

Jarosław Kucharski
Uniwersytet Ignatianum w Krakowie
<https://orcid.org/0000-0001-6129-4477>
<https://doi.org/10.35765/slowniki.450>

Utylitaryzm

Streszczenie

DEFINICJA POJĘCIA: Utylitaryzm to zespół teorii etycznych, których cechą wspólną jest uzależnienie oceny działań moralnych od ich konsekwencji, mierzonych za pomocą wpływu na szczęście, przyjemność, dobrobyt bądź ilość zaspokojonych preferencji wszystkich zainteresowanych stron. W artykule wskazane są główne rodzaje utilitaryzmu.

ANALIZA HISTORYCZNA POJĘCIA: Pokazane są historyczne źródła utilitaryzmu. Analiza wychodzi od wskazania niektórych źródeł preutilitarystycznych. Następnie omówione są główne elementy teorii klasycznych i wybranych dwudziestowiecznych utilitarystów (J. Bentham, J.S. Mill, H. Sidgwick, E. Moore, R.B. Brandt, R.M. Hare, P. Singer) w zakresie różnych rodzajów utilitaryzmu.

UJĘCIE PROBLEMOWE POJĘCIA: Utylitaryzm staje także przed wyzwaniami i krytyką. Omówione są wybrane argumenty krytyczne względem utilitaryzmu, takie jak: problem „odrażającej konkluzji”, nieintuicyjne i, jak się wydaje, formułowane wbrew ustalonej moralności propozycje rozstrzygnięcia dylematów moralnych, niektóre różnice między moralnością chrześcijańską a utilitaryzmem oraz problem (zbyt) wysokich wymogów moralnych stawianych przez utilitaryzm.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Etyka utilitarystyczna pomimo swoich trudności znajduje zastosowanie między innymi w ocenie opłacalności wdrażanych środków i terapii medycznych lub polityk społecznych. Podniesione są problem odpowiedzialności za przyszłe pokolenia i argument, iż Zasada Bezstronności

powinna obowiązywać również w aspekcie czasowym. Wskazane zostanie także wyzwanie, jakim jest adaptacja obiektywnej teorii wartości w utilitaryzmie.

Słowa kluczowe: utilitaryzm, John Stuart Mill, Jeremy Bentham, Henry Sidgwick, Derek Parfit

Definicja pojęcia

Utylitaryzm to zespół teorii etycznych, rozwijanych systematycznie od początku XIX wieku. Ich cechą wspólną jest uzależnienie oceny moralnej danego działania (lub normy) przede wszystkim od konsekwencji, podjęcie próby wskazania niezależnego kryterium oceny norm i działań moralnych (takim kryterium mogą być np. szczęście, dobrobyt) oraz weryfikacja obowiązujących norm moralnych pod kątem przynoszonych przez nie konsekwencji. Podstawową zasadą utilitaryzmu jest Zasada Maksymalizacji. Głosi ona, iż dany czyn jest moralnie godziwy wtedy, gdy maksymalizuje wartość lub wartości przyjęte jako podstawowe w danym wariantcie utilitaryzmu. Równie ważną zasadą jest Zasada Bezstronności: każdy liczy się tak samo i każdy tylko za jednego. Pozostałe reguły są wtórne wobec tych zasad. Myśl utilitarystyczna wywarła duży wpływ na współczesną metaetykę, etykę praktyczną, bioetykę, filozofię polityczną i społeczną, antropologię filozoficzną, etykę zarządzania i biznesu oraz etykę maszyn i sztucznej inteligencji.

Utylitaryzm można podzielić ze względu na:

- przedmiot wyboru,
- czynniki będące podstawą określenia wartości danego działania,
- liczbę poziomów myślenia moralnego,
- źródło wartości.

Podział ze względu na przedmiot wyboru:

- Utylitaryzm czynów: aktu oceny należy dokonać przed każdym działaniem, tak aby zawsze wybrać takie, które będzie maksymalizowało pożądane konsekwencje. Przykład: utilitaryzm J. Benthama.
- Utylitaryzm reguł: należy dążyć do stworzenia takiego zestawu reguł, którego przestrzeganie będzie umożliwiało uzyskanie optymalnej użyteczności działań (optymalnej korzyści dla wszystkich zainteresowanych stron). Przykład: utilitaryzm R.B. Brandta.

Podział ze względu na czynniki będące podstawą określenia wartości danego działania:

- Utylitaryzm hedonistyczny: należy dążyć do jak największej przyjemności jak największej liczby osób. Jest to jedna z form monizmu aksjologicznego (poglądu głoszącego, iż tak naprawdę istnieje tylko jedna wartość, pozostałe są do niej sprowadzalne). Przykład: utilitaryzm J. Benthama.

- Utylitaryzm idealistyczny: przyjmuje, że istnieje wiele wartości (pluralizm aksjologiczny), które a) nie są do siebie nawzajem sprowadzalne oraz b) mogą lub powinny być maksymalizowane. Właściwe działanie dąży do maksymalizacji różnych rodzajów wartości (takich jak np. przyjemność, miłość, wiedza, piękno itp.). Właściwy czyn to taki, który realizuje największą osiągalną sumę tych wartości. Jedną z form utylitaryzmu idealistycznego jest utylitaryzm preferencji, głoszący, iż źródłem tych wartości są interesy oraz preferencje czujących istot. Przykład: propozycja J.E. Moore'a.

Podział ze względu na liczbę poziomów myślenia moralnego:

- Utylitaryzm jednopoziomowy: rozważań utylitarystycznych (np. rachunku utylitarystycznego) należy dokonywać w obrębie tego samego podmiotu i w stosunku do tych samych działań lub reguł. Przykład: utylitaryzm J. Benthama.
- Utylitaryzm dwupoziomowy: trzeba odróżnić poziom działania bezpośredniego, w którym nie ma potrzeby krytycznej refleksji nad przestrzeganiem reguł, od poziomu krytycznego, gdzie jest możliwość weryfikacji reguł, służących do kierowania działaniami. Przykład: system R.M. Hare'a.

Podział ze względu na źródło wartości:

- Utylitaryzm obiektywizujący: utylitaryzm taki przyjmuje, iż istnieją niezależne od zapatrywań członków danej społeczności wartości lub zbiór wartości, które należy maksymalizować. System reguł powinien dążyć do zwiększania tak określonych wartości. Przykładem takiego utylitaryzmu może być system Benthama.
- Utylitaryzm subiektywizujący: wartości, które należy maksymalizować, są zależne od interesów i planów zaangażowanych stron. To je należy brać pod uwagę, a system etyczny powinien dążyć do ich maksymalizacji. Przykład: utylitaryzm preferencji P. Singera.

Powyższy podział nie jest kompletny. Na jego podstawie można jednak pokazać, jak złożonym i wielogłosowym nurtem jest utylitaryzm.

Analiza historyczna pojęcia

Historycznych źródeł idei utylitarystycznych można szukać już w starożytności. Do weryfikacji obowiązujących norm moralnych nawoływali

sofiści. Arystoteles nie tylko sformułował koncepcję szczęścia jako ostatecznego celu ludzkiego życia, ale również próbował wyrazić ją w formie dającej przynajmniej możliwość empirycznej weryfikacji. Uwzględniania znaczenia przyjemności w życiu szczęśliwym domagali się epikurejczycy. Augustyn i Tomasz głosili potrzebę podporządkowania innych aktywności osiągnięciu szczęścia, będącego – zgodnie ze stanowiskiem chrześcijańskim – tożsamym ze zbawieniem. Średniowieczni myśliciele wprowadzili także do etyki pojęcie proporcjonalności, które wymaga odpowiedniego oceniania środków i celów. Akwinata zaproponował również uwzględnianie konsekwencji w ocenie moralnej działania. Skupienie się na konsekwencjach jako głównym wyznaczniku wartości moralnej postulował Machiavelli („skutek usprawiedliwia tego, kogo obciąża działanie”). Zauważyć należy, że naczelną zasadą utilitaryzmu: „najlepsze jest takie działanie, które zapewnia największe szczęście największej liczbie osób”, zanim zostało przytoczone przez F. Hutchesona, zostało sformułowane przez G.W. Leibniza (Hutcheson, 2004, s. X i s. 125). Hutcheson także zalecał stosowanie obliczeń w etyce, tak aby uwzględniać zarówno konsekwencje działań, jak i okoliczności, w których były dokonane. Listę koncepcji prowadzących do utilitaryzmu należy też uzupełnić o D. Hume’a, z jego postulatem odejścia od metafizycznego ugruntowania etyki. Cały empiryzm brytyjski poszukiwał już empirycznych odpowiedników takich pojęć jak „dobry” i „zły”, proponując jako ich odpowiedniki właśnie przyjemność i cierpienie.

Za pierwszego, klasycznego utilitarystę uznaje się Jeremy’ego Benthama. Od niego pochodzi sformułowanie zasady utilitarystycznej, która „aprobuje lub gani wszelką działalność zależnie od tego, czy wykazuje ona tendencję do powiększania czy zmniejszania szczęścia obchodzącej nas strony” (Bentham, 1958, s. 18). Bentham zauważa, że „Natura poddała rodzaj ludzki rządowi dwu zwierzchnich władców: przykrości i przyjemności. Im tylko dane jest wskazywać, co powinniśmy czynić, oraz stanowić o tym, co będziemy czynili” (Bentham, 1958, s. 17). Bentham wprost utożsamia wartość moralną działań z ich wpływem na przyjemność bądź jej brak. Wszelkie inne wartości (takie jak pomyślność, miłość, sprawiedliwość) są sprowadzalne do bilansu przyjemności i przykrości. W tym sensie Bentham jest monistą aksjologicznym. Przyjemność i przykreść postrzega on jako „wspólny mianownik” innych wartości. Wszystko, co jest cenione jako wartość, w ostatecznym

rozrachunku jest wartością dzięki temu, że sprawia przyjemność zainteresowanym stronom. Dzięki temu można skonstruować system, w którym rozstrzygalne będą dylematy etyczne (a nawet będą mogły zostać zlikwidowane) – przeliczenie danych wartości na ilość wytwarzanych przez nie przyjemności i przykrości pozwoli na określenie, która z danych opcji jest warta wyboru.

Bentham poza zaproponowaniem uniwersalnej podstawy wszelkich wartości przedstawia także kryteria mierzenia wartości poszczególnych działań. Przy ich ocenie należy uwzględnić: intensywność, trwałość, pewność lub niepewność, bliskość lub dalekość, płodność, czystość oraz rozciągłość. Intensywność odnosi się do poziomu odczucia danej przyjemności, trwałość – do jej długości i możliwości odczuwania przez dłuższy czas, pewność lub niepewność to prawdopodobieństwo wystąpienia danego skutku, bliskość lub dalekość wyraża horyzont czasowy, w jakim dany skutek wystąpi. Płodność oznacza prawdopodobieństwo wystąpienia kolejnych pozytywnych skutków, zaś czystość – brak związanej z daną przyjemnością rzeczywistej lub potencjalnej przykrości. Rozciągłość umożliwia zaś stwierdzenie, ile osób będzie ponosić konsekwencje (oraz przyjemności/przykrości) związane z danym działaniem (Bentham, 1958, s. 52–54).

