

Rafał KOPEĆ

AUTONOMIA SYSTEMÓW BOJOWYCH

Abstrakt:

Autonomiczne systemy bojowe są jednym z najbardziej obiecujących kierunków rozwoju broni. Autonomizacja pozwolić ma na urzeczywistnienie wizji robotolucji, czyli rewolucji w sposobie prowadzenia działań zbrojnych polegającej na szerokim wykorzystaniu robotów. Jest to zagadnienie niezmiernie istotne w kontekście prognozowania kierunków rozwoju sił zbrojnych, ale narażone na liczne uproszczenia z powodu zawłości terminologicznych. Artykuł podejmuje wyzwanie określenia istoty autonomii systemów bojowych. Autor odpowiada na pytania o założenia autonomii w odniesieniu do maszyn, odróżnia autonomię od pojęć pokrewnych, określa stopnie autonomii oraz jej kryteria względem systemów bojowych., wreszcie staje wobec dylematu, czy autonomia systemów bojowych jest kwestią przyszłości czy elementem teraźniejszości.

Słowa kluczowe: autonomia, roboty bojowe, robotolucja, broń inteligentna, pętla decyzyjna.

Wprowadzenie

Prowadzenie działań zbrojnych podlega nieustannym przekształceniom. Rewolucyjne zmiany w tym zakresie zyskują miano tzw. rewolucji w sprawach wojskowych (Kopeć 2014). Być może stoimy obecnie u progu kolejnej rewolucji, związanej tym razem z wykorzystaniem robotów na polu bitwy. Zwolennicy takiego podejścia (Windeck 2014, s. V) mówią o tzw. robotolucji (*robotolucja*). Termin ten, powstały z połączenia słów robotyzacja i rewolucja, opiera się na założeniu, że robotyzacja prowadzi do rewolucyjnych zmian i stanowi punkt zwrotny w historii militarnej. Pozostaje pytanie, jak bliska jest ta perspektywa.

Nie jest to pytanie trywialne, gdyż od odpowiedzi na niego zależy określenie kierunków przyszłego rozwoju sił zbrojnych. W ten kontekst wpisują się chociażby sugestie dokonania przeskoku generacyjnego. W polskich warunkach propagatorem tej idei był gen. Stanisław Koziej (2008, s. 38), który namawiał do skupienia się na pracach nad przyszłościowymi rodzajami broni, zakładając, że takie systemy jak bezpilotowe aparaty latające „z pewnością za 15 – 20 lat zdominują powietrzny wymiar operacji wojskowych”. Teza taka postawiona została w 2008 roku, więc horyzont czasowy, w którym ta wizja ma

się zmaterializować, nieubłagane się zbliża. Tymczasem uzbrojone bezpilotowe statki latające, awangarda robotyzacji pola bitwy, nadal stanowią w najlepszym razie niszę i są użyteczne tylko w ściśle określonych warunkach (w sytuacji całkowitego braku obrony powietrznej i sił powietrznych przeciwnika).

Jedną z najpoważniejszych słabości obecnie stosowanych systemów jest możliwość zakłócenia kanałów łączności, albo wręcz przejścia kontroli nad aparatami bezzałogowymi przez przeciwnika¹. Najprawdopodobniej taki scenariusz udało się zrealizować Irańczykom, którzy sprowadzili na ziemię i przejęli w stanie praktycznie nieuszkodzonym amerykański rozpoznawczy dron RQ-170 Sentinel². Rozwiązaniem tych problemów mogą być aparaty autonomiczne, gdzie „ośrodek decyzyjny” w postaci jakiejś formy sztucznej inteligencji zostanie umieszczony na pokładzie maszyny.

W pełni użyteczne uderzeniowe bezpilotowe statki latające, być może będące pokłosiem obecnie realizowanego programu badawczo-rozwojowego UCLASS (Unmanned Carrier-Launched Airborne Surveillance and Strike) zapewne nie wejdą do użytku przed rokiem 2030 (Gajzler 2015, s. 14). Jednym z najważniejszych kierunków rozwojowych przyszłych aparatów bezzałogowych ma być zwiększenie autonomii (podkreśla to m.in. opracowany przez amerykański Departament Obrony w 2013 Unmanned Systems Integrated Roadmap FY 2013-2038³). Podkreśla się, że zwiększona autonomia może być rozwiązaniem problemu kontroli nad łączami przesyłu danych, ale również sprawdzi się w sytuacji rosnącej dynamiki pola walki, gdzie środowisko jest zbyt skomplikowane, by człowiek mógł dalej bezpośrednio kierować systemami bojowymi⁴.

W kontekście prognozowania kierunków rozwoju sił zbrojnych należy więc określić dokładnie, czym jest autonomia maszyn bojowych oraz odpowiedzieć na pytanie, na ile jest ona kwestią przyszłości, czy też może pewne jej formy funkcjonują już w obecnie użytkowanych systemach bojowych. Te dwa pytania porządkują treść niniejszego artykułu.

¹ A. Kim i in., *Cyber Attack Vulnerabilities Analysis for Unmanned Aerial Systems*, American Institute of Aeronautics and Astronautics, <https://engineering.purdue.edu/HSL/uploads/papers/cybersecurity/cyber-attack-lit-review.pdf>.

² K. Hartman, Ch. Steup, *The Vulnerability of UAVs to Cyber Attacks – An Approach to the Risk Assessment*, NATO CCD COE Publications, s. 7-8, https://ccdcoe.org/cycon/2013/proceedings/d3r2s2_hartmann.pdf.

³ *Unmanned Systems Integrated Roadmap FY 2013-2038*, Department of Defense, s. 15-16, <http://archive.defense.gov/pubs/DOD-USRM-2013.pdf>.

⁴ *Joint Doctrine Note 2/11; The UK Approach to Unmanned Aircraft Systems*, Ministry of Defence, s. 5-10, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/33711/20110505JDN_211_UAS_v2U.pdf.

Autonomia a artefakty

Zastosowanie pojęcia autonomii w stosunku do artefaktów jest przedmiotem wielu kontrowersji. Termin ten bowiem pochodzi z języka greckiego i powstał z połączenia słów *autos* – samo, *nomos* – prawo. Dosłowne rozumienie autonomii oznacza zatem, że byt autonomiczny jest w stanie nałożyć na samego siebie własne normy. Tak właśnie zakłada klasyczna filozoficzna koncepcja autonomii – autonomicznym jest ten, kto ma zdolność do przestrzegania prawa moralnego, które sam sobie narzucił (Smith, Kraemer 2006, s. 74). W sensie religijnym autonomia oznacza wolną wolę, nawet jeśli byt autonomiczny został stworzony przez wyższą istotę. Byt ten może stać się autonomiczny względem swojego twórcy, jeśli twórca wyposaży go w możliwość rozwijania (i dyspozycje do rozwijania) wolnej woli. Silne i normatywne rozumienie autonomii oznacza zdolność do wylamania się z prawa moralnego kreatora i wdrożenia swojego własnego. Autonomia jest więc relacją pomiędzy stwórcą (kreatorem) a stworzeniem.

Co to oznacza w przypadku maszyn? Autonomiczne artefakty powinny posiadać zdolność do rozwijania i podążania za prawem moralnym niezależnym od ich kreatora, nawet przeciwstawnym mu. Maszyny w pełni autonomiczne w sensie normatywnym musiałyby charakteryzować się wolną wolą. Takie rozumienie autonomii, zwłaszcza w kontekście sztucznej inteligencji, jest odmienne od rozumienia funkcjonalnego, typowego dla dyskusji o autonomicznych robotach.

Zupełna niezależność robotów od ludzkiej kontroli, mimo iż jest popularnym i nośnym motywem literatury i filmu *science-fiction*, pozostaje w najlepszym razie kwestią odległej przyszłości (wypadałoby powiedzieć „najgorszym”, bo obecne w kulturze wizje rebelii robotów zazwyczaj zakładają podporządkowanie lub zagładę rodzaju ludzkiego). Pełnię autonomii rezerwujemy bowiem tylko dla istot świadomych. Trzeba mieć przy tym na uwadze fakt, że intersubiektywne twierdzenie o posiadaniu lub nieposiadaniu przez inną istotę świadomości jest praktycznie niemożliwe. Niepozostawiająca wątpliwości weryfikacja, czy dany byt (np. maszyna) posiada świadomość jest niewykonalna (według Alana Turinga, można by to zrealizować tylko, gdyby było się tą maszyną), a konsekwentne żądanie takiej weryfikacji prowadzi do solipsyzmu, czyli stwierdzenia, że rzeczywistość jest jedynie zbiorem subiektywnych wrażeń podmiotu poznającego (Piesko 2002, s. 84). W kontekście maszyn możemy obecnie mówić jedynie o tym, że przejawiają one „autonomiczno-podobne” funkcje. Funkcjonowanie autonomiczne będzie więc oznaczało „zdolność systemu, platformy czy oprogramowania, do realizacji zadania bez ludzkiej interwencji, przy użyciu zachowań będących wynikiem interakcji pomiędzy programem komputerowym a zewnętrznym środowiskiem” (Kuptel, Williams 2014, s. 9).

Autonomia a automatyzacja

Definicje autonomii odnoszące się tylko do działania przy braku interwencji operatora są zdecydowanie niewystarczające. Przykładem tego sposobu definiowania autonomii jest definicja Departamentu Obrony USA, która zakłada, że autonomia oznacza zdolność do działania w realnym środowisku bez żadnej formy kontroli przez dany okres Lin, Bekey, Abney 2008, s. 103). Kryterium czasowe określa, jak długo maszyna funkcjonować może bez odwołania się do ludzkiej interwencji, np. wskazania celu bądź autoryzacji ataku (Thomas 2015, s. 137). W tym przypadku maszyna zdalnie sterowana charakteryzować się będzie zerową autonomią, a – przykładowo – miny autonomią nieskończoną. Mogą one bowiem działać, po tym jak zostaną uzbrojone, zupełnie bez odwoływania się do ludzkiej autoryzacji.

W takim ujęciu zarówno prosta mina, jak i wyrafinowany system pokroju fikcyjnego HAL-a z filmu *2001 Odyseja kosmiczna* Stanleya Kubricka, w równej mierze spełniałyby kryteria autonomii. Tego rodzaju podejście wynika z utożsamienia dwóch różnych terminów – autonomii i automatyzacji. Zdolność do działania bez interwencji operatora charakteryzuje bowiem już systemy określane jako automatyczne. Czym zatem różni się autonomia od automatyzacji i w jakiej relacji względem siebie znajdują się te dwa pojęcia?

Oba pojęcia odnoszą się do procesów, które mogą być wykonywane od początku do końca bez ludzkiej interwencji. Automatyzacja to jednak tylko i wyłącznie działanie bez ludzkiej interwencji, nieuwzględniające zdolności do samokierowania i podejmowania decyzji. Jest to rutynowe realizowanie sekwencji krok po kroku. Z kolei systemy autonomiczne są niezależne w ustalaniu i realizowaniu swoich własnych zadań (Marra, McNeil 2012, s. 1150). Oznacza to posiadanie pewnego poziomu zdolności do adaptowania się do kompleksowych i nieprzewidywalnych sytuacji. Odróżnia to autonomię od automatyzacji, która jest prostą zależnością typu wejście-zestaw reguł-wyjście, co powoduje, że zachowania są zdeterminowane i w dużym stopniu przewidywalne. Autonomię w kontekście automatyzacji rozumieć można dwojako (Johnson 2014):

1. Autonomia jako krańcowa automatyzacja. Automatyzacja zakłada mechanizację zadania – jest ono postrzegane jako sekwencja sformalizowanych operacji, które maszyna jest w stanie wykonać samodzielnie. Krańcowa automatyzacja oznacza, że maszyna jest w stanie samodzielnie wykonać wszystkie lub prawie wszystkie operacje. Na drugim końcu skali znajdują się systemy, które wymagają ingerencji (kontroli, podejmowania decyzji) w czasie wykonywania tych operacji. Według tego rozumienia autonomia ma swoje granice, a człowiek zawsze kontroluje maszynę – nawet jeśli nie znajduje się w pętli decyzyjnej (nie jest operatorem), pozostaje tym, który zaplanował cały proces i poszczególnie kroki w jego ramach (jest kreatorem).

2. Autonomia jako pojęcie odmienne od automatyzacji. Autonomiczne systemy są odmienne względem automatycznych, gdyż ich zachowanie nie jest

zaprogramowane. Autonomia dopuszcza co prawda funkcjonowanie w oparciu o instrukcje dotyczące tego, co i w jaki sposób maszyna ma robić, ale jest to sytuacja zupełnie odmienna niż zaplanowanie wszystkich sekwencji działania. Instrukcje pozwalają jedynie na ogólne zarządzanie funkcjonowaniem maszyny, a nie na jej kontrolowanie. Zachowanie maszyny, będące odpowiedzią na bodźce środowiska, nie może być w pełni wyspecyfikowane lub zaprogramowane. Autonomia zakłada podejmowanie własnych decyzji, chociaż w oparciu o reguły i ograniczenia. Z kolei systemy zautomatyzowane, chociaż działają niezależnie od zewnętrznego wpływu lub kontroli, podążają po wcześniej zdefiniowanej ścieżce, a ich zachowanie może być zawczasu w pełni wyspecyfikowane.

Można też spojrzeć na zależności autonomii i automatyzacji nieco inaczej i potraktować automatyzację jako niezbędny, ale niewystarczający, komponent autonomii. Automatyzacja, czyli zdolność do działania bez zewnętrznej interwencji, będzie więc pierwszym elementem autonomii. Drugim elementem będzie zdolność do podejmowania decyzji, która zakłada posiadanie woli (możliwości wyboru w działaniu lub myśleniu) i zamiaru (intencji, celowego dążenia do realizacji zadań). Prawdziwa autonomia wymaga jeszcze trzeciego komponentu, czyli zdolności do uczenia się – wyciągania konkluzji na podstawie doświadczeń i włączania ich w proces realizacji kolejnych działań (Marra, McNeil, op.cit. s. 1150-1151).

Stopniowanie autonomii

Podział systemów na autonomiczne i automatyczne jest oczywiście uproszeniem. Stopień autonomii scharakteryzować możemy nieco bardziej szczegółowo za pomocą kilku kryteriów, wskazanych przez Paula Scharre (2014, s. 13):

1. Stopień kontroli:

- a. systemy semiautonomiczne (*human in the loop*): możliwość działania automatycznego przez pewnego momentu (np. do wystąpienia sytuacji wymagającej podjęcia decyzji określonego rodzaju), potem konieczność ingerencji człowieka;
- b. systemy nadzorowane (*human on the loop*): możliwość nieprzerwanego działania autonomicznego, przy zachowaniu przez człowieka możliwości nadzorowania maszyny i ingerencji, kiedy uzna to za stosowne;
- c. systemy w pełni autonomiczne (*human out of the loop*): możliwość nieprzerwanego działania autonomicznego i brak możliwości ingerencji człowieka;

2. Kompleksowość maszyny:

- a. systemy automatyczne (*automatic*): prosta, mechaniczna odpowiedź na bodziec środowiska (np. w systemach wyzwalających wybuchy min);

b. systemy zautomatyzowane (*automated*): bardziej złożone systemy kształtujące zachowanie nie tylko w oparciu o bodźce środowiska, ale również na podstawie zespołu reguł, np. pojazdy zdolne do samodzielnej nawigacji i kierowania;

c. systemy autonomiczne (*autonomous*): systemy mające zdolność do samokierowania (*self-direction*), czyli takiego kształtowania swojego zachowania, które nie wypływa bezpośrednio ze sposobu zaprogramowania; możliwości te mogą być dodatkowo uzupełnione przez zdolność samouczenia (*self-learning*);

3. Typ funkcji, która jest zautomatyzowana. Trudno mówić o stopniu autonomizacji bez wskazania, o jakie działania chodzi, bo różne działania charakteryzują się różnym stopniem komplikacji i ryzyka wynikającego z ewentualnej błędnej decyzji, np. maszyna może być w pełni autonomiczna w zakresie nawigowania, ale zupełnie nieautonomiczna (czyli *human-controlled*) w zakresie wybierania miejsca, do którego zmierza.

Wyznaczenie precyzyjnych i nie budzących wątpliwości granic pomiędzy stopniami w ramach poszczególnych kryteriów jest niezmiernie trudne. Nie ma wyraźnej granicy między poziomami kompleksowości – systemami automatycznymi, zautomatyzowanymi, autonomicznymi. W uproszczeniu kontinuum rozciąga się od prostych układów funkcjonujących na zasadzie bodziec-reakcja do systemów zdolnych do wyciągania konkluzji na podstawie informacji na temat środowiska pozyskanych za pomocą sensorów. Różnica polega na kompleksowości myślenia – sprowadza się do określenia, kiedy mamy do czynienia, a kiedy nie, z analizowaniem informacji. Jest więc płynna i trudna do uchwycenia. W przypadku trzeciego kryterium trudno wręcz mówić o wyznaczeniu konkretnych stopni.

William C. Marra i Sonia K. McNeil (op. cit. s. 1151-1155) określają te kryteria inaczej, w nieco odmienny sposób rozkładając akcenty, natomiast wyróżnione przez nich trzy cechy określające autonomię, zasadniczo pokrywają się z cechami wskazanymi przez Scharre'a.

Cechy te to:

1. Niezależność – częstotliwość interwencji operatora niezbędna do funkcjonowania maszyny. Jeśli maszyna potrzebuje ciągłej interwencji operatora, jest po prostu zdalnie sterowana – to druga strona kontinuum, sytuująca się na przeciwnym biegunie względem pełnej autonomii. W owym kontinuum można wyróżnić wiele stadiów pośrednich, które najlepiej opisuje spektrum autonomii, opisane w dalszej części tekstu.

2. Dyskrecjonalność – możliwość podejmowania decyzji w sposób swobodny. Określa ona zdolność do zmiany planów operacyjnych w celu wypełnienia zadania bez interwencji operatora. Ta cecha decyduje o możliwości samodzielnego dostosowania środków do celów, przy założeniu, że cele są określone przez człowieka. Pełna autonomia może oznaczać również zdolność do wyznaczania celów.

3. Adaptacyjność – zdolność do skutecznego funkcjonowania w warunkach niepewnego i nieprzewidywalnego środowiska. Natężenie tej cechy

decyduje o tym, z jakimi rodzajami problemów maszyna jest w stanie sobie poradzić, do jakich nieplanowanych sytuacji oferowanych przez środowisko jest w stanie się dostosować.

Poszczególne stopnie autonomii w ramach danych kryteriów mogą pozostawać względem siebie we wzajemnie krzyżujących się relacjach. Przykładowo, dany system może być autonomiczny w realizacji funkcji nawigowania, nadzorowany w przypadku funkcji selekcji celów i semiautonomiczny, gdy chodzi o ich rażenie. Co więcej, dana funkcja może być realizowana w różnych trybach, np. gdy operator może przejść z trybu nadzorowanego na semiautonomiczny czy wręcz na tryb ręcznego, zdalnego sterowania.

Wewnątrz pętli

Mówiąc o obecności bądź nieobecności człowieka w pętli decyzyjnej (*loop*) mamy na myśli pętlę OODA (*observe-orient-decide-act*) Johna Boyda. Tłumaczy ona sposób podejmowania decyzji, zarówno człowieka (schemat stworzony został na potrzeby podejmowania decyzji przez pilotów myśliwskich w czasie walki powietrznej), jak i maszyny. Poszczególne elementy pętli to: obserwacja (prowadzona przy pomocy zmysłów w taki sposób, by zaabsorbować tak wiele informacji, jak to możliwe), orientacja (synteza tych informacji i konwersja tej wiedzy pod kątem decyzji, która ma być podjęta), decyzja („zważenie” dostępnych opcji i podjęcie decyzji), działanie w oparciu o podjętą uprzednio decyzję (Fadok 1994, s. 16).

Stopień autonomiczności maszyny można określić przez jej zależność od człowieka w zakresie podejmowania decyzji w ramach pętli. W debatach na temat autonomii dominuje uproszczone stanowisko, wedle którego człowiek jest albo w pętli albo poza nią. Pamiętać jednak należy, że cykle pętli wielokrotnie się powtarzają, niejednokrotnie zachodząc na siebie. Pętle mogą mieć różny zakres – mogą być węższe lub szersze w zależności od tego, jak szeroko zostało zdefiniowane zadanie stojące przed maszyną (Marra, McNeil, op. cit. s. 1179). Maszyny mogą charakteryzować się różnym stopniem autonomii w odniesieniu do poszczególnych elementów pętli decyzyjnej, np. być autonomiczne w obserwacji, a półautonomiczne w podejmowaniu decyzji. Samo podejmowanie decyzji również może być zróżnicowane, jeśli chodzi o stopień autonomii w zależności od tego, czego dotyczy decyzja (np. inaczej może być traktowana decyzja odnośnie do zmiany kierunku ruchu, a inaczej decyzja o otwarciu ognia).

Stopień kontroli człowieka nad maszyną, inaczej niezależność maszyny od operatora, w sposób bardziej szczegółowy przedstawia spektrum autonomii skonstruowane przez R. Parasuramana, T. Sheridana i Ch. Wickensa (2000, s. 286-297). Autorzy ci posługują się terminem automatyzacja, jednakże jego rozumienie jest zbliżone do rozumienia terminu autonomia prezentowanego w niniejszym opracowaniu. Spektrum to wyróżnia więcej stopni niż stosunkowo

prosta kategoryzacja na systemy półautonomiczne, nadzorowane i autonomiczne.

Tab. 1: Spektrum autonomii

Poziom	Charakterystyka
1	Człowiek wykonuje wszystkie czynności, komputer nie oferuje żadnej pomocy
2	Komputer oferuje zestaw alternatywach działań do wyboru, oraz...
3	Dokonuje selekcji, wybierając kilka z nich, albo...
4	Sugeruje jedną z alternatyw i...
5	Realizuje działanie zgodnie z sugestią, jeśli człowiek wyrazi zgodę, albo...
6	Daje człowiekowi określony czas na zablokowanie decyzji, po czym – jeśli człowiek nie zablokuje jej – w sposób automatyczny przystępuje do działania, albo...
7	Przystępuje do działania automatycznie, po czym przekazuje człowiekowi stosowną informację
8	Przekazuje informację po zrealizowaniu działania, ale tylko w przypadku, jeśli człowiek takiej informacji zażąda
9	Przekazuje informację po zrealizowaniu działania, ale tylko w przypadku, jeśli on sam – czyli komputer – uzna to za stosowne
10	Komputer decyduje o wszystkich działaniach i podejmuje je w sposób autonomiczny, zupełnie ignorując człowieka

Źródło: R. Parasuraman, T. Sheridan, Ch. Wickens 2000, s. 287.

Poziomy od 2 do 5 alokują zdolność decyzyjną pomiędzy człowiekiem a maszyną. Poziomy 6 do 9 dają inicjatywę maszynie, pozostawiając człowiekowi – tylko na pierwszym z nich – możliwość aprobaty lub zablokowania podjętej przez maszynę decyzji. Na pozostałych poziomach człowiek otrzymuje tylko coraz bardziej ograniczone prawo do informacji, a na ostatnim szczeblu zostaje całkowicie wykluczony z pętli decyzyjnej.

Autonomia a broń „inteligentna”

Autonomiczne systemy bojowe traktowane są często jako kolejne ogniwo rozwoju tzw. broni inteligentnej (*smart weapon*). Określenie „broń inteligentna” jest jednak pojęciem co prawda nośnym i działającym na wyobraźnię, ale niezbyt precyzyjnym. Pojęciem, które lepiej oddaje zasadę działania tego typu broni, jest określenie *precision-guided munition* (PGM) – amunicja precyzyjnego rażenia, a dosłownie amunicja precyzyjnie kierowana. Oznacza to broń, które używa układu poszukującego do namierzenia energii będącej odbiciem celu lub punktu celowania, przetwarzając ją w komendy kierujące broń do celu (Puckett 2004, s. 648). W przypadku broni inteligentnej kluczowy jest więc system kierowania. Wyróżniamy trzy rodzaje systemów kierowania, kluczowych dla broni precyzyjnego rażenia:

1. kierowanie programowe: pocisk porusza się po zaprogramowanej wcześniej trasie,

2. kierowanie zdalne: pocisk porusza się na podstawie sygnałów zewnętrznych,

3. samonaprowadzanie: pocisk działa w sposób autonomiczny.

Pojawia się dylemat, na ile pojęcia „broń inteligentna” i „autonomiczne systemy bojowe” są ze sobą tożsame. Czy każdą broń działającą za zasadzie samonaprowadzania nazwać można autonomiczną?

Typologia stosowana przez Marynarkę Wojenną Stanów Zjednoczonych wyróżnia trzy grupy uzbrojenia charakteryzującego się mniejszym lub większym poziomem „inteligencji” (Sharkey 2010, s. 377):

1. Broń działająca według przygotowanego wcześniej scenariusza (*scripted*): w tej kategorii mieści się broń kierowana działająca na zasadzie programowej (np. pociski manewrujące), na zasadzie samonaprowadzania (np. torpedy z samonaprowadzaniem czy rakiety funkcjonujące w oparciu o zasadę *fire&forget*, tzw. „odpal i zapomnij”), a także systemy przeciwlotnicze, w tym systemy typu CIWS (*Close-In Weapon Systems*, systemy obrony bezpośredniej, stosowane przede wszystkim na okrętach do obrony przed pociskami przeciwokrętowymi). Kluczowym czynnikiem jest działanie w ramach prostego schematu bodziec-reakcja, bez udziału myślenia deliberatywnego (czyli takiego, gdzie decyzja podjęta przez maszynę nie prostym rezultatem zarejestrowanego bodźca). Z tego powodu nie mogą być traktowane jako broń autonomiczna, nawet jeśli cechują się wysokim stopniem niezależności od operatora (jak w przypadku systemów samonaprowadzających się). Rodzaj systemu kierowanie nie determinuje więc wprost, czy dana broń mieści się w kategorii broni autonomicznej czy też nie.

2. Broń nadzorowana (*supervised*): ta kategoria obejmuje zarówno systemy nadzorowane (w myśl kryterium stopnia kontroli Scharre’a), jak i „niższą” kategorię systemów semiautonomicznych (czyli systemy *human in the loop* oraz *human on the loop*). Broń nadzorowana realizuje wszystkie bądź wybrane funkcje z pomocą operatora, który ma podejmować

decyzje. Do tej kategorii zaliczają się np. uzbrojone drony, jak MQ-1 Predator czy MQ-9 Reaper. Spełniają one więc tylko częściowo – na tyle, na ile potrafią działać niezależnie – kryteria pozwalające zaliczyć tego typu broń do grupy broni autonomicznej.

3. Broń inteligentna (*intelligent*): w tej kategorii broni inteligencja maszyny jest wykorzystywana do podejmowania decyzji, postrzegania i interpretowania znaczenia informacji pozyskanych z sensorów, rozwiązywania problemów i współpracy z innymi systemami. Pozostaje pytanie, czy broń spełniająca tak wyśrubowane wymagania została już skonstruowana.

Autonomiczne systemy bojowe postrzegać można również szczytowy etap rozwoju broni inteligentnej, w myśl kontinuum zaprezentowanego przez Armina Krishana w książce *Killer Robots* (2009). Autor wyróżnia następujące rodzaje broni:

1. Broń zrobotyzowana – skomputeryzowana, wyposażona w sensory, zdalnie sterowana lub autonomiczna, np. pociski manewrujące; trzeba zaznaczyć, że autor postrzega roboty w sposób ogólny jako maszyny programowalne, mogące charakteryzować się różnym stopniem niezależności od operatora (od zdalnego sterowania, poprzez nadzór, aż do autonomii).

2. Broń bezzalógowa – zrobotyzowana platforma bojowa możliwa do powtórnego wykorzystania, nie przeznaczona do misji „samobójczych”, w czasie których byłaby niszczone, np. drony (ale z wyłączeniem np. precyzyjnie naprowadzanych pocisków).

3. Broń autonomiczna – broń skomputeryzowana, będąca w stanie samodzielnie, bez ludzkiej ingerencji, wybierać i razić cele.

Typologia ta w sposób syntetyczny, choć nieco uproszczony sposób prezentuje linię rozwojową broni inteligentnej. Poszczególne kategorie nie charakteryzują się jednak niewyłącznością, a zastosowanie kryterium wielorazowości użycia do wyróżnienia drugiej kategorii może budzić wątpliwości (z punktu widzenia zdolności maszyny do niezależnego działania w złożonym środowisku nie ma znaczenia fakt, czy maszyna ta jest możliwa do jednokrotnego czy wielokrotnego użycia).

Kryteria autonomii

Podsumowując omówione powyżej zagadnienia można pokusić się o wyszczególnienie zespołu cech, którymi powinien charakteryzować się autonomiczny system bojowy.

System ten powinien łącznie spełniać trzy warunki. Powinien być:

1. Niezależny
2. Asertywny
3. Adaptacyjny

Ad. 1. Rozpatrując kryterium niezależności szczególną uwagę należy zwrócić na systemy nadzorowane (*human on the loop*). Jeśli systemy zdalnie sterowane (*remote controlled*) oraz semiautonomiczne (*human in the loop*) w sposób oczywisty nie mogą być traktowane jako autonomiczne, a systemy *human out of the loop* sytuują się na przeciwnym biegunie kontinuum, to sytuacja z systemami nadzorowanymi nie jest jednoznaczna.

Z jednej strony obecność człowieka w pętli decyzyjnej, jego kontrola nad wyborem i porażeniem celów, sugerowałaby, że tego typu systemy należy wykluczyć z grupy systemów autonomicznych. W myśl tego rozumowania różnica między systemami półautonomicznymi a nadzorowanymi jest w zasadzie niewielka – w obu przypadkach mamy do czynienia z kontrolą nad działaniami maszyny sprawowaną przez człowieka. Polega ona albo na konieczności autoryzacji ataku (systemy semiautonomiczne), albo może na możliwości jego powstrzymania (systemy nadzorowane). Różnica ta pod względem formalnym jest stosunkowo niewielka (w obu przypadkach mamy do czynienia z kontrolą człowieka), pod względem funkcjonalnym zaś jest

ogromna. Może się bowiem okazać, że na dynamicznym polu walki, z powodu krótkiego czasu pozostawionego człowiekowi na ewentualne weto, prawo do jego stosowania będzie iluzoryczne, w praktyce jedynie nominalne (Crootof 2015, s. 1849).

Żeby więc uznać system nadzorowany za zdecydowanie nieautonomiczny (np. dla potrzeb prawnej regulacji tych systemów) trzeba by założyć, że systemy takie w każdej sytuacji dają człowiekowi odpowiedni czas na ewentualną decyzję wstrzymującą atak. Jak jednak precyzyjnie określić „odpowiedni czas”? Czyż nie zależy to od celu i sytuacji? Na te pytania nie ma jednoznacznej odpowiedzi.

W takim razie, różnica między systemami kontrolowanymi przez człowieka (czyli nadzorowanymi), a systemami w pełni autonomicznymi, nie okazuje się tak znacząca, jak się to pozornie wydaje. Szczególnie z konstrukcyjnego punktu widzenia jest ona niewielka – systemy nadzorowane i tak muszą mieć wbudowaną zdolność do autonomicznego działania, tylko nie zawsze musi ona być wykorzystana. Definicja autonomii nie powinna więc czynić różnicy między tymi trybami, bo w każdym przypadku osiągnięto zdolność do autonomicznego działania.

Biorąc pod uwagę kryterium niezależności za systemy zdalnie sterowane uznamy np. niemieckie bomby kierowane z okresu II wojny światowej Fritz X, a także pierwsze pokolenie przeciwpancernych pocisków kierowanych z lat 50. i 60. – francuskie SS.10 i ENTAC, szwajcarsko-niemieckie Cobra, radzieckie 3M6 Trzmiel, 3M11 Falanga i 9M14 Maultka (generalnie systemy te odchodzą w przeszłość). Systemy semiautonomiczne to np. uzbrojone drony w rodzaju Predatora i Reapera. Przykładami systemów nadzorowanych będą głównie systemy przeciwlotnicze i przeciwrakietowe, np. amerykański system okrętowy Aegis, izraelski system przeciwrakietowy Iron Dome, a także rozwijane systemy C-RAM (*Counter Rocket, Artillery, and Mortar* – systemy zwalczające rakiety oraz pociski artyleryjskie i moździerzowe, np. amerykański Centurion i niemiecki MANTIS) oraz okrętowe systemy typu CIWS, np. amerykański Phalanx czy holenderski Goalkeeper). Systemami w pełni autonomicznymi będą miny, rakiety samonaprowadzające się odpalane na zasadzie *fire&forget*, przeciwpancerne systemy artyleryjskie z głowicami samonaprowadzającymi (szwedzko-francuski BONUS, niemiecki SMArt, amerykański SADRAM) oraz hipotetyczne autonomiczne roboty bojowe. Z punktu widzenia kryterium niezależności do systemów autonomicznych zaliczają się dwie ostatnie grupy.

Ad. 2. Precyzyjne określenie takiego stopnia asertywności maszyny, od którego uznać ją za autonomiczną jest zadaniem niezmiernie skomplikowanym. Przypomnijmy, że chodzi o to, na ile maszyna jest zdolna do podejmowania decyzji w sposób swobodny, na ile może samodzielnie decydować o sposobie wypełnienia zadania. Z konieczności przyjęto więc uproszczone, ale możliwe do stosunkowo precyzyjnej operacjonalizacji kryterium. Za maszyny autonomiczne uznano te, gdzie operator nie wskazuje konkretnego celu. System ma niszczyć

cele o danych charakterystykach (mogą to być np. pojazdy opancerzone albo osoby mające przy sobie broń). Z kolei systemy charakteryzujące się specyficznym wskazywaniem celów przez operatora – wskazuje on konkretny cel – np. dany pojazd czy daną osobę – nie mogą być uznane za autonomiczne. Z tego punktu widzenia np. pociski manewrujące rażące zaprogramowany wcześniej cel nie będą autonomiczne. Za nieautonomiczne uznamy także rakiety *fire&forget* działające w trybie LOBL (*lock on before launch* – wskazywanie celu przed odpaleniem) np. rakiety typu powietrze-powietrze. Rakiety samonaprowadzające działające w trybie LOAL (*lock on after launch* – wskazywanie celu po odpaleniu) sytuują się na pograniczu autonomii, np. rakiety przeciwokrętowe bywają odpalane w kierunku wykrytej wcześniej grupy okrętów, ale samodzielnego wyboru ostatecznego celu dokonuje głowica samonaprowadzająca. Podobnie jest z artyleryjską amunicją przeciwpancerną, z tym, że tu już regułą jest odpalenie w kierunku grupy celów, a nie celu pojedynczego. Z kolei systemy defensywne, np. przeciwlotnicze czy przeciwrakietowe, a także uzbrojeni roboty-strażnicy (np. stosowany do nadzorowania strefy zdemilitaryzowanej na Półwyspie Koreańskim Samsung Techwin SGR-A1) zdecydowanie są systemami autonomicznymi z tego punktu widzenia.

Ad. 3. Kryterium zdecydowanie najtrudniejszym do precyzyjnego stopniowania jest adaptacyjność. Generalnie można przyjąć tu podział na systemy reaktywne oraz refleksyjne (deliberatywne). Gdzie wyznaczyć granicę między nimi? Pewną odpowiedzią może być stopień ustrukturyzowania środowiska, w ramach którego systemy te działają. Systemy działające w środowiskach wysoce ustrukturyzowanych nie muszą rozważać wielu zmiennych, więc działają raczej w oparciu o zasadę bodziec-reakcja, a nie w oparciu o konkluzje wyciągane na podstawie analizy tych zmiennych (czyli w oparciu o myślenie). Nie chodzi tu tylko o stopień komplikacji środowiska, w którym funkcjonuje maszyna (np. operowanie w przestrzeni powietrznej jest mniejszym wyzwaniem dla maszyny niż działanie w środowisku lądowym), ale także o rodzaj wykonywanych zadań (w tym przede wszystkim o rodzaj zwalczanych celów). System mający za zadanie samodzielnie wykryć i zniszczyć pojazd opancerzony czy nadlatujący samolot będzie zdecydowanie bardziej reaktywny niż system mający np. zidentyfikować i wyeliminować wrogich rebeliantów noszących cywilnie ubrania. Podobnie, łatwiej wyobrazić sobie bezzalogowy bombowiec strategiczny (zadanie, które ma do wykonania jest stosunkowo proste) niż autonomiczny samolot wielozadaniowy, zdolny do stawienia czoła nieprzewidywalności pola walki.

Z tego punktu widzenia archetypem systemu reaktywnego będą miny, za systemy reaktywne można uznać też rakiety samonaprowadzające oraz przeciwpancerne systemy artyleryjskie. Do tej grupy zaliczymy także systemy typu CIWS (niszczą wszystkie nadlatujące rakiety) czy, już z pewnymi wątpliwościami, system Iron Dome (niszczy nadlatujące rakiety, jeśli kierują się

na obszary zamieszkane; system prowadzi więc już pewien rodzaj analizy). Na pograniczy systemów refleksyjnych sytuują się kompleksowe systemy przeciwlotnicze, np. Aegis – mogą one operować w skomplikowanym środowisku, zwłaszcza jeśli funkcjonują w sytuacji kryzysu, a nie otwartego konfliktu. Klasyfikacja potencjalnych celów nie jest zadaniem łatwym i nawet, jeśli mamy do czynienia z systemem nadzorowanym i ostateczną decyzję podejmuje człowiek, może dochodzić do tragicznych pomyłek. Do takiej doszło m.in. 3 lipca 1988 roku, kiedy omyłkowo zestrzelono irański samolot pasażerski z 290 osobami na pokładzie przez wyposażony w system Aegis amerykański krążownik USS Vincennes (Rochlin 1989).

Wnioski

Autonomizacja bezzałogowych systemów bojowych jest jednym z najbardziej obiecujących kierunków rozwojowych techniki wojskowej. Autonomia to jednak nie tylko niezależność od zewnętrznego operatora, ale przede wszystkim pewny poziom zdolności do adaptowania się do kompleksowych i nieprzewidywalnych sytuacji występujących w realnym środowisku. W sposób oczywisty określenie „pewny poziom” jest pojęciem nieostrym i dokonano jego przybliżenia za pomocą kryteriów niezależności, adaptacyjności i asertywności.

Systemy bojowe spełniające kryteria autonomii istnieją i są używane przez siły zbrojne. Jak dotychczas ich proliferacja jest stosunkowo ograniczona – są to tylko systemy defensywne i stacjonarne (w tym sensie, że niezdolne do autonomicznego poruszania się – niektóre systemy, np. okrętowe, są bowiem zamontowane na mobilnych platformach, ale mobilność nie jest w ich przypadku funkcją poddaną autonomizacji). Z prawdziwą rewolucją będziemy mieli do czynienia wówczas, gdy pojawią się systemy mobilne i ofensywne, spełniające przy tym wszystkie wymienione kryteria autonomii. Wydaje się jednak, że nawet w przypadku systemów powietrznych, których rozwój jest najbardziej zaawansowany, jest to jednak perspektywa stosunkowo odległa.

Literatura:

- Crootof, R., 2015, *The Killer Robots Are Here, Legal and Policy Implications*, *Cardozo Law Review*, 36: 1837.
- Fadok, D. S., 1994, *John Boyd and John Warden: Air Power's Quest for Strategic Paralysis*, Air University Press, Maxwell.
- Gajzler, M., 2015, *Bezzałogowa przyszłość?*, *Dziennik zbrojny. Analiza*, nr 1.
- Johnson, D. G., Noorman M. E., 2014, *Responsibility Practices in Robotic Warfare*, *Military Review*, maj-czerwiec.
- Kopeć, R., 2014, *Revolution in Military Affairs as Technology-Driven Change in Military Operation*, w: *Encyclopedia of Information Science and Technology, Third Edition*, Khosrow-Pour M. (red.), Hershey, s. 6534-6542.

- Koziej, S., 2008, *Kierowanie bezpieczeństwem narodowym. Skrypt internetowy*, Warszawa.
- Krishan, A., 2009, *Killer Robots: Legality and Ethicality of Autonomous Weapon*, Ashgate Press, Farnham.
- Kuptel, A., Williams, A. (red.), 2014, *Role of Autonomous Systems in Gaining Operational Access. Policy Guidance – Autonomy in Defence Systems*, Supreme Allied Commander Transformation HQ, Norfolk.
- Lin, P., Bekey, G., Abney, K., 2008, *Autonomous Military Robotics: Risk, Ethics, and Design*, US Department of Navy, Office of Naval Research, Waszyngton.
- Marra, W.C., McNeil, S.K., 2012, *Understanding “the Loop”: Regulating the Next Generation of War Machines*, Harvard Journal of Law & Public Policy, 36.
- Parasuraman, R., Sheridan, T., Wickens, Ch., 2000, *A Model for Types and Levels of Human Interaction with Automation*, IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 30, 3.
- Piesko, M., 2002, *O subtelnej różnicy między słabą a mocną wersją koncepcji sztucznej inteligencji na przykładzie tekstu Turinga*, Zagadnienia filozoficzne w nauce, nr XXXI.
- Puckett, Ch. B., 2004, *In this era of “smart weapon”. Is a state under an international legal obligation to use precision-guided technology in armed conflict*, Emory International Law Review, 18, 2.
- Rochlin, G. I., 1989, *Iran Air Flight 655 and the USS Vincennes. Complex, Large-Scale Military Systems and the Failure of Control*, w: *Social Responses to Large Technical Systems, Control and Anticipation*, red. T. R. La Porte, Berkeley, s. 99-125.
- Scharre, P., 2014, *Robotics on the Battlefield. Part I: Range, Persistence and Daring*, Waszyngton.
- Sharkey, N., 2010, *Saying ‘No?’ to Lethal Autonomous Targeting*, Journal of Military Ethics, 9, 4.
- Smith, C. T. A., Kraemer, F., 2006, *Robots, Dennett and the autonomous: a terminological investigation*, Minds and Machines, 16.
- Thomas, B. T., 2015, *Autonomous weapon systems: The anatomy of autonomy and the legality of lethality*, Houston Journal of International Law, 37, 1.
- Windeck, A., 2014, *Preface*, [w:] *Robots on the Battlefield, Contemporary Issues and Implications for the Future*, R. Doaré i in. (red.), Fort Leavenworth.

Summary

Autonomy of combat systems

Autonomous combat systems are one of the most promising directions of weapon development. Weapons' autonomization could bring the vision of robolution to the reality, which means the possibility of waging war with a wide range of robots. This is a crucial issue when it comes to the forecast of the military development. However, there is

a danger of oversimplifying this development due to terminological loophole. The aim of the article is to determine the essence of autonomy of combat systems. It raises the questions about the conditions of a machine's autonomy, about the discrimination between autonomy and related terms, about the steps of autonomy and its criterions in regard to combats systems. Finally, it answers the question whether autonomous combat systems are an element of the reality or a matter of future.

Key words: autonomy, combat robots, robolution, intelligent weapon, decision cycle.