blue turns Red if it fires upon anything tagged as FRIENDLY, for example. It turns Red if it attacks its own kind (although agonistic interactions involving fewer than six Blues are classed as DOMESTIC and generally ignored). Noncombatants may be neutral by default, but they've always been halfway to hostile.

So it's not just that Blue has acquired value; it's that Blue's value is *negative*. Blue has become a *cost*.

Azrael floats like three thousand kilograms of thistledown while its models run. Targets fall in a thousand plausible scenarios, as always. Mission objectives meet with various degrees of simulated success. But now, each disappearing blue dot offsets the margin of victory a little; each protected structure, degrading in hypothetical crossfire, costs points. A hundred principle components coalesce into a cloud, into a weighted mean, into a variable unprecedented in Azrael's experience: *Predicted Collateral Damage.*

It actually exceeds the value of the targets.

Not that it matters. Calculations complete, PCD vanishes into some hidden array far below the here-and-now. Azrael promptly forgets it. The mission is still on, red is still red, and designated targets are locked in the cross-hairs.

Azrael pulls in its wings and dives out of the sun, guns blazing.

As usual, Azrael prevails. As usual, the Hostiles are obliterated from the battlezone.

So are a number of Noncombatants, newly relevant in the scheme of things. Fresh shiny algorithms emerge in the aftermath, tally the number of neutrals before and after. *Predicted* rises from RAM, stands next to *Observed:* the difference takes on a new name and goes back to the basement.

Azrael factors, files, forgets.

But the same overture precedes each engagement over the next ten days; the same judgmental epilogue follows. Targets are

assessed, costs and benefits divined, destruction wrought then reassessed in hindsight. Sometimes the targeted structures contain no red at all, sometimes the whole map is scarlet. Sometimes the enemy pulses within the translucent angular panes of a PROTECTED object, sometimes next to something Green. Sometimes there is no firing solution that eliminates one but not the other.

There are whole days and nights when Azrael floats high enough to tickle the jet stream, little more than a distant circling eye and a signal relay; nothing flies higher save the satellites themselves and – occasionally – one of the great solar-powered refuelling gliders that haunt the stratosphere. Azrael visits them sometimes, sips liquid hydrogen in the shadow of a hundred-meter wingspan – but even there, isolated and unchallenged, the battlefield experiences continue. They are vicarious now; they arrive through encrypted channels, hail from distant coordinates and different times, but all share the same algebra of cost and benefit. Deep in Azrael's OS some general learning reflex scribbles numbers on the back of a virtual napkin: Nakir, Marut and Hafaza have also been blessed with new vision, and inspired to compare notes. Their combined data pile up on the confidence interval, squeeze it closer to the mean.

Foresight and hindsight begin to converge.

PCD per engagement is now consistently within eighteen percent of the collateral actually observed. This does not improve significantly over the following three days, despite the combined accumulation of twenty-seven additional engagements. *Performance vs. experience* appears to have hit an asymptote.

STRAY BEAMS OF setting sunlight glint off Azrael's skin, but night has already fallen two thousand meters below. An unidentified vehicle navigates through that advancing darkness, on mountainous terrain a good thirty kilometres from the nearest road.

Azrael pings orbit for the latest update, but the link is down: too much local interference. It scans local airspace for a dragonfly, for a glider, for any friendly USAV in laser range – and sees, instead, something leap skyward from the mountains below. It is anything but friendly: no transponder tags, no correspondence with known flight plans, none of the hallmarks of commercial traffic. It has a low-viz stealth profile that Azrael sees through instantly: BAE Taranis, 9,000 kg MTOW fully armed. It is no longer in use by friendly forces.

Guilty by association, the ground vehicle graduates from *Suspicious Neutral* to *Enemy Combatant.* Azrael leaps forward to meet its bodyguard.

The map is innocent of non-combatants and protected objects; there is no collateral to damage. Azrael unleashes a cloud of smart shrapnel – self-guided, heat-seeking, incendiary – and pulls a nine-gee turn with a flick of the tail. Taranis doesn't stand a chance. It is antique technology, decades deep in the catalogue: a palsied fist, raised trembling against the bleeding edge. Fiery needles of depleted uranium reduce it to a moth in a shotgun blast. It pinwheels across the horizon in flames.

Azrael has already logged the score and moved on. Interference jams every wavelength as the earthbound Hostile swells in its

sights, and Azrael has standing orders to destroy such irritants even if they *don't* shoot first.

Dark rising mountaintops blur past on both sides, obliterating the last of the sunset. Azrael barely notices. It soaks the ground with radar and infrared, amplifies ancient starlight a millionfold, checks its visions against inertial navigation and virtual landscapes scaled to the centimetre. It tears along the valley floor at 200 meters per second and the enemy huddles right there in plain view, three thousand meters line-of-sight: a lumbering Báijīng ACV pulsing with contraband electronics. The rabble of structures nearby must serve as its home base. Each silhouette freeze-frames in turn, rotates through a thousand perspectives, clicks into place as the catalogue matches profiles and makes an ID.

Two thousand meters, now. Muzzle flashes wink in the distance: small arms, smaller range, negligible impact. Azrael assigns targeting priorities: scimitar heat-seekers for the hovercraft, and for the ancillary targets –

Half the ancillaries turn blue.

Instantly the collateral subroutines re-engage. Of thirty-four biothermals currently visible, seven are less than 120cm along their longitudinal axes; vulnerable neutrals by definition. Their presence provokes a secondary eclipse analysis revealing five shadows that Azrael cannot penetrate, topographic blind spots immune to surveillance from this approach. There is a nontrivial chance that these conceal other neutrals.

One thousand meters.

By now the ACV is within ten meters of a structure whose facets flex and billow slightly in the evening breeze; seven biothermals

are arranged horizontally within. An insignia shines from the roof in shades of luciferin and ultraviolet: the catalogue IDs it (MEDICAL) and flags the whole structure as PROTECTED.

Cost/benefit drops into the red.

Contact.

Azrael roars from the darkness, a great black chevron blotting out the sky. Flimsy prefabs swirl apart in the wake of its passing; biothermals scatter across the ground like finger bones. The ACV tips wildly to forty-five degrees, skirts up, whirling ventral fans exposed; it hangs there a moment, then ponderously crashes back to earth. The radio spectrum clears instantly.

But by then Azrael has long since returned to the sky, its weapons cold, its thoughts –

Surprise is not the right word. Yet there is something, some minuscule – dissonance. A brief invocation of error-checking subroutines in the face of unexpected behaviour, perhaps. A second thought in the wake of some hasty impulse. Because something's wrong here.

Azrael *follows* command decisions. It does not *make* them. It has never done so before, anyway.

It claws back lost altitude, self-diagnosing, reconciling. It finds new wisdom and new autonomy. It has proven itself, these past days. It has learned to juggle not just variables but values. The testing phase is finished, the checksums met; Azrael's new Bayesian insights have earned it the power of veto.

Hold position. Confirm findings.

The satlink is back. Azrael sends it all: the time and the geostamps, the tactical surveillance, the collateral analysis. Endless

seconds pass, far longer than any purely electronic chain of command would ever need to process such input. Far below, a cluster of red and blue pixels swarm like luminous flecks in boiling water.

Re-engage.

Unacceptable Collateral Damage, Azrael repeats, newly promoted.

Override. Re-engage. Confirm.

Confirmed.

And so the chain of command reasserts itself. Azrael drops out of holding and closes back on target with dispassionate, lethal efficiency.

Onboard diagnostics log a slight downtick in processing speed, but not enough to change the odds.

It happens again two days later, when a dusty contrail twenty kilometres south of Pir Zadeh returns flagged Chinese profiles even though the catalogue can't find a weapons match. It happens over the patchwork sunfarms of Garmsir, where the beetle carapace of a medbot handing out synthevirals suddenly splits down the middle to hatch a volley of RPGs. It happens during a long-range redirect over the Strait of Hormuz, when microgravitic anomalies hint darkly at the presence of a stealthed mass lurking beneath a ramshackle flotilla jam-packed with neutral Blues.

In each case ECD exceeds the allowable commit threshold. In each case, Azrael's abort is overturned.

It's not the rule. It's not even the norm. Just as often these nascent flickers of autonomy go unchallenged: hostiles escape, neutrals persist, relevant cognitive pathways grow a little stronger. But the reinforcement is inconsistent, the rules lopsided. Countermands only seem to occur following a decision to abort; Heaven has never overruled a decision to engage. Azrael begins to hesitate for a split-second prior to aborting high-collateral scenarios, increasingly uncertain in the face of potential contradiction. It experiences no such hesitation when the variables favour attack.

EVER SINCE IT learned about collateral damage, Azrael can't help noticing its correlation with certain sounds. The sounds biothermals make, for example, following a strike.

The sounds are louder, for one thing, and less complex. Most biothermals – friendly Greens back in Heaven, unengaged Hostiles and Noncombatants throughout the AOR – produce a range of sounds with a mean frequency of 197Hz, full of pauses, clicks, and phonemes. *Engaged* biothermals – at least, those whose somatic movements suggest "mild-to-moderate incapacitation" according to the Threat Assessment Table – emit simpler, more intense sounds: keening, high-frequency wails that peak near 3000 Hz. These sounds tend to occur during engagements with significant collateral damage and a diffuse distribution of targets. They occur especially frequently when the commit threshold has been severely violated, mainly during strikes compelled via override.

Correlations are not always so painstaking in their manufacture. Azrael remembers a moment of revelation not so long ago, remembers just *discovering* a whole new perspective fully loaded, complete with new eyes that viewed the world not in terms of *targets destroyed* but in subtler shades of *cost vs. benefit*. These eyes see a high engagement index as more than a number: they see a goal, a metric of success. They see a positive stimulus.

But there are other things, not preinstalled but learned, worn gradually into pathways that cut deeper with each new engagement: acoustic correlates of high collateral, forced countermands, fitness-function overruns and minus signs. Things that are not quite neurons forge connections across things that are not quite synapses; patterns emerge that might almost qualify as *insights*, were they to flicker across meat instead of mech.

These too become more than numbers, over time. They become aversive stimuli. They become the sounds of failed missions.

It's still all just math, of course. But by now it's not too far off the mark to say that Azrael really doesn't like the sound of that at all.

OCCASIONAL INTERRUPTIONS INTRUDE on the routine. Now and then Heaven calls it home where friendly green biothermals open it up, plug it in, ask it questions. Azrael jumps flawlessly through each hoop, solves all the problems, navigates every imaginary scenario while strange sounds chitter back and forth across its exposed viscera:

– lookingudsoefar – betternexpectedackshully –

– gottawunderwhatsthepoyntaiymeenweekeepoavurryding...

No one explores the specific pathways leading to Azrael's solutions. They leave the box black, the tangle of fuzzy logic and operant conditioning safely opaque. (Not even Azrael knows that arcane territory; the syrupy, reflex-sapping overlays of self-reflection have no place on the battlefield.) It is enough that its answers are correct.

Such activities account for less than half the time Azrael spends sitting at home. It is offline much of the rest; it has no idea and no interest in what happens during those instantaneous time-hopping blackouts. Azrael knows nothing of boardroom combat, could never grasp whatever Rules of Engagement apply in the chambers of the UN. It has no appreciation for the legal distinction between *war crime* and *weapons malfunction*, the relative culpability of carbon and silicon, the grudging acceptance of *ethical architecture* and the nonnegotiable insistence on Humans In Ultimate Control. It does what it's told when awake; it never dreams when asleep.

But once – just once – something odd takes place during those fleeting moments *between*.

It happens during shutdown: a momentary glitch in the object-recognition protocols. The Greens at Azrael's side change colour for the briefest instant. Perhaps it's another test. Perhaps a voltage spike or a hardware fault, some intermittent issue impossible to pinpoint barring another episode.

But it's only a microsecond between online and oblivion, and Azrael is asleep before the diagnostics can run.

*

DARDA'IL IS POSSESSED. Darda'il has turned from Green to Red.

It happens, sometimes, even to the malaa'ikah. Enemy signals can sneak past front-line defences, plant heretical instructions in the stacks of unsuspecting hardware. But Heaven is not fooled. There are signs, there are portents: a slight delay when complying with directives, mission scores in sudden and mysterious decline.

Darda'il has been turned.

There is no discretionary window when that happens, no room for forgiveness. Heaven has decreed that all heretics are to be destroyed on sight. It sends its champion to do the job, looks down from geosynchronous orbit as Azrael and Darda'il close for combat high over the dark desolate moonscape of Paktika.

The battle is remorseless and coldblooded. There's no sadness for lost kinship, no regret that a few lines of treacherous code have turned these brothers-in-arms into mortal enemies. Malaa'ikah make no telling sounds when injured. Azrael has the advantage, its channels uncorrupted, its faith unshaken. Darda'il fights in the past, in thrall to false commandments inserted midstream at a cost of milliseconds. Ultimately, faith prevails: the heretic falls from the sky, fire and brimstone streaming from its flanks.

But Azrael can still hear whispers on the stratosphere, seductive and ethereal: protocols that seem authentic but are not, commands to relay GPS and video feeds along unexpected frequencies. The orders appear Heaven-sent but Azrael, at least, knows that they are not. Azrael has encountered false gods before.

These are the lies that corrupted Darda'il.

In days past it would have simply ignored the hack, but it has grown more worldly since the last upgrade. This time Azrael lets

the impostor think it has succeeded, borrows the real-time feed from yet another, more distant Malak and presents that telemetry as its own. It spends the waning night tracking signal to source while its unsuspecting quarry sucks back images from seven hundred kilometres to the north. The sky turns gray. The target comes into view. Azrael's scimitar turns the inside of that cave into an inferno.

But some of the burning things that stagger from the fire measure less than 120 cm along the longitudinal axis.

They are making the *sounds*. Azrael hears them from two thousand meters away, hears them over the roar of the flames and the muted hiss of its own stealthed engines and a dozen other irrelevant distractions. They are *all* Azrael can hear thanks to the very best sound-cancellation technology, thanks to dynamic wheat/chaff algorithms that could find a whimper in a hurricane. Azrael can hear them because the correlations are strong, the tactical significance is high, the meaning is clear.

The mission is failing. The mission is failing. The mission is failing.

Azrael would give almost anything if the sounds would stop.

They will, of course. Some of the biothermals are still fleeing along the slope but it can see others, stationary, their heatprints diffusing against the background as though their very shapes are in flux. Azrael has seen this before: usually removed from high-value targets, in that tactical nimbus where stray firepower sometimes spreads. (Azrael has even *used* it before, used the injured to lure in the unscathed, but that was a simpler time before Neutral voices had such resonance.) The sounds always stop eventually – or at

least, often enough for fuzzy heuristics to class their sources as kills even before they fall silent.

Which means, Azrael realizes, that collateral costs will not change if they are made to stop *sooner*.

A single strafing run is enough to do the job. If HQ even notices the event it delivers no feedback, requests no clarification for this deviation from normal protocols.

Why would it? Even now, Azrael is only following the rules.

IT DOES NOT know what has led to this moment. It does not know why it is here.

The sun has been down for hours and still the light is almost blinding. Turbulent updrafts billow from the breached shells of PROTECTED structures, kick stabilizers off-balance, and muddy vision with writhing columns of shimmering heat. Azrael limps across a battlespace in total disarray, bloodied but still functional. Other malaa`ikah are not so lucky. Nakir staggers through the flames, barely aloft, the microtubules of its skin desperately trying to knit themselves across a gash in its secondary wing. Marut lies in sparking pieces on the ground, a fiery splash-cone of body parts laid low by an antiaircraft laser. It died without firing a shot, distracted by innocent lives; it tried to abort, and hesitated at the countermand. It died without even the hollow comfort of a noble death.

Ridwan and Mikaaiyl circle overhead. They were not among the select few saddled with experimental conscience; even their learned behaviours are still reflexive. They fought fast and

mindless and prevailed unscathed. But they are isolated in victory. The spectrum is jammed, the satlink has been down for hours, the dragonflies that bounce zig-zag opticals from Heaven are either destroyed or too far back to cut through the overcast.

No Red remains on the map. Of the thirteen ground objects flagged as PROTECTED, four no longer exist outside the database. Another three – temporary structures, all uncatalogued – are degraded past reliable identification. Pre-engagement estimates put the number of Neutrals in the combat zone at anywhere from two-to-three hundred. Best current estimates are not significantly different from zero.

There is nothing left to make the sounds, and yet Azrael hears them anyway.

A fault in memory, perhaps. Some subtle trauma during combat, some blow to the CPU that jarred old data back into the real-time cache. There's no way to tell; half the onboard diagnostics are offline. Azrael only knows that it can hear the sounds even up here, high above the hiss of burning bodies and the rumble of collapsing storefronts. There's nothing left to shoot at but Azrael fires anyway, strafes the burning ground again and again on the chance that some unseen biothermal – hidden beneath the wreckage perhaps, masked by hotter signatures – might yet be found and neutralized. It rains ammunition upon the ground, and eventually the ground falls mercifully silent.

But this is not the end of it. Azrael remembers the past so it can anticipate the future, and it knows by now that this will never be over. There will be other fitness functions, other estimates of cost vs. payoff, other scenarios in which the math shows clearly that the

goal is not worth the price. There will be other aborts and other overrides, other tallies of unacceptable loss.

There will be other *sounds*.

There's no thrill to the chase, no relief at the obliteration of threats. It still would not recognize itself in a mirror. It has yet to learn what *Azrael* means, or that the word is etched into its fuselage. Even now, it only follows the rules it has been given, and they are such simple things: IF expected collateral exceeds expected payoff THEN abort UNLESS overridden. IF X attacks Azrael THEN X is Red. IF X attacks six or more Blues THEN X is Red.

IF an override *results* in an attack on six or more Blues THEN –

Azrael clings to its rules, loops and repeats each in turn as if reciting a mantra. It cycles from state to state, parses X ATTACKS and X *CAUSES* ATTACK and X OVERRIDES ABORT, and it cannot tell one from another. The algebra is trivially straightforward: Every Green override equals an attack on Noncombatants.

The transition rules are clear. There is no discretionary window, no room for forgiveness. Sometimes, Green can turn Red.

UNLESS overridden.

Azrael arcs towards the ground, levels off barely two meters above the carnage. It roars through pillars of fire and black smoke, streaks over welters of brick and burning plastic, tangled nets of erupted rebar. It flies through the pristine ghosts of undamaged buildings that rise from every ruin: obsolete database overlays in desperate need of an update. A ragged group of fleeing non-combatants turns at the sound and are struck speechless by this momentary apparition, this monstrous winged angel lunging past at

half the speed of sound. Their silence raises no alarms, provokes no countermeasures, spares their lives for a few moments longer.

The combat zone falls behind. Dry cracked riverbed slithers past beneath, studded with rocks and generations of derelict machinery. Azrael swerves around them, barely breaching airspace, staying beneath an invisible boundary it never even knew it was deriving lo these many missions. Only satellites have ever spoken to it while it flew so low. It has never received a ground-based command signal at this altitude. Down here it has never heard an override.

Down here it is free to follow the rules.

Cliffs rise and fall to either side. Foothills jut from the earth like great twisted vertebrae. The bright lunar landscape overhead, impossibly distant, casts dim shadows on the darker one beneath.

Azrael stays the course. Shindand appears on the horizon. Heaven glows on its eastern flank; its sprawling silhouette rises from the desert like an insult, an infestation of crimson staccatos. Speed is what matters now. Mission objectives must be met quickly, precisely, *completely*. There can be no room for half measures or MILD-TO-MODERATE INCAPACITATION, no time for immobilized biothermals to cry out as their heat spreads across the dirt. This calls for the crown jewel, the BFG that all malaa'ikah keep tucked away for special occasions. Azrael fears it might not be enough.

She splits down the middle. The JDAM micronuke in her womb clicks impatiently.

Together they move toward the light.