

R. EUGENE BALES

## Czego należy oczekiwać od utilitaryzmu czynów?\*

*Dwa rodzaje utilitaryzmu*

Zwykło się tradycyjnie dzielić teorie etyczne na teleologiczne i deontologiczne. Według teorii teleologicznych słusność, niesłusność czy też powinność moralna (*moral rightness, wrongness, and obligatoriness*) zależą od wartości konsekwencji. Niektórzy spośród teleologów utrzymują np., że słusność, niesłusność lub powinność czynu zależą od tego, w jakim stopniu przyczynia się on do stworzenia wewnątrznie dobrych lub złych układów spraw i sytuacji (*contribution towards intrinsically good or bad states of affairs*).

Teorie deontologiczne przeczą, jakoby słusność, niesłusność i powinność czynu były jedynie zależne od wartości jego konsekwencji. Innymi słowy, deontologowie twierdzą, że istnieją moralnie znaczące racje inne niż te, na które powołują się zwolennicy teorii teleologicznych. Zazwyczaj deontologowie próbują okazać, że czyny są słuszne, niesłuszne lub powinny być spełnione, ponieważ są czynami pewnego rodzaju bądź też są nakazane lub zakazane przez pewne zasady moralne. Etycy z reguły zatem należą do jednego z dwóch obozów: obozu, który nalega na przestrzeganie zasad, bądź obozu, który zaleca maksymalizację wartości wewnętrznej.

Respektowanie zasad i maksymalizacja wartości wewnętrznej, traktowane jako kryteria powinności, nie muszą bynajmniej pozostawać w sprzeczności. Jednakże zarówno deontologowie, jak teleologowie zgodnie twierdzą, że bywają czyny, które nie zwiększają wartości wewnętrznej i które zarazem są rzeczywiście zgodne z uznawanymi przez większość

---

\* Wcześniejsze wersje tego artykułu pt. *Act-Utilitarianism: Ethical Theory or Decision-Making Procedure?* były przedstawione w Philosophy Club, California State College w Hayward oraz na posiedzeniu American Philosophical Association, Western Division, Portland (Oregon) w marcu 1969 r.

zasadami moralnymi. Deontologowie są skłonni sądzić, że czyny te są słuszne, ponieważ są zgodne z zasadami. Teleologowie są skłonni przypuszczać, że czyny te są złe, mimo że istotnie zgodne są z zasadami.

Pośród teorii teleologicznych pierwsze miejsce zajmują najróżniejsze rodzaje utylitaryzmu. Niniejszy artykuł stanowi próbę oceny wartości argumentów wysuwanych czasami na poparcie tezy, iż utylitaryzm czynów nie może być adekwatną teorią etyczną. Jak wskazuje na to nazwa, utylitaryzm czynów jest jednym z rodzajów utylitaryzmu. Od niedawna przyjęło się odróżniać utylitaryzm czynów (*act-utilitarianism*) od utylitaryzmu zasad (*rule-utilitarianism*). Rozróżnienie to ma istotne znaczenie dla pojęcia utylitaryzmu czynów.

Można ogólnie powiedzieć, że — według utylitaryzmu czynów — określony czyn (w przeciwieństwie do typu lub klasy czynów) jest słuszny wtedy i tylko wtedy, jeżeli jego użyteczność, tzn. wkład do wewnętrznie dobrych układów spraw i sytuacji, jest nie mniejsza niż jakiegokolwiek innego czynu alternatywnego; jest zaś niesłuszny wtedy, jeżeli jego użyteczność jest mniejsza niż jakiegokolwiek innego czynu alternatywnego. Według utylitaryzmu zasad, czyn (ogólnie rzecz biorąc) jest słuszny (lub niesłuszny) wtedy i tylko wtedy, jeżeli jest zgodny (lub niezgodny) z zasadą zależną w pewien sposób od użyteczności.

Utylitaryzm zasad jest jedną z najnowszych prób pogodzenia etyki teleologicznej z deontologiczną. Zwolennicy utylitaryzmu zasad zgadzają się z tymi deontologami, którzy twierdzą, że tym, co decyduje, iż czyn słuszny jest słuszny, jest jego zgodność z zasadą moralną. Zarazem jednak godzą się oni z teleologami o tyle, że uzależniają uznanie domniemych zasad moralnych od wartości ich konsekwencji. Mówi się niekiedy, że zwolennicy utylitaryzmu zasad są deontologami w odniesieniu do czynów, natomiast teleologami w odniesieniu do zasad. Utylitaryzm zasad jest więc rodzajem teorii „warstwowej”. Słuszność lub niesłuszność poszczególnych czynów określa się przez odniesienie do swoistych zasad moralnych; swoistość zaś zasad zależy od ich użyteczności. Odwołanie się w ostatecznej instancji do jednej i tylko jednej zasady moralnej — zasady użyteczności — ma uczynić utylitaryzm zasad typem utylitaryzmu, a nie formalizmu. Jednakże zasadniczym motywem rozwoju utylitaryzmu zasad nie była jedynie chęć teoretycznego pogodzenia etyki teleologicznej i deontologicznej.

Utylitaryzm zasad powstał jako alternatywa wobec utylitaryzmu czynów i jest próbą odparcia obiegowych zarzutów przeciwko temu drugiemu. Dokonana przez Rossa druzgocząca krytyka utylitaryzmu skierowana była przede wszystkim przeciwko temu, co później określono mianem utylitaryzmu czynów<sup>1</sup>. Zwolennicy utylitaryzmu zasad są niezmiernie

<sup>1</sup> W. D. Ross, *The Right and the Good*, Oxford 1930; por. zwłaszcza rozdz. II.

czuli na owe rzekomo niezgodne z intuicją implikacje utilitaryzmu czynów. „Utilitaryzm czynów, przyjąwszy nawet powszechnie uznane przez utilitarystów założenia dotyczące kryterium wartości, prowadzi w rezultacie do konsekwencji, których nie sposób uznać. Wynika zeń, że jeśli zatrudniło się chłopca do strzyżenia trawnika i po skończonej pracy prosi on o wypłatę, to powinno mu się wypłacić obiecaną sumę tylko wtedy, jeżeli nie można wymyślić lepszego zużytkowania posiadanych pieniędzy. Wynika zeń, że jeżeli przynosi się do domu swą comiesięczną pensję, to powinno się ją użyć na potrzeby swoje i swej rodziny tylko wtedy, jeżeli nie może ona być wydana z większym pożytkiem na zaspokojenie potrzeb innych ludzi. Wynika zeń nadto, że jeżeli ojciec twój jest chory i nie ma żadnej nadziei na poprawę stanu jego zdrowia, a utrzymywanie go przy życiu jest tylko stratą twojej energii i nie przysparza radości innym, wówczas — jeżeli potrafisz skrócić jego życie bez wywoływania publicznego zgorszenia i nie dając nikomu złego przykładu — jest twoim obowiązkiem podjąć odpowiednią decyzję i skrócić cierpienia ojca. Zaletą utilitaryzmu zasad jest to, iż przynajmniej w swych niektórych wersjach unika takich niepożądanych implikacji”<sup>2</sup>.

Wydaje się oczywiste, że pewien rodzaj reguł rządzących w sytuacjach zawierania umów, odpowiedzialności rodzicielskiej, troski o ludzi starszych itd. można uzasadnić, odwołując się do rozumowania utilitarystycznego. Utilitaryzm zasad utrzymuje, że gdyby kryterium użyteczności odnosić do zasad, a nie — jak zalecają zwolennicy utilitaryzmu czynów — do poszczególnych czynów, wówczas można by zbudować teorię, która odwołuje się ostatecznie do jednej i tylko jednej najwyższej zasady moralnej — zasady użyteczności — i która zarazem unika wątpliwych implikacji utilitaryzmu czynów. Zbudowaną w ten sposób teorią jest oczywiście utilitaryzm zasad. Zastosowana do konkretnych sytuacji, teoria ta ma w przybliżeniu funkcjonować następująco. (Wypada uprzedzić czytelnika, że to, co teraz nastąpi, jest zaledwie zarysem argumentacji, której z racji zwięzłości brak subtelności i wyrafinowania większości argumentów wysuwanych na rzecz utilitaryzmu zasad. Ponadto nie uwzględnia się różnic między różnymi typami utilitaryzmu zasad).

Wszelkie umowy są niezmiernie użytecznym instrumentem społecznym; są one jednak możliwe jedynie w tych społeczeństwach, w których są respektowane i w których obie umawiające się strony przypisują pewne znaczenie przestrzeganiu warunków umowy. Rozważmy więc np. dwie reguły, które mogą być odnoszone do sytuacji umowy: „Dotrzymuj warunków umowy, dopóki nie rozejrzysz się za innymi sposobami ulokowania posiadanych pieniędzy” i „Dotrzymuj warunków umowy tylko wtedy,

<sup>2</sup> R. B. Brandt, *Toward a Credible Form of Utilitarianism*, w: *Morality and the Language of Conduct*, Hector Neri Castañeda (ed.), Detroit 1963, s. 109 i n.

jeżeli nie możesz lepiej ulokować swych pieniędzy”. Zapewne żadna z tych reguł nie jest właściwą kandydatką na uznawaną przez utilitaryzm zasad zasadę obowiązującą w sytuacjach umowy. Mimo to zbadajmy je ze względów heurystycznych. Według utilitaryzmu czynów, twierdzą zwolennicy utilitaryzmu zasad (tacy np. jak Brandt), powinno się postępować zgodnie z drugą regułą; to jednak szybko doprowadziłoby do zniesienia wszelkich umów. Fakt użyteczności umów świadczy więc, iż raczej pierwsza reguła zdaje się być bardziej użyteczna na dłuższą metę; byłaby ona wobec tego regułą wyznaczającą słuszność lub niesłuszność stanowionego przez nią czynu. Według utilitaryzmu zasad, rzeczywiście zatem powinniśmy zapłacić chłopcu za strzyżenie trawnika; a niepożądana implikacja utilitaryzmu czynów, iż powinniśmy mu wypłacić obiecaną sumę tylko wtedy, gdy nie znaleźliśmy lepszej lokaty naszych pieniędzy, zostaje całkowicie uchylona.

Argument ten spotkał się z natychmiastowymi, niezmiernie różnorodnymi reakcjami. Niektórzy zwolennicy utilitaryzmu czynów po prostu zaprzeczyli zarzutom typu Brandta. Jan Narveson np. poświęcił swą wydaną niedawno książkę — *Morality and Utility*<sup>3</sup> okazaniu, że utilitaryzm czynów nie implikuje bynajmniej tak wielu intuicyjnie fałszywych konsekwencji, jak to przypisują mu jego krytycy. Inni filozofowie, zarówno utilitaryści, jak nieutilitaryści, wątpią, czy utilitaryzm zasad rzeczywiście unika lub może uniknąć „...przynajmniej niektórych spośród owych wątpliwych implikacji”<sup>4</sup>. Spór utilitaryzmu czynów z utilitaryzmem zasad doczekał się wcale obszernej literatury. Spór ten ma jednak drugorzędne znaczenie dla problemów poruszanych w niniejszym szkicu. Jednakże rozróżnienie pomiędzy utilitaryzmem czynów a utilitaryzmem zasad ma szczególną doniosłość, jako że kwestionowane przeze mnie argumenty przeciwko utilitaryzmowi czynów byłyby jawnie niedorzeczne, gdyby odnosić je do utilitaryzmu zasad.

### *O pewnym typie argumentacji przeciwko utilitaryzmowi czynów*

W kwestionowanym przeze mnie typie argumentacji zwraca się szczególnie uwagę na praktyczne trudności i paradoksy wynikające z prób stosowania utilitaryzmu czynów w konkretnych sytuacjach moralnych. Same zaś argumenty występują w rozmaitych wersjach. Słabsze wersje stosowane są zwykle, by okazać, że ze względu na trudności związane z określeniem, który spośród możliwych dla nas w danej chwili czynów rzeczy-

<sup>3</sup> J. Narveson, *Morality and Utility*, Baltimore 1967.

<sup>4</sup> Por. np. D. Lyons, *Forms and Limits of Utilitarianism*, London 1965, oraz R. E. Bales, *Act- vs. Rule-Utilitarianism*, Stanford 1968 (rozprawa doktorska).

wiecie przyczyni się do stworzenia wewnętrznie dobrych układów spraw i sytuacji nie mniej niż jakikolwiek inny alternatywny czyn, dążenie do postępowania według zaleceń utilitaryzmu czynów jest niepraktyczne lub wręcz sprzeczne samo w sobie. Mocniejsze wersje mają okazać, że każda poważna próba rozumowania w myśl zasad utilitaryzmu czynów prowadzi do *regressus ad infinitum*. Jeszcze inne odwołują się do kryterium generalizacji: „Co by było, gdyby każdy rozumował według zasad utilitaryzmu czynów?” Moim zdaniem, wszystkie te argumenty zasadzają się na notorycznym nieporozumieniu i w związku z tym przynajmniej niektóre spośród podnoszonych przez nie kwestii są w gruncie rzeczy pseudoproblemami. Sądzę, że pamiętając o wprowadzonym rozróżnieniu, łatwo będzie można dostrzec, że argumenty te wcale nie są lepsze, jak również zrozumieć, jakie są tego powody.

By wyrazić moje zamierzenia możliwie jasno, chciałbym na początek podkreślić kilka spraw.

Po pierwsze, nie lekceważę trudności związanych z praktycznym zastosowaniem utilitaryzmu czynów. Przeciwnie, zamierzam okazać, że trudności te wcale nie mają katastrofalnych skutków dla utilitaryzmu czynów, tak jak to utrzymują niektórzy krytycy.

Po wtóre, nie zamierzam w niniejszym szkicu bronić utilitaryzmu czynów przed wszelkimi możliwymi zarzutami. Nie zamierzam zwłaszcza okazać, że utilitaryzm czynów jest prawdziwy, lub też, że bardziej zbliża się do prawdy niż jakakolwiek inna teoria etyczna czy nawet jakikolwiek typ utilitaryzmu. Pragnę jedynie wykazać, że pewne rozpowszechnione ostatnio argumenty przeciwko utilitaryzmowi czynów wcale nie są najlepsze.

Po trzecie, argumenty, które chciałbym rozważyć, odnoszą się, zdaniem krytyków, do różnych rodzajów utilitaryzmu czynów. Moja krytyka dotyczy jedynie tych argumentów, które wymierzone są przeciwko utilitaryzmowi czynów jako teorii obiektywnej słuszności, tzn. takiej teorii, która utrzymuje, że czyny słuszne to takie czyny, których aktualne konsekwencje (w przeciwieństwie do konsekwencji przewidzianych) optymalnie przyczyniają się do stworzenia wewnętrznie dobrych układów spraw i sytuacji.

### *Trzy wersje przedstawionej argumentacji i kilka odpowiedzi*

Słabsze odmiany analizowanego przeze mnie typu argumentacji mogą przebiegać następująco. Ocena konsekwencji nawet najbliższych czynów jest zadaniem niezwykle złożonym. Ocena i porównywanie konsekwencji czynów alternatywnych bez wątplenia powiększa złożoność zada-

nia. Często mamy tak mało informacji lub jesteśmy tak zaangażowani osobiście, że nie możemy rzetelnie rozważyć wszelkich za i przeciw nawet wtedy, gdy usiłujemy to uczynić. Dalej, zastanawianie się i porównywanie względnych korzyści płynących z alternatywnych czynów może być zajęciem niezwykle czasochłonnym. W pewnych wypadkach okoliczności mogą być tego rodzaju, że gdybyśmy spróbowali rozważyć i porównać wszelkie korzyści alternatywnych czynów, wówczas skutecznie wybralibyśmy jedną tylko możliwość; tak jest np. w przypadku tonącego człowieka lub obietnicy, że uczyni się pewną rzecz w najbliższym czasie. Bo gdybyśmy spokojnie poczęli się zastanawiać i porównywać względne korzyści różnych przypadków udzielania pomocy tonącemu bądź nieudzielania pomocy tonącym, wówczas jednoznacznie opowiedzielibyśmy się przeciwko pomocy. Gdybyśmy więc spokojnie usiedli, by rozważyć i porównać względne korzyści płynące z wywiązywania się lub łamania obietnic, to w rzeczy samej skutecznie złamalibyśmy obietnicę. Jeżeli zatem zdarzają się przypadki, w których czyn, uznany według kryteriów utylitaryzmu czynów za słuszny, musi być dokonany szybko, jeżeli w ogóle ma być dokonany — a przypadki takie niewątpliwie istnieją — próba stosowania w takich przypadkach teorii utylitaryzmu czynów wydaje się być równoważna z wyborem przez zaniechanie — czynu, który według kryteriów utylitaryzmu czynów jest niesłuszny.

Oczywiście, istnieją również inne sytuacje. Niekiedy żadna okoliczność nie powstrzyma nas od bezustannej deliberacji. Jednakże przedstawiony argument pozostaje ważny i dopóki nie przedstawi się jakiegoś, właściwego utylitaryzmowi czynów, uzasadnienia ograniczenia czasu podejmowania decyzji, każdy możliwy czyn alternatywny, „...jakkolwiek błahy, może stać się, a według niektórych powinien stać się, przedmiotem utylitarystycznego rachunku; nie tylko te czyny, o których zazwyczaj niewiele myślę, lecz także i te, które mogę spełniać z rozmysłem, jeżeli zdecyduję się to uczynić”<sup>5</sup>. Byłoby jednak czymś bez wątpienia absurdalnym zastanawiać się przed spełnieniem wszelkich możliwych czynów błahych, włączając w to te czyny, które mógłbym spełnić, „gdybym zdecydował się to uczynić”. Czy powinienem postąpić jeden krok w lewo, czy też może dwa kroki? Krok w prawo i krok wstecz? Zawsze mamy przed sobą nieskończoną możliwość zachowań nieistotnych. Czy wszystkie one muszą być przedmiotem utylitarystycznego rachunku? Jeżeli tak, to wydaje się, że każda poważna próba zbliżenia utylitaryzmu czynów do rzeczywistości skutecznie gwarantuje, że teoria ta nigdy nie zostanie zastosowana w praktyce; wydaje się czymś nieprawdopodobnym, że spełnimy czyn nie mniej przyczyniający się do stworzenia wewnętrznie dobrych ukła-

<sup>5</sup> A. Duncan-Jones, *Utilitarianism and Rules*, „The Philosophical Quarterly” 1957, nr 7, s. 366.

dów spraw i sytuacji niż jakikolwiek inny czyn możliwy, jeśli siądziemy i zaczniemy rozmyślać o możliwościach alternatywnych, próbując rozważyć i porównać ich konsekwencje.

Mocniejsze wersje tego argumentu przedstawia się niekiedy następująco. Rozważmy drugi spośród wymienionych przypadków, to znaczy taki, w którym decyzja jest czymś nagannym. Upraszczając, przyjmijmy, że osoba działająca może spełnić dwa czyny *A* lub *B*. Który z nich powinna spełnić? Jeżeli sprawca czynu jest konsekwentnym utylitarystą czynów, oceni i porówna możliwe konsekwencje *A* i *B*, i spełni ten czyn, który ma prawdopodobnie lepsze konsekwencje. Będzie więc, krótko mówiąc, kalkulował. Jednakże czynność kalkulacji jest sama w sobie czynnością, którą sprawca może wybrać lub zaniechać. Tak więc na scenie pojawia się trzeci czyn *C* — czyn kalkulacji. Czy sprawca ma więc spełnić *A*, *B*, czy może *C*? Jeżeli jest on konsekwentnym zwolennikiem utylitaryzmu czynów, wszystkie te możliwości będą dla niego wystarczającą sposobnością do zastanowienia się; w związku z tym pojawia się możliwość czwarta — *D*, będąca czynem zastanawiania się nad możliwymi konsekwencjami *A*, *B* i *C*. Jednakże *D*, jako możliwość alternatywna, może być także przedmiotem refleksji. Sprawca znajduje się zatem w sytuacji *regressus ad infinitum*.

Inna wersja tego argumentu prowadzi do regresu w nieco odmienny sposób. Ponieważ w każdym momencie mamy przed sobą nieskończenie wiele czynów możliwych, zarówno istotnych, jak nieistotnych, możemy zatem zastanawiać się dowolnie długo, który spośród owych czynów spełnić. Zgodzimy się, że dowolnie długie zastanawianie się nad błahostkami jest czymś absurdalnym; jeżeli jednak zwolennik utylitaryzmu czynów musi zastanowić się nad ograniczeniem czasu deliberacji — co niewątpliwie czyni, — jeżeli ma uzasadnić ograniczenie czasu deliberacji w myśl zasad utylitaryzmu czynów — wówczas ponownie popada w *regressus ad infinitum*.

Obie wersje przedstawionego argumentu mają podkreślić niezbędnosć reguł podejmowania decyzji moralnych. Słabsza wersja akcentuje praktyczną niezbędnosć reguł, mocniejsza — niezbędnosć teoretyczną; stąd też rozróżnienie pomiędzy słabszą i mocniejszą wersją argumentu.

W odpowiedzi na przedstawione argumenty zwolennicy utylitaryzmu czynów wprowadzają najczęściej pojęcie „reguł praktycznych” (*rules-of-thumb*). Są to praktycznie nieodzowne wytyczne postępowania, zazwyczaj respektowane przez nas na zasadzie przyzwyczajenia, które z punktu widzenia teorii nie mają najmniejszego znaczenia dla określenia słusznosci lub niesłusznosci czynu. Zwolennicy utylitaryzmu czynów mogą oczywiście twierdzić (i często tak czynią), że poprzez doświadczenie odkrywamy, iż dotrzymywanie obietnic, prawdomówność, pomoc bliźniemu w chwili-

li nieszczęścia itd., są tymi rodzajami czynów, które na ogół prowadzą do lepszych wewnętrznie układów spraw i sytuacji niż np. powstrzymywanie się od takich czynów w sytuacjach, w których byłoby rzeczą właściwą je spełnić. Aby więc ułatwić sobie decyzję i działanie, konstruujemy takie reguły praktyczne, jak: „Nie kłam”, „Nie łam przyrzeczeń”, „Pomagaj swoim bliźnim, gdy spotka ich nieszczęście” itd., których respektowanie jest, według utylitaryzmu czynów, całkowicie uzasadnione. Jest ono uzasadnione na mocy tych samych racji, jakie wysuwali krytycy w przytoczonych argumentach. Oczywiście żaden zwolennik utylitaryzmu czynów nie będzie twierdził, że słuszność lub niesłuszność czynu zależy od reguły praktycznej bądź że reguły te określają lub stanowią nasze obowiązki. Ostatecznie dla rozstrzygnięcia kwestii słuszności czy niesłuszności poszczególnego czynu jest istotne tylko to, czy przyczynia się on do stwarzania wewnętrznie dobrych układów spraw i sytuacji nie mniej niż jakikolwiek czyn alternatywny. Mimo to zwolennik utylitaryzmu czynów może przyznać rację swoim krytykom, iż są to reguły praktycznie niezbędne. „Prawdą jest — pisze np. Smart — że możemy świadomie przyzwyczajać się do zachowania zgodnego z pewnymi zasadami, takimi jak np. zasada dotrzymywania obietnicy, w przekonaniu, że postępowanie w myśl tych zasad jest na ogół bardziej użyteczne, wiedząc nadto, że częstokroć nie mamy po prostu czasu, żeby rozważyć indywidualne za i przeciw. Naturalnie, postępując zgodnie z przyzwyczajeniem, nie zastanawiamy się, ani tym bardziej nie podejmujemy decyzji. Zwolennik utylitaryzmu czynów uznaje owe zasady jedynie za zwykłe zasady praktyczne i stosuje je jako przybliżone wytyczne postępowania (*rough guides*). Zazwyczaj postępuje on według nich wtedy, jeżeli nie ma czasu na zastanawianie się nad możliwymi konsekwencjami lub gdy pożytek płynący z zastanowienia jest niewspółmiernie mniejszy od ewentualnej niekorzystnej straty czasu”<sup>6</sup>. Ale krytycy odpowiadają zwykle, że wprowadzając pojęcie zasad praktycznych, utylitaryzm czynów po prostu przesuwając problem na dalszy plan. Zwolennik utylitaryzmu czynów mówi wówczas o zasadach praktycznych w ogóle. Jednakże w poszczególnych przypadkach musi on nadal zastanawiać się, czy postępować według zasady praktycznej, czy też rozważyć dany przypadek niezależnie od niej. Jeżeli nie zastanowi się nad konkretnym przypadkiem, trudno pojąć, w jaki sposób uzasadni konieczność respektowania zasady praktycznej odnoszącej się do tego przypadku. Jeśli zaś będzie się zastanawiał, na powrót wyłoni się problem *regressus ad infinitum*<sup>7</sup>.

<sup>6</sup> J. J. C. Smart, *Outline of a System of Utilitarian Ethics*, London 1961, s. 30.

<sup>7</sup> W opublikowanym niedawno podręczniku (*Moral Philosophy*, New York, London 1967) R. T. Garner i B. Rosen z Ohio State University przeprowadzają podobne rozumowanie. A. Duncan-Jones (op. cit.) kwestionuje możliwość właściwego utylitaryzmowi czynów uzasadnienia tych zasad.



Trzecia wersja rozważanego argumentu polega na stosowaniu kryterium generalizacji: „Co by było, gdyby każdy rozumował zgodnie z zasadami utylitaryzmu czynów?” Przykład argumentacji tego rodzaju przedstawia, jak sądzę, R. B. Brandt w swej *Ethical Theory*<sup>8</sup>. W zasadzie przedstawia on kilka argumentów, które sprowadzają się ostatecznie do następującego rozumowania. Brandt podaje przykład Francuza myślącego według zasad utylitaryzmu czynów, który, przebywając podczas wojny w Anglii, stoi wobec następującego problemu: czy respektować zalecenie rządu dotyczące ograniczeń w korzystaniu z gazu i elektryczności i utrzymać maksymalną temperaturę w swym pokoju 50° F, czy też zużywając odpowiednio więcej gazu i elektryczności, podnieść temperaturę do 70° F. Przyjmijmy, że rozumuje on tak: „Zdecydowana większość Anglików postąpi zapewne zgodnie z zaleceniