

Etyka

* Opisz różnicę między moralnością operacyjną, moralnością funkcjonalną i pełną sprawczością moralną i określ odpowiedzialność projektanta za każdą z nich.

* Wymień cztery rodzaje agentów etycznych (wpływ etyczny, ukryty, wyraźny, pełny) i podaj przykład każdego z nich.

* Opisz co najmniej jeden problem z każdym z Trzech Zasad Robotyki Asimova jako wskazówką do opracowania robotów moralnych.

Przegląd

Wyraźnie widoczne użycie dronów przez wojsko USA zwiększyło dyskomfort społeczny związany z robotami i przyspieszyło rozwój etyki robotów. Opinia publiczna zadaje pytania takie jak: Czy uzbrajanie robotów w broń nie jest nieetyczne? Ta kwestia jest gorąco dyskutowana w kręgach sztucznej inteligencji. Kolejne częste pytanie opinii publicznej, czasami stwierdzenie brzmi: jeśli programiści zastosują Trzy Prawa Asimova, czy to nie wystarczy, aby uchronić nas przed obrażeniami, śmiercią lub powstaniem? To pytanie odnosi się do odwiecznie popularnych opowieści science fiction Asimova „Ja, Robot”, ale ignoruje fakt, że historie te dotyczyły tego, w jaki sposób te trzy prawa powodowały konflikty i niezamierzone konsekwencje. Trzecie pytanie postawione przez opinię publiczną brzmi: czy nie jest nieetyczne tworzenie robotów i traktowanie ich jak niewolników? Te pytania stają się tak powszechne, że Unia Europejska przeprowadziła niedawno wieloletnie badanie RoboLaw. W badaniu przeanalizowano, czy roboty były zwolnione z przepisów dotyczących produktów i ludzi, a tym samym wymagały specjalnych przepisów, a jeśli tak, to jakie te przepisy miałyby być. Odpowiedzi na te trzy pytania mają konsekwencje dla projektowania systemów robotów. Pojawił się czwarty temat etyki dotyczący społecznego wpływu zmiany miejsca pracy w wyniku zwiększonego korzystania z robotów, ale ponieważ temat ten nie jest wyraźnie problemem projektowania robotów, wykracza poza nasz zakres

Projektanci często wikłają się w spekulacje na temat przyszłych możliwości robotów („Pewnego dnia, gdy roboty będą miały pełną autonomię...”) i zapominają o swoich zawodowych obowiązkach projektowania bezpiecznych systemów dla dowolnych możliwości autonomicznych, które są zgodne z dobrymi czynnikami ludzkimi i zasadami inżynierii. Na przykład niedawna katastrofa samochodu Tesla najwyraźniej miała miejsce w sytuacji, gdy autonomia nawigacyjna zawiodła, a człowiek nie przejął kontroli; wydaje się to być przykładem problemu kontroli człowieka poza pętlą, opisanego w Części 18. W tym przypadku sądy nie pociągnęły producenta do odpowiedzialności za zaprojektowanie systemu, który naruszał dobrze znane zasady dotyczące czynnika ludzkiego, ale z pewnością podnosi kwestię etyki zawodowej, która jest domeną moralności operacyjnej.

Względy etyczne sztucznej inteligencji i robotyki są motywowane co najmniej czterema kategoriami zastosowań, których użycie wiąże się z naszym postrzeganiem moralności. Jedną kategorią to aplikacje, w których roboty są zaprojektowane do „czynienia dobra”, a zatem społeczeństwo jest zobowiązane do ich używania, tak jak pracownicy służby zdrowia byli zobowiązani do korzystania z penicylin i maszyn płucno-sercowych. Dżokeje-roboty na wielbłądach, które zastępują dzieci i roboty ratownicze, to dwa przykłady robotów, które mają „czynić dobro”. Inna kategoria dotyczy robotów, które są używane, ponieważ są moralnie lepszymi agentami. Roboty straży granicznej w Korei Południowej obserwujące strefę zdemilitaryzowaną nie dają się przekupić i nie zawahają się zareagować na intruzów. Są więc bardziej niezawodni niż ludzcy strażnicy i moralnie lepsi. Trzecia kategoria obejmuje aplikacje, w których oczekujemy, że roboty będą nas odpowiednio traktować. Ta kategoria jest szczególnie widoczna w debacie o uzbrojonych robotach, ale obejmuje ona również sytuacje, w których

autonomiczny samochód może zdecydować się na działanie, które rani cię, ale ratuje wiele innych. Innym przykładem odpowiedniego leczenia jest oszustwo: Czy dopuszczalne jest, aby robot kłamał, abyśmy zachowywali się w naszym najlepszym interesie? Czwarta kategoria jest odwrotnością trzeciej; zawiera aplikacje, w których powinniśmy odpowiednio traktować roboty. Na przykład, w którym momencie robot zajmujący się opieką zdrowotną i usługami przechodzi od bycia inteligentnym asystentem do bycia niewolnikiem?

W tej Części postaram się dać czytelnikowi ramy do myślenia o etyce projektowania. Najpierw definiuje etykę i etykę zawodową. Przedstawia cztery kategorie podmiotów etycznych Moora. Klasyfikacja sugeruje, że większość autonomicznych robotów prawdopodobnie nie będzie bezpośrednio zajmowała się etyką lub będzie działać jako w pełni etyczni podmioty. Kolejne cztery sekcje obejmują moralność i hierarchię moralną zaproponowaną przez Wallacha i Allena: moralność operacyjną, moralność funkcjonalną i pełną sprawczość moralną. Obejmuje aktualny stan wiedzy w zakresie projektowania funkcjonalnych i pełnych czynników moralnych. Część następnie analizuje trzy prawa Asimova i wyjaśnia, dlaczego nie są one wystarczające do dobrego projektu. Następnie Część przeformułowuje etykę pod kątem rodzajów sztucznej inteligencji potrzebnej do wspierania projektów etycznych lub agentów etycznych.

Rodzaje etyki

ETYKA ZAWODOWA

Etyka w robotyce ma dwa znaczenia. W sensie filozoficznym „Etyka jest gałęzią filozofii zajmującą się tym, jak należy określać wartości moralne, w jaki sposób można osiągnąć moralny skutek w określonych sytuacjach, jak rozwija się zdolność moralna lub sprawczość moralna i jaka jest jej natura oraz jakie wartości moralne postępują ludzie i trzymaj się.” Drugim znaczeniem etyki w robotyce jest etyka zawodowa. Etyka zawodowa „dotyczy kwestii moralnych, które wynikają ze specjalistycznej wiedzy, którą zdobywają profesjonaliści, oraz tego, w jaki sposób należy zarządzać wykorzystaniem tej wiedzy podczas świadczenia usług na rzecz społeczeństwa”. Projektanci robotów muszą być świadomi obu rodzajów etyki. Podczas gdy publiczne dyskusje skupiają się przede wszystkim na filozoficznych aspektach robotów, projektanci są odpowiedzialni za przestrzeganie etyki zawodowej ze względu na kwestie odpowiedzialności. Może to być skomplikowane, ponieważ nie istnieje „zawód inteligentnego robotyka”, a zatem etyka zawodowa specyficzna dla robotyki nie została jasno ustalona. Jednak robotyka AI opiera się na zawodach, które mają etykę, takich jak inżynieria, informatyka, psychologia, a także obszary zastosowań medycyny i pielęgniarstwa, które również mają etykę zawodową. Organizacje zawodowe promują etykę w swoich dziedzinach, a programy przyznające stopnie naukowe w tym zawodzie oferują zajęcia lub moduły z etyki. Na przykład Association for Computing Machinery ma 24-częściowy kodeks etyki i postępowania zawodowego. Na przykład część 2.5 kodeksu stanowi: „Przedstaw kompleksowe i dokładne oceny systemów komputerowych i ich wpływu, w tym analizę możliwych zagrożeń”. Kod ACM wyjaśnia, że stwierdzenie, że robot jest autonomiczny, a zatem projektant nie ponosi odpowiedzialności za jakiegokolwiek odchylenia, nie jest profesjonalne. Części 3.4 i 3.5, „Upewnij się, że użytkownicy i osoby, na które system będzie miał wpływ, mają jasno wyartykułowane potrzeby podczas oceny i projektowania wymagań; później system musi zostać poddany walidacji w celu spełnienia wymagań” oraz „Wypowiedzenie i wsparcie polityk, które chronią godność użytkowników i innych osób dotkniętych systemem komputerowym” są szczególnie istotne. Prawdziwy informatyk od samego początku włączyłby interakcję człowieka i robota w projekt robota.

Kategoryzacja pełnomocników etycznych

Z pragmatycznego punktu widzenia projektant musi określić, kiedy roboty muszą być etyczne i w jakim stopniu muszą być zaprogramowane, aby były etyczne. Zarówno Moor, Wallace, jak i Allen uważają

robota za podlegający gdzieś spektrum rozważań etycznych. Spektrum odzwierciedla omówienie autonomii operacyjnej w Części 4, gdzie robot może mieć dobrze ograniczone możliwości autonomiczne lub może mieć pozwolenie na rozumowanie i rozluźnianie ograniczeń.

Cztery kategorie Moora

Moor przedstawia cztery kategorie agentów w odniesieniu do etyki. Kategorie te są pomocne projektantowi w formułowaniu etycznych oczekiwań robota i dyskusji o moralności.

Cztery kategorie Moora to:

PRZEDSTAWICIELE WPŁYWU ETYCZNEGO

Środki wpływu etycznego. Są to roboty przeznaczone do zastosowań wspierających, jak nazywa Moor, normy etyczne. Roboty poszukiwawczo-ratownicze mają pomagać w działaniach, które mają konsekwencje etyczne: przyspieszenie reakcji ratujących życie i ożywienie gospodarcze są pożądane i wartościowe. Zauważ, że robot nie oczekuje rozumowania na temat etyki; jest to robot etyczny po prostu ze względu na zastosowanie.

NIEWŁAŚCIWE PRZEDSTAWICIELE ETYCZNE

Niejawni agenci etyczni. Są to roboty, które nie są zaprogramowane do uwzględniania etyki, ale ich programowanie musi być zaprojektowane tak, aby zminimalizować negatywne skutki etyczne. Moor podaje przykład autopilotów w samolotach. Pierwotnie autopiloty zwracali kontrolę pilotom w momencie, gdy dochodzili do granic swojego działania. Ta nagła zmiana czasami zaskakiwała pilotów i prowadziła do wypadków. Jednak autopilot może wyczuć, kiedy zbliża się do swoich granic. Tak więc, ponieważ stawką jest życie, autopilot musi być zaprogramowany tak, aby ostrzegał pilotów, że być może będą musieli przygotować się do przejścia w najbliższej przyszłości. Innym przykładem jest ukryte oczekiwanie, że szpitalny robot dostarczający leki nie uderzy pacjentów chodzących o kulach, mimo że pisanie oprogramowania do wykrywania powierzchni o cienkich profilach, takich jak kule, jest trudniejsze niż wykrywanie ludzi i ścian.

WYRAŹNI AGENCI ETYCZNI

Wyraźne środki etyczne. Są to roboty, które mogą wnioskować o etycznych konsekwencjach działań. Wracając do przykładów we wstępie, można sobie wyobrazić, że samochód autonomiczny można zaprogramować tak, aby poświęcał swoich pasażerów w celu ocalenia większej liczby pasażerów w innych samochodach. Wydaje się, że w zasięgu ręki są autonomiczne samochody, które mogą dzielić się informacjami i, w ciągu milisekund, przewidywać prawdopodobieństwo śmierci i obrażeń.

PEŁNI AGENCI ETYCZNI

Pełni agenci etyczni. Są to roboty, które mogą podejmować, wykonywać i uzasadniać decyzje etyczne, w przeciwieństwie do jawnych agentów etycznych, które są ograniczone do określonych sytuacji i zakresów działań. Jak zauważa Moor, projekt w pełni etycznych agentów jest przedmiotem wielu dyskusji na temat etyki robotów, chociaż nie jest jasne, czy można zbudować robota z wymaganą ilością inteligencji i prawdziwą autonomią polityczną. Cztery kategorie dotyczą nie tylko definicji etyki, ale także definicji autonomii operacyjnej. Kategoria ukrytego agenta etycznego wzmacnia pojęcie etyki zawodowej. Inteligentny robot musi być bezpieczny i musi uwzględniać najlepsze praktyki i najwyższe standardy zawodu robotyka. Wyraźny agent etyczny podejmuje decyzje etyczne, ale w wyraźnych granicach. Realizuje swoje istniejące programowanie dotyczące etyki, podczas gdy pełnoprawny agent etyczny może generować nowe programy. Różnicę między pełnoprawnym a pełnoprawnym podmiotem etycznym można wyrazić za pomocą poziomów terminologii inicjatywnej z Części 4. W tej

terminologii w pełni sprawny podmiot etyczny ma autonomię przymusu, ponieważ inicjatywa może zmienić zasady lub granice swojego działania w odniesieniu do etyki, podczas gdy wyraźny podmiot etyczny ma jedynie autonomię systemowo-stanową.

Kategorie moralności

Wendall i Allen przedstawiają podobną, ale inną taksonomię niż Moor. Ich taksonomia odzwierciedla przejście od robota jako urządzenia lub narzędzia do w pełni autonomicznego agenta pod względem moralności, którą przejawia.

MORALNOŚĆ OPERACYJNA

* moralność operacyjna, gdzie odpowiedzialność i odpowiedzialność za moralną funkcję robota leży w gestii projektantów i użytkowników narzędzi. Odpowiedzialność i odpowiedzialność są ekwiwalentem etyki zawodowej i kładą nacisk na autonomiczne umiejętności, które zawsze w pewnym sensie wiążą się z moralnością. Nawet jeśli robot nie jest deterministyczny, niekoniecznie oznacza to, że etyka zawodowa i odpowiedzialność zawodowa są sprzeczne.

MORALNOŚĆ FUNKCJONALNA

* moralność funkcjonalna, gdzie robot ma autonomiczne zdolności (np. autopilot) lub pracuje w aplikacjach z wrażliwością etyczną (np. pomoce w podejmowaniu decyzji medycznych) nadal mieści się w etyce zawodowej, ponieważ autonomia jest funkcją projektu. Na przykład, jeśli zawiedzie zdolność autonomicznego prowadzenia samochodu i osoba zostanie zabita, projektant może zostać pociągnięty do odpowiedzialności.

PEŁNA SPRAWCZOŚĆ MORALNA

* pełna sprawczość moralna, gdzie robot jest godny zaufania i ma wysoce wyrafinowane zdolności deliberacyjne, które obejmują podejmowanie decyzji w sytuacjach etycznych.

Etyka programowania

Robot z funkcjonalną lub pełną sprawczością moralną będzie musiał zostać zaprogramowany, aby miał etykę. Nie wiadomo dokładnie, w jaki sposób zostałyby to wdrożone. Filozofowie spekulowali na temat tego, w jaki sposób można osiągnąć takie programowanie, analizując, w jaki sposób uważają, że ludzie nabywają etyki. Robotycy zaczęli przyglądać się bardziej realistycznym i praktycznym narzędziom obliczeniowym, ale nie ma konsensusu ani nowego trendu.

Podejścia z filozofii

Filozofowie dyskutowali o tym, jak ich zdaniem ludzie nabywają etyki przy założeniu, że roboty będą powielać ten proces.

KONSEKWENCJALIZM

DEONTOLOGIA

ETYKA WARTOŚCI

Istnieją trzy podejścia ogólnie: konsekwencjalizm, w którym robot uzasadnia konsekwencje opcji; deontologia, gdzie robot uzasadnia opcje pod względem swoich obowiązków, obowiązków i praw; oraz etyka cnót, w której robot rozwija dobry charakter, aby mógł reagować w najbardziej moralny sposób. Te trzy podejścia mogą uchwycić sposób, w jaki ludzie uczą się zachowywać etycznie, ale nie jest jasne, jak wdrożyć to algorytmicznie. Wszystkie trzy wymagają postępów w reprezentacji i rozumowaniu.

Chociaż etyka cnót może wydawać się najłatwiejsza, ponieważ wytwarza zachowania, a nie prowadzi rozważania, oznacza to, że do wytworzenia tych zachowań stosuje się rozumowanie analogiczne; rozumowanie analogiczne jest nadal poza zakresem praktycznej sztucznej inteligencji. Alternatywą dla metod odgórnych są metody oddolne, w których robot uczy się być dobrym. Jest to zbyt ogólne, aby mogło być przydatne dla projektantów. Przypomnij sobie z Części 16, że jest wiele rodzajów uczenia się i że uczenie się wymaga sprecyzowania tego, czego się uczy lub przynajmniej określenia wkładu do uczenia się nowych pojęć.

Podejścia od robotyki

Istniejące roboty zazwyczaj podlegają wpływowi etycznemu lub niejawną agencji etycznej. Obecne roboty nie zostały zaprogramowane z możliwościami rozważania potrzebnymi do bardziej wyrafinowanych, jednoznacznych reakcji etycznych. Nie oznacza to, że pewnego dnia aplikacja wezwie robota do działania z wyraźnym lub pełnym zaangażowaniem etycznym. Jednym z przykładów jest skupienie się Arkina na robotach jako potencjalnie zdolnych do prowadzenia operacji wojskowych w sposób bardziej moralny niż sposób, w jaki czasami postępują ludzie. Ludzie znani są z tego, że nadużywają swojej pełnej moralności do popełniania zbrodni wojennych, podczas gdy roboty, jako jawne podmioty moralne, nie przekraczają granic reguł zaangażowania. Roboty prawdopodobnie wykorzystywałyby deliberatywne algorytmy, takie jak etyczny gubernator.

Trzy prawa robotyki Asimova

Słynna seria opowiadań „Ja, Robot” Isaaca Asimova została osadzona w przyszłości, w której roboty rządzą się trzema zasadami, z których pierwsza to, że robot nie może wyrządzić krzywdy człowiekowi. Zasady te są często przytaczane jako ostateczny cel robotów, które wchodzi w interakcje z ludźmi, ignorując fakt, że historie dotyczyły wbudowanych konfliktów i wystarczającej niejasności, aby prowadzić do zabawnych i niezamierzonych konsekwencji. Reguły były narzędziem literackim celowo zaprojektowanym tak, aby wspierać fabuły. Ponieważ zasady są mocno ugruntowane jako de facto złoty standard moralności funkcjonalnej robotów, ważne jest, aby zrozumieć zasady.

TRZY PRAWA ROBOTYKI

Trzy prawa wzięte dosłownie z tego to:

1. Robot nie może zranić człowieka ani przez beczynność dopuścić do zranienia człowieka.
2. Robot musi być posłuszny rozkazom ludzi, z wyjątkiem sytuacji, w których takie rozkazy byłyby sprzeczne z Pierwszym Prawem.
3. Robot musi chronić swoją egzystencję, o ile taka ochrona nie jest sprzeczna z Pierwszym lub Drugim Prawem.

Problemy z Trzema Prawami

Te trzy prawa, zbiorowo i indywidualnie, są problematyczne dla etycznego projektowania robotów. Murphy i Woods wymieniają i omawiają wady; istotę ich analizy podsumowano poniżej.

Łącznie te trzy prawa zakładają, że wszystkie roboty mają przynajmniej funkcjonalną moralność i prawdopodobnie są nawet w pełni moralnymi podmiotami. Jednak roboty mogą nie mieć wystarczającej woli i zdolności poznawczych, aby podejmować decyzje moralne, które mają wpływ etyczny. Przepisy lekceważą moralność operacyjną i nie uwzględniają odpowiedzialności wynalazców i producentów za decyzje projektowe.

Pierwsze prawo określa robota jako odpowiedzialnego agenta bezpieczeństwa, ale w rzeczywistości robot jest produktem. Jako produkt projektanci i producenci ponoszą odpowiedzialność za bezpieczeństwo i muszą być w stanie udowodnić, że postępowali rozsądnie i rozważnie.

Drugie prawo określa, że roboty muszą być posłuszne ludziom bez ograniczeń co do tego, którzy ludzie i w jakich okolicznościach. W rzeczywistości ludzie pracują w wyraźnych hierarchiach i nie każdy odpowiada za robota i jego zadanie. Wyobraź sobie terrorystę porywającego robota, aby wywołać chaos gospodarczy. Najwyraźniej polecenia niektórych osób miałyby wyższy priorytet niż polecenia innych osób.

Trzecie prawo mówi, że robot powinien się chronić. To prawo sugeruje, że robot ma wystarczające inteligentne możliwości, aby projektować i monitorować niebezpieczne warunki. W rzeczywistości oczekuje się, że ludzie przejmą kontrolę, gdy robot nie może się kontrolować, aby uratować robota przed awarią (lub uratować ludzi).

Trzy prawa odpowiedzialnej robotyki

TRZY PRAWA ODPOWIEDZIALNEJ ROBOTYKI

Murphy i Woods idą dalej i proponują przepisanie Trzech Praw Asimova, które nazwali trzema prawami odpowiedzialnej robotyki:

1. Człowiek nie może wdrożyć robota bez systemu, który spełnia prawne normy bezpieczeństwa i najwyższą etykę zawodową.
2. Robot musi odpowiadać ludziom odpowiednio do ich ról.
3. Robot musi być wyposażony w wystarczająco usytuowaną autonomię, aby chronić swoją egzystencję, o ile taka ochrona zapewnia płynne przekazywanie kontroli, które nie jest sprzeczne z pierwszym i drugim prawem.

Alternatywne pierwsze prawo wprowadziłoby moralność operacyjną jako podstawowy „przypadek” rozmieszczenia robota i przeniosłoby odpowiedzialność za moralne działania na projektanta i operatora. To alternatywne pierwsze prawo miałoby zastosowanie do uzbrojenia robotów i ograniczeń ich obecnych możliwości. Zwróć uwagę, że pierwsze prawo alternatywne wymaga od społeczności robotycznej wrażliwości na przepisy, a budowanie czegoś tylko dlatego, że możesz, nie uchyla odpowiedzialności etycznej. Alternatywne drugie prawo uznaje relacje niehierarchiczne oraz potrzebę bezpieczeństwa i hierarchii dowodzenia. Pozwala także na przypadki, gdy robot jest mądrzejszy od człowieka i niezależnie od tego, czy człowiek chce mieć kontrolę, robot będzie działał lepiej (np. łączyć przewody). Innym aspektem drugiego prawa jest to, że sposobem odpowiedzi nie musi być język naturalny. Robot szpitalny dostarczający leki, przechwycony przez gościa, może potrzebować powiedzieć „nie”, aby wskazać, że nie może przyjmować zamówień ani odpowiadać na pytania. Alternatywne trzecie prawo sprzeciwia się założeniu, że ludzie mogą rozwiązać każdy problem w dowolnym czasie, co nie jest prawdą, jak wykazały badania nad problemem kontroli człowieka poza pętlą, omówione w Części 18. To trzecie prawo faktycznie zachęca do rozwoju autonomii, przynajmniej na skalę taktyczną, aby umożliwić robotowi monitorowanie i prognozowanie przyszłego stanu własnego bezpieczeństwa.

Sztuczna inteligencja i etyka wdrażania

Robot z funkcjonalną lub pełną sprawczością moralną musi być pod pewnymi względami inteligentny. Potrzebne są reprezentacje wiedzy, aby umożliwić kodyfikację wyraźnych zasad etycznych, takich jak zasady zaangażowania wojska lub konwencja genewska, oraz ukrytych zasad etycznych. Planowanie i

rozwiązywanie problemów, wnioskowanie i poszukiwanie są potrzebne do rozważania stanu robota oraz do wnioskowania o tym, jak wpływa lub wpłynie na ludzi. Zrozumienie świata, a zwłaszcza tego, co robią ludzie, będzie prawdopodobnie wymagało wizji komputerowej. Uczenie się jest reklamowane jako sposób na przyswojenie przez roboty zasad i zachowań etycznych, chociaż nie jest jasne, czego robot miałby się uczyć i z jakich danych wejściowych. Istnieje ukryte oczekiwanie, że roboty z funkcjonalną i pełną sprawczością moralną będą w stanie wydawać rozkazy innemu robotowi i sprawić, by zrozumiał intencje pomimo niejasności w języku naturalnym.

Podsumowanie

Wszystkie roboty mają wpływ etyczny, choć może to być ukryty, ale roboty niekoniecznie są podmiotami moralnymi, które mają albo funkcjonalne, albo pełne moralne działanie. Projektanci ponoszą odpowiedzialność za bezpośrednie negatywne konsekwencje (moralność operacyjną), których można się spodziewać; Inteligencja robota musi zapewniać podejmowanie decyzji i wykonanie w czasie rzeczywistym (moralność funkcjonalna), ale projektanci nadal ponoszą odpowiedzialność. Trzy prawa Asimova przyciągnęły publiczną wyobraźnię jako wygodny punkt odniesienia dla etyki robotów, ale prawa te są głęboko wadliwe i nie można ich wdrożyć i stosować w praktyce. Wracając do pytań postawionych we wstępie, odpowiedzi mogą być znacznie inne niż oczekiwano. Czy uzbrajanie robotów w broń nie jest nieetyczne? Decyzja o uzbrojeniu robotów jest decyzją człowieka, a zatem należy do dziedziny polityki. Z pewnością nieetyczne jest dla projektanta robotyki budowanie broni w robotach bez odpowiednich środków bezpieczeństwa lub gwarantowanej niezawodności. Pytanie: Jeśli programiści zastosują Trzy Prawa Asimova, czy to nie wystarczy, aby uchronić nas przed obrażeniami, śmiercią lub powstaniem? powinien w tym momencie wywołać chichot u czytelnika, ponieważ prawa nie nadają się do stosowania jako realistyczne ramy programowania etycznych robotów. Prawa pokazują, jak trudno jest uchwycić złożoność ludzkiego społeczeństwa. Część nie odnosił się bezpośrednio do Czy nie jest nieetyczne tworzenie robotów i traktowanie ich jak niewolników? To pytanie nie dotyczy projektu inteligentnego robota, ale raczej ludzkiej etyki. Powinno być jednak jasne, że zdolności wymagane do tego, aby robot miał pełną sprawczość moralną, oznacza, że ma on inteligencję podobną do rówieśników. Nie jest jasne, kiedy robot przekracza próg od bycia narzędziem do bycia kognitywnym podmiotem z prawami do ochrony równymi co najmniej naczelnym, jeśli nie człowiekiem.