Pierwszy system utylitarystyczny był nakierowany na wprowadzenie empirycznego rachunku do etyki, likwidację nierozstrzygalnych sporów moralnych, wskazanie metody wyznaczania wartości danych czynów, uniezależnienie tej wartości od rozstrzygnięć tradycyjnych, wskazanie potrzeby weryfikacji istniejących norm moralnych oraz uwzględnienie społecznego kontekstu i okoliczności w ocenie wartości moralnej czynu. Celem Benthama było doprowadzenie do stworzenia takiego systemu, który rzeczywiście będzie promował powszechną szczęśliwość. Namysł moralny miał polegać na sprawdzeniu, jakie konsekwencje przyniesie dane działanie według zaproponowanych kryteriów. Po dokonaniu tego rachunku utylitarystycznego i porównaniu jego wyników z alternatywnymi możliwościami działania można było podjąć decyzję co do słusznego postępowania.

Dzieło Benthama kontynuował i rozwijał John Stuart Mill. Zgadza się on z głównym poglądem poprzednika: „czyny są dobre, jeśli przyczyniają się do szczęścia, złe, jeżeli przyczyniają się do czegoś przeciwnego. Przez szczęście rozumie się przyjemność i brak cierpienia, przez

nieszczęście – cierpienie i brak przyjemności” (Mill, 2005, s. 10). Nie jest jasne, czy Bentham lub Mill rozumieli Zasadę Bezstronności w wymiarze przestrzeganym (czy chodziło im głównie o aktualnie istniejących ludzi), czy także w wymiarze czasowym (czy mieli na myśli wszystkich odczuwających ludzi, niezależnie od momentu urodzenia). Mill wprowadza modyfikacje w trzech aspektach. Po pierwsze proponuje rozróżnienie pomiędzy przyjemnościami ilościowymi i jakościowymi. Stwierdza: „lepiej być niezadowolonym człowiekiem niż zadowoloną świnia; lepiej być niezadowolonym Sokratesem niż zadowolonym głupcem” (Mill, 2005, s. 11). Mill uzupełnia w ten sposób brak zróżnicowania rodzajów przyjemności, *de facto* wprowadza pluralizm przyjemności. Okazuje się bowiem, że przyjemności jakościowe (właściwe Sokratesowi) są więcej warte niż przyjemności ilościowe (właściwe wszystkim, także głupcom). Pomiedzy nimi istnieje nieredukowalna różnica – nie można wymienić przyjemności płynącej z obcowania z arcydziełem na przyjemność płynącą z zaspokojenia głodu. Aby odróżnić, które przyjemności są jakiego rodzaju, należy odwołać się do sądu osób, które znają je obie i porównać je parami. Zatem do odróżniania rodzajów przyjemności niezbędnymi są sędziowie kompetentni.

Po drugie Mill przypisał Benthamowi sformułowanie Zasady Bezstronności: „każdy liczyć się ma za jednego i nikt za więcej niż jednego” (Mill, 2005, s. 84). Utylitysta podkreśla tym samym silny egalitaryzm. Każdy ma takie samo prawo do szczęścia i szczęście (przyjemności) każdej jednostki liczy się tak samo. W konsekwencjach danego działania liczą się przyjemności, które ono powoduje dla dowolnej osoby odczuwającej przyjemność. Podkreślenie tego warunku wskazuje na daleko idący uniwersalizm zasad utylitarystycznych. Uwypukla, że w rachunku utylitarystycznym liczą się przyjemności, nie osoby. Oznacza to także, że nie należy utożsamiać utylitaryzmu z egoizmem.

Po trzecie, Mill jest uważany za prekursora utylitaryzmu reguł. Zauważa, że istnieją co najmniej dwa poziomy reguł: ogólna zasada użyteczności i reguły drugiego stopnia (reguły szczegółowe). Można dokonywać namysłu nad każdym konkretnym działaniem bądź też nad regułami, ponieważ nie zawsze będzie czas na dokładne rozważenie konsekwencji danego działania.

Ludzkość miała do chwili obecnej czas wypracować sobie pewne określone poglądy na skutki, jakie mają pewne czyny dla jej szczęścia. Te przekazane nam poglądy stanowią dla ogółu reguły moralne. Są one nimi także i dla filozofa, póki nie uda mu się znaleźć lepszych (Mill, 2005, s. 33).

Istnieją więc pewne reguły, które konkretyzują ogólną zasadę utilitaryzmu. Część z nich została odkryta w toku rozwoju ludzkości. Mill zauważa, że reguły te były nakierowane na osiągnięcie pomyślności, lecz podlegają one udoskonalaniu. Można uznać, że reguły są czymś w rodzaju „moralności tymczasowej” postulowanej przez Kartezjusza – należy ich przestrzegać, a w miarę możliwości dokonywać namysłu nad ich ulepszeniem.

Henry Sidgwick przeciwstawiał utilitaryzm etyce zdroworozsądkowej (jak sam ją nazywał – intuicyjnej). Wskazywał, że etyka zdroworozsądkowa nie potrafi wskazać swojej głównej zasady. Jest ona *de facto* subiektywną postawą, polegającą na indywidualnych intuicjach moralnych. Jej źródła to partykularny zmysł moralny i normy odziedziczone przez kulturę. Inaczej sprawa ma się z utilitaryzmem. Jest on oparty na solidnej podstawie: obiektywnej zasadzie utilitarystycznej, nakazie uniwersalizacji oraz możliwości racjonalnego namysłu i weryfikacji norm moralnych. Sidgwick uważa, że to utilitaryzm umożliwia przekroczenie egoizmu (indywidualnego lub grupowego) – właśnie dzięki rzetelnemu namysłowi (przyjęcie punktu widzenia Wszechświata (Sidgwick, 1962, 382)). Jednakże angielski filozof ma świadomość, że dążenie do całkowicie bezosobowego wyboru nie jest praktycznie możliwe do osiągnięcia dla człowieka. Ten bowiem zawsze będzie zdradzał tendencję do preferowania tych osób, na których mu zależy. Ponadto Sidgwick sformułował paradoks hedonizmu:

Szczęście może być osiągnięte w większym stopniu, jeżeli zakres, w jakim świadomie do niego dążymy, zostanie starannie ograniczony. I to nie tylko dlatego, że działanie może być bardziej skuteczne, jeśli nasz wysiłek jest tymczasowo skoncentrowany na realizacji bardziej ograniczonych celów – choć jest to bez wątpienia ważny powód – ale także dlatego, że najpełniejszy rozwój szczęśliwego życia dla każdej jednostki wydaje się wymagać, aby miała ona inne zewnętrzne obiekty zainteresowania poza szczęściem innych świadomych istot (Sidgwick, 1962, 405).

Rozróżnia tutaj pomiędzy szczęściem jako celem a szczęściem jako motywem. Sidgwick zauważa, że jeżeli głównym motywem działających podmiotów będzie szczęście (rozumiane utilitarystycznie lub

hedonistycznie), to ich działanie nie będą tak efektywne, jak w przypadku, gdy będą mieli inne motyw. Po sformułowaniu tego paradoksu Sidgwick zwraca uwagę, że może on dotyczyć samego utylitaryzmu:

Na gruncie zasad utylitarystycznych słuszne może być robić i prywatnie zalecać, pod pewnymi warunkami, to, za czym niesłusznie byłoby opowiadać się otwarcie; słusznie może być jedną grupę ludzi uczyć otwarcie tego, czego źle byłoby uczyć innych; i do pomyślenia jest, że słusznie może być robić coś, o ile da się to zrobić względnie dyskretnie, co źle byłoby robić na oczach świata (...) doktryna głosząca, iż moralność ezoteryczna jest pożyteczna sama powinna mieć status doktryny ezoterycznej (Sidgwick, 1962, s. 489–490, polskie tłumaczenie za: Williams, 1999, s. 346–347).

Sidgwick wprost zaleca uznanie utylitaryzmu za ezoteryczną teorię moralną i wskazuje, że powinna ona raczej być stosowana przez elity niż przez ogół. Takie odejście od jawności ma prowadzić do bardziej efektywnej realizacji zasady utylitaryzmu. Sidgwick w swoim paradoksie przyznaje, iż sam utylitaryzm jest na tyle racjonalny oraz wymaga tak wysokich kompetencji filozoficznych i samokontroli, iż lepiej (z punktu widzenia Wszechświata) będzie, jeżeli nie stanie się on doktryną wyznawaną przez większość ludzi.

George E. Moore postanowił rozwiązać przynajmniej część trudności zidentyfikowanych przez Sidgwicka. Po pierwsze przyjął, iż „dobry” jest własnością prostą, niedefiniowalną, a co za tym idzie – nie można sprowadzić go do przyjemności lub innych rodzajów dóbr. Termin „dobry” może być orzekany o różnych podmiotach, lecz jest do nich nieredukowalny. System Moore’a jest pierwszym utylitaryzmem idealistycznym – przyjmującym pluralistyczną i nieredukowalną koncepcję dobra. Jako taki może rozwiązać paradoks hedonizmu Sidgwicka – nie ma potrzeby ukrywania ani ograniczania dążenia do dobra, ponieważ wiele rzeczy może być określane jako dobre. Główna zasada utylitaryzmu pozostaje bez zmian – nadal należy dążyć do maksymalizacji dobra dla jak największej liczby ludzi. Nie trzeba jednak tej zasady czynić ekskluzywną (Moore, 1980).

Utylitaryzm idealny postanowił ulepszyć Richard B. Brandt, przekształcając go w kompletny utylitaryzm reguł. Podobnie jak Sidgwick, Brandt poszukuje metody weryfikacji intuicyjnych reguł moralnych. W tym celu proponuje metodę „kontrolowanej postawy”, czyli takiej, w której przyjmuje się postawę bezstronną (lub spełniającą warunek

odwracalności, czyli zastosowanie tej samej zasady do głoszącego ją podmiotu), doinformowaną (nie zmienia się pod wpływem świadomości nowych, relewantnych faktów), będącej w normalnym stanie umysłowym (polegającym na niepodleganiu obciążeniom psychicznym) i nieprzyjmującej zasad, które nadmiernie komplikowałyby system, co prowadziłoby do zbytniego uszczegółowienia, a w niektórych przypadkach być może do sprzeczności między przyjętymi regułami (Brandt, 1996, s. 425–426). Metoda ta umożliwiłaby sformułowanie kilku reguł, które, o ile byłyby całkowicie przestrzegane, doprowadziłyby do maksymalizacji pluralistycznie rozumianej użyteczności. Jak zauważa Brandt:

utilitaryzm reguł głosi, że w poszczególnych przypadkach powinniśmy kierować się ogólnymi przepisami, to zaś, czy dany przepis ogólny powinien zostać przyjęty zależy od użyteczności powszechnego respektowania tego przepisu (Brandt, 1996, s. 673–674).

System ten ma mieć przewagę nad utilitaryzmem czynów, ponieważ, po pierwsze, nie ma już potrzeby każdorazowego namysłu w sytuacji wyboru (a co za tym idzie – mozolnego liczenia), jakie działania będą w danym przypadku maksymalizować użyteczność. Po drugie, utilitaryzm nie musi być już ezoteryczny – podążanie za regułami i same reguły mogą być jawne. System ma być prosty, a sposób ustalania reguł znany, przy tym reguł nie powinno być zbyt wiele (metoda kontrolowanej postawy). Jeżeli okazuje się, że któraś z wybranych reguł nie działa, to w ramach przyjętej metody poddaje się ją weryfikacji. Według Brandta ma również przewagę nad propozycjami etyki intuicjonistycznej (np. etyka obowiązków *prima facie*) dzięki temu, że proponuje metodę i zestaw reguł rozstrzygających spory moralne.

Jednakże utilitaryzm reguł jest odejściem od założeń klasycznego utilitaryzmu. Nie chodzi już bowiem o to, aby każde działanie moralnie godziwe maksymalizowało użyteczność, lecz o wypracowanie zestawu reguł, które w wyjątkowym przypadku (wszyscy przestrzegają reguł) maksymalizują użyteczność. W sytuacji realnej (nie wszyscy przestrzegają reguł) można wskazać dwa warianty. Albo utilitaryści reguł godzą się na osiąganie wystarczającego stopnia użyteczności (rezygnując z maksymalizacji), albo też powinni budować taki zestaw reguł, który będzie maksymalizował użyteczność w sytuacji, gdy nie wszyscy przestrzegają reguł.

Problem metody rozstrzygania norm moralnych zaprzętał również głowę Richarda M. Hare'a. W ramach utylitaryzmu, aby uniknąć problemów, których świadomy był Sidgwick, oraz aby doprecyzować kontrolowaną postawę Brandta, zaproponował on dwupłaszczyznowe myślenie moralne. Zauważył, że na co dzień podmioty moralne kierują się zasadami domniemanymi. Są one formułowane na poziomie intuicyjnym i na ogół wystarczają do właściwej oceny niewymagających sytuacji moralnie znaczących. Problem rodzi się wtedy, gdy zasady domniemane wchodzą ze sobą w konflikt. Wówczas należy przejść na poziom tak zwanego „myślenia krytycznego”.

Celem myślenia krytycznego jest wyselekcjonowanie najlepszego do zastosowania w myśleniu intuicyjnym zespołu zasad domniemanych. (...) Funkcją myślenia krytycznego, oprócz selekcjonowania zasad domniemanych, jest również rozwiązywanie konfliktów między tymi zasadami (Hare, 2001, s. 68).

Zasada Maksymalizacji użyteczności ma być naczelną zasadą rozstrzygania sporów. Hare jest świadomy, że do racjonalnego rozstrzygnięcia sporów wymagany jest trening – hipotetyczną istotę, która zawsze myśli na poziomie krytycznym, nazywa archaniołem. Jednocześnie zauważa, że myślenie krytyczne nie ma na celu stworzenia jednego, uniwersalnego zespołu reguł. Ma ono interweniować wtedy, gdy reguły domniemane przestają działać optymalnie. Zbiór reguł domniemanych jest zmienny. Jak zauważa: „Dobrym wodzem jest ten, kto wygrywa bitwy, a nie ten, kto ma najlepsze zasady domniemane” (Hare, 2001, s. 70).

Jako ostatni zostanie przedstawiony pogląd utylitaryzmu preferencji. Skupia się on na źródle informacji na temat dóbr, które powinny być maksymalizowane. Chodzi o tak zwany utylitaryzm preferencji lub utylitaryzm skupiony na interesach. Jak zauważa Peter Singer, w jego przypadku „najlepsze konsekwencje» są rozumiane raczej jako to, co ogólnie sprzyja realizacji interesów tych, których dotyczy, iż jedynie zwiększa przyjemność i zmniejsza ból” (Singer, 2003, s. 28). Singer wskazuje, iż wartości, które powinny być maksymalizowane, właśnie dlatego są wartościami, iż są czyimiś interesami bądź szerzej – że przez jakąś istotę myślącą są preferowane w stosunku do innych. Rolą utylitaryzmu jest przedstawienie metody oceniania preferencji wielu żywiących je istot w taki sposób, aby możliwa była maksymalizacja dobrobytu rozumianego jako wypełnienie preferencji. Utylitaryzm preferencji stara się

uniknąć problemów związanych z narzucaniem innym sposobu, w jaki powinni żyć (i czego powinni pragnąć). Co więcej, pomimo że uznaje wiele wartości, to możliwa jest jego interpretacja jako monizmu aksjologicznego – wszelkie wartości mogą być sprowadzone do zaspokojenia lub niezaspokojenia preferencji (Janikowski, 2016, s. 80–81) Co istotne – mowa tu przede wszystkim o racjonalnych preferencjach (a zatem przemyślanych, podejmowanych lub wzmacnianych przez poinformowane podmioty, świadome swoich uwarunkowań). W przypadku utilitaryzmu preferencji można podać parafrazę zdania Benthama/Milla: każda preferencja liczy się tak samo i każda tylko za jedną.

Ujęcie problemowe pojęcia

Utilitaryzm wzbudzał (i wzbudza) wiele kontrowersji. Etyka oparta na racjonalnym „wylicznaniu”, wartościowaniu konsekwencji, której celem jest selekcja i zmiana tradycyjnych reguł moralnych, została poddana krytyce ze strony etyk deontologicznych oraz opartych na pojęciu cnoty. Krytycznie odnosi się do niej również etyka chrześcijańska. Poniżej zostanie przedstawione jedynie kilka z tych argumentów krytycznych.

Jak zauważa Bernard Williams, utilitaryzm dostarcza argumentów, iż pewne działania są moralnie słusznie, pomimo że zawierają w sobie element działania nagannego. Przedstawia przykład, w którym Jim otrzymuje propozycję od szeryfa miasteczka na południu USA. Schwytano tam dwudziestu Indian i skazano ich na śmierć za antyrządowe protesty. Jim, jako gość w miasteczku, otrzymuje propozycję: jeżeli uśmierci osobiście jednego z Indian, pozostali odejdą wolni. Z punktu widzenia utilitaryzmu czynów, a zapewne również z punktu widzenia myślenia krytycznego, wykonanie wyroku przez Jima jest działaniem lepszym niż zezwolenie na śmierć dwudziestu osób (Smart & Williams, 1973, s. 98–99). Williams rozwija ten zarzut, stwierdzając, iż utilitaryzm nie uwzględnia szkód, jakie płyną z posiadania prawego charakteru, wierności przyjętym zasadom (*integrity*), zmniejsza znaczenie wrażliwego sumienia oraz subiektywnego poczucia, iż zrobiło się coś złego (skrupułów). Williams zauważa, że wszystkie te zjawiska posiadają jednak pozytywną wartość moralną, a niszczenie ich za pomocą rachunku utilitarystycznego będzie raczej prowadzić do zmniejszenia oporów przed

podejmowaniem w przyszłości działań tradycyjnie ocenianych jako niemoralne. To z kolei może prowadzić raczej do zmniejszenia powszechnej użyteczności niż do jej maksymalizacji.

Odpowiadając na ten argument, John J.C. Smart najpierw zauważa, że aby podważyć utilitaryzm, trzeba skonstruować tak daleko odbiegający od rzeczywistości przykład. Demaskuje on metodę antyutilitarystów: konstruują oni przykłady, w których działanie w innych sytuacjach naganne staje się właściwe. Zauważa przy tym, że argumenty takie opierają się na połączeniu oceny działania i oceny podmiotu. Proponuje, aby uznać, że moralnie dobra osoba to taka, która zazwyczaj realizuje najlepszy możliwy rezultat, niż przeciętnie się to dzieje. Analogicznie osoba zła moralnie to taka, która czyni to w sposób gorszy niż przeciętna. Dobry motyw to taki, który zasadniczo prowadzi do korzystnych skutków, zły – do niekorzystnych (Smart & Williams, 1973, s. 48). Osoba taka jak Jim może być moralnie dobrą osobą i działając z dobrych pobudek, uczynić to, co innym może wydawać się naganne. Smart za pomocą tego rozróżnienia chce pokazać, iż utilitaryzm także będzie wspierał moralnie pozytywne cechy takie jak prawy charakter czy skrupuły. Jednakże samo przestrzeganie zasad zawartych w sumieniu nie gwarantuje właściwego działania. Tutaj pojawia się utilitarystyczny pomysł dotyczący czujności – zwolennik tej etyki będzie wymagał (również od siebie) sprawdzania, czy posiadane zasady domniemane są zgodne z Zasadą Maksymalizacji. Utylitaryzm jawi się więc jako stawiający o wiele większe wymagania moralne niż etyka deontologiczna czy etyka cnoty. Nie godzi się na pozostanie na poziomie przestrzegania reguł czy też na poziomie kształtowania charakteru, lecz wskazuje, iż zawsze trzeba być gotowym na zmianę swoich zasad w imię realizacji naczelnej zasady moralnej.

O wiele poważniejszy zarzut w stronę utilitaryzmu wysunął Derek Parfit. Nazwał go „odrażającą konkluzją”. Przedstawił ją następująco:

dla dowolnej populacji liczącej przynajmniej dziesięć miliardów ludzi cieszących się bardzo wysoką jakością życia musi istnieć jakaś znacznie większa możliwa do wyobrażenia populacja, której egzystencja – o ile wszystko inne pozostaje równe – byłaby lepsza, mimo że życie jej członków ledwo przekracza minimalny poziom, który czyni możliwym uznanie go za życie warte tego, by nim żyć (Parfit, 2012, s. 444).

„Odrażająca konkluzja” polega na tym, że dla każdego stanu o ustalonej wartości użyteczności można pomyśleć taki stan, w którym łączna suma użyteczności jest większa, pomimo że użyteczność tej sytuacji dla każdego z podmiotów czujących jest mniejsza. Z moralnego punktu widzenia należałoby więc preferować drugi stan (większa liczba minimalnie zadowolonych podmiotów) niż pierwszy (mniejsza liczba o wiele bardziej zadowolonych podmiotów). Parfit, przedstawiając „odrażającą konkluzję”, zwraca uwagę na pewną cechę utilitaryzmu, w szczególności utilitaryzmu preferencji: jeżeli nie istnieje obiektywna hierarchia wartości, wówczas spełnienie dowolnej wartości zwiększa globalną sumę szczęścia. Zasada Maksymalizacji nakazuje zwiększanie tej sumy. Zatem: lepiej jest, aby istniało wiele czujących podmiotów znajdujących się tuż nad progiem wartości, które „czynią życie wartym przeżycia”, niż aby istniało mniej podmiotów, które posiadają ze wszech miar szczęśliwe życie.

„Odrażająca konkluzja” zwraca uwagę na konsekwencje oderwania wartości użyteczności jako takich od wartości podmiotów, które ich doświadczają. Wskazuje, iż liczą się same pożądane odczucia (przyjemności, szczęście, wartości), a nie podmioty. Nieistotna jest również agregacja tychże własności w samych podmiotach; „każde pożądane odczucie liczy się tak samo i każde tylko za jednego”. W takiej formie „odrażająca konkluzja” stanowi wyzwanie dla zastosowań utilitaryzmu w polityce społecznej. Na jej podstawie można bowiem uznać za moralnie właściwy nawet najbardziej niesprawiedliwy rozkład dóbr i użyteczności, byleby spełniał warunek efektywności w sensie Pareto. Utilitaryzm poszukuje sposobu uniknięcia „odrażającej konkluzji”. Jednym ze sposobów jest próba stworzenia (na bazie Milowskich przyjemności jakościowych oraz hierarchii preferencji) obiektywnej hierarchii wartości użyteczności, dzięki którym ulegnie zmianie sam rachunek utilitarny (Janikowski, 2016, s. 75–80).

Utilitaryzm został także poddany krytyce zewnętrznej za uzależnienie wartości albo od bycia przedmiotem preferencji istoty czującej, albo też od przynoszenia takiej istocie przyjemności lub innych odczuć określanych jako dobre. Oznacza to, że istota, która nie posiada w stosunku do siebie takich preferencji (podobnie jak inne istoty w stosunku do niej) oraz która nie jest w stanie odczuwać bólu lub innej nieprzyjemności (lub przyjemności), nie posiada żadnej wartości. Na takich

lub podobnych przesłankach zbudowane są utylitarystyczne argumenty na rzecz moralnej dopuszczalności aborcji, eutanazji, a nawet niekiedy dzieciobójstwa. Takie konkluzje są odrzucane między innymi przez przedstawicieli etyki chrześcijańskiej. Wskazują oni, że odebranie ludziom niezależnej od wszystkiego wartości wewnętrznej może prowadzić do katastrofalnej równi pochyłej.

Należy wspomnieć o jeszcze jednej właściwości systemów utylitarystycznych, jaką są wysokie wymagania moralne. Jeżeli przystać na stwierdzenie, iż najważniejsza jest Zasada Maksymalizacji, zaś pozostałe powinny być jej podporządkowane, to za niemoralne należałoby uznać realizację w danej sytuacji którejkolwiek z niższych sum użyteczności, jeżeli można zrealizować większą. W myśl tego argumentu utylitarysta powinien zawsze dążyć do zidentyfikowania i zaspokojenia maksymalnej możliwej sumy użyteczności. Takie poziomy wymogów moralnych jest porównywalny z wymaganiem realizacji cnót heroicznych obecnych u Arystotelesa czy też w etyce chrześcijańskiej. Utylitaryzm idzie jednak o krok dalej. Realizacja maksymalnie dobrego czynu staje się obowiązkiem utylitarysty. Realizacja jakiegokolwiek stanu rzeczy o mniejszej wartości globalnej niż najlepszy nie jest realizacją Zasady Maksymalizacji. Próby zmniejszenia tak wysokich wymagań moralnych były podejmowane w ramach utylitaryzmu reguł oraz utylitaryzmu dwóch poziomów myślenia. W ich przypadku stan, w którym wszyscy realizują Zasadę Maksymalizacji, zostaje uznany za odległy ideał, do którego prowadzi stopniowe włączanie coraz to większej liczby reguł utylitarystycznych do moralności. Jednakże takie ujęcie zakłada jakąś formę moralności ezoterycznej – utylitaryści są świadomi, jaka moralność powinna obowiązywać, jednak prezentują takie zasady moralne, które mają największe szanse na bycie zaakceptowanymi przez społeczeństwo.

Refleksja systematyczna z wnioskami i rekomendacjami

Utylitaryzm postawił etyce i praktyce społecznej nowe wyzwania. Stał się źródłem zarówno inspiracji, jak i punktem odniesienia do krytyki. Poprzez niekiedy bezkompromisowe podejście do tradycyjnych rozstrzygnięć problemów moralnych spowodował, iż także krytycy

podejścia konsekwencjalistycznego na nowo musieli sformułować argumenty na rzecz obowiązywalności niektórych norm moralnych. Ponadto dzięki utylityzmowi rozbudzona została dyskusja na temat moralności ezoterycznej, wartości życia oraz sprawiedliwości społecznej. Argumentów utylitystycznych w debacie etycznej i społecznej nie można zlekceważyć. Przykładem zastosowania kategorii utylitystycznych do praktyk społecznych może być działalność komisji do spraw refundacji leków i terapii. Opracowanie wskaźników takich jak QALY (wskaźnik jakości życia skorygowany o liczbę lat – *quality adjusted life years*), ICER (stosunek dodatkowych kosztów do dodatkowych korzyści – *the incremental costeffectiveness ratio*) czy QUALE (liczba oczekiwanych lat życia z uwzględnieniem ich jakości – *quality adjusted life expectancy*); problem, komu jako pierwszemu udzielać świadczeń zdrowotnych – zarówno ratujących życie, jak i podnoszących jego długość lub jakość (i jakie kryteria przy tym zastosować) – są problemami, w których również stosuje się kryteria utylitystyczne i które trudno jest rozstrzygnąć bez wzięcia pod uwagę bilansu dóbr. Utylityzm, w szczególności utylityzm czynów, jest też rozważany jako jedna z kandydatur na system etyczny dla maszyn opartych na sztucznej inteligencji. W przeciwieństwie do ludzi maszyny będą w stanie bardzo precyzyjnie operować na dużych zbiorach danych, przez co mogą wykorzystać algorytmy oparte np. na wytycznych Benthama. Problemem oczywiście pozostaje nieintuicyjność rozwiązań, do których mogą w ten sposób dojść.

Przed samym utylityzmem również stoją wyzwania. Po pierwsze, wciąż trwa dyskusja na temat „odrażającej konkluzji”. Podejmowane są coraz to nowe próby jej odrzucenia bądź osłabienia. Po drugie, trwa dyskusja na temat odpowiedzialności osób żyjących aktualnie względem przyszłych pokoleń. Z jednej strony pojawia się argument, iż powinno się włączyć do aktualnie przeprowadzanego rachunku utylitarne przyszłe możliwe stany istot, których jeszcze nie ma, a które mogą mieć preferencje np. korzystania z przyrody w sposób podobny do nas. Argument ten jest uzasadniany wymogiem poszerzenia bezstronności o wymiar czasowy. Ponieważ „każdy liczy się tak samo i każdy tylko za jednego”, nie należy uznawać, że moment urodzenia ma w tym przypadku jakieś znaczenie. Zatem należy dbać o pomyślność przyszłych pokoleń, bez faworyzowania tych aktualnie istniejących bądź tych, które zaistnieją w niewielkim horyzoncie czasowym (Mulgan, 2022, s. 5, 21).

Z drugiej strony stawiany jest argument, iż przyszłe pokolenia nie istnieją, nie istnieją więc także ich preferencje (ani dobro). Tym, co powinno być brane pod uwagę, są przede wszystkim interesy, dobra i preferencje osób aktualnie istniejących, także te dotyczące przyszłych pokoleń. Takie jednak postawienie sprawy powoduje, iż może pojawić się zarzut paternalizmu – aktualnie istniejące osoby mają decydować o dobru osób, które są zaledwie możliwe, bez uwzględniania ich faktycznych preferencji. Spór pozostaje otwarty i niewątpliwie wymaga rzetelnego namysłu nad odpowiedzialnością aktualnych utylitarystów za dobrobyt przyszłych pokoleń.

Utylitaryzm jednak może także inspirować do: a) poszukiwania płaszczyzny porozumienia pomiędzy różnymi systemami wartości, b) poszukiwania rozwiązań dylematów etycznych, c) krytycznego podejścia do obowiązujących zasad moralnych. W przypadku a) taki był cel utylitaryzmu. Pytanie: „Które wartości należy maksymalizować?” jest nadal ważne. Nawet jeżeli nie obowiązuje teza o wzajemnej redukowalności wartości, to nadal warto poszukiwać odpowiedzi na pytanie: „Co czyni wartości wartościami?” oraz „Jaka jest hierarchia wartości?”. Utylitaryzm, pomimo że wskazywał tendencje naturalistyczne (np. utożsamienie wartości z przyjemnością), domaga się racjonalnego uzasadnienia uznawania i hierarchizacji danych wartości. Odnośnie do b) argumentacja utylitarystyczna jest w stanie zidentyfikować pozorne konflikty moralne. Nie oznacza to, że potrafi rozstrzygnąć wszystkie, lecz wskazuje jako na pozorne te, które zachodzą np. pomiędzy zasadami domniemanymi bez ich krytycznej analizy. W przypadku c) utylitaryzm nie zadowala się odpowiedzią: dane wartości są wartościowe z natury, tradycji bądź „głębokiego poznania”. Jest ciągłym wyzwaniem dla zastanej moralności, domaga się racjonalnych uzasadnień norm moralnych. Dzięki temu jest czynnikiem wpływającym pozytywnie na postęp moralny (bo każdy liczy się tak samo i każdy tylko za jednego). Pozostanie więc ważnym punktem odniesienia dla współczesnej etyki.

BIBLIOGRAFIA

- Bentham, J. (1958). *Wprowadzenie do zasad moralności i prawodawstwa*. Przeł. B. Nawroczyński. Warszawa: PWN.
- Brandt, R.B. (1996). *Etyka. Zagadnienia etyki normatywnej i metaetyki*. Przeł. B. Stanosz. Warszawa: PWN.
- De Lazari-Radek, K., & Singer, P. (2014). *The Point of View of the Universe. Sidgwick and Contemporary Ethics*. Oxford: Oxford University Press.
- Hare, R.M. (2001). *Myślenie moralne. Jego płaszczyzny, metody i istota*. Przeł. J. Margański, Warszawa: Fundacja Aletheia.
- Hutcheson, F. (2004). *An Inquiry into the Original of Our Ideas of Beauty and Virtue in Two Treatises*. Indianapolis: Liberty Fund.
- Janikowski, W. (2016). *Wolna wola i użyteczność. Podstawy etyki utilitaryzmu preferencji*. Warszawa: Semper.
- Klimowicz, E. (1974). *Utylitaryzm w etyce*. Warszawa: PWN.
- Mill, J.S. (2005). *Utylitaryzm*. Przeł. M. Ossowska. W: J.S. Mill, *Utylitaryzm. O wolności* (s. 3–89). Warszawa: PWN.
- Moore, G.E. (1980). *Etyka*. Przeł. Z. Szawarski. Warszawa: PWN.
- Mulgan, T. (2020). *Utilitarianism*. Cambridge: Cambridge University Press.
- Parfit, D. (2012). *Racje i osoby*. Przeł. W.M. Hensel, & M. Warchała. Warszawa: PWN.
- Saja, K. (2008). *Język etyki a utilitaryzm. Filozofia moralna Richarda M. Hare'a*. Kraków: Aureus.
- Sidgwick, H. (1962). *The methods of ethics*. London: Palgrave Macmillan.
- Singer, P. (2003). *Etyka praktyczna*. Przeł. A. Sagan. Warszawa: Książka i Wiedza.
- Smart, J.J.C., & Williams, B. (1973). *Utilitarianism: for and against*. Cambridge University Press.
- Williams, B., (1999). *Punkt widzenia Wszechświata. Sidgwick i ambicje etyki*. Przeł. M. Szczubiałka. W: B. Williams, *Ile wolności powinna mieć wola* (s. 329–353). Warszawa: Fundacja Aletheia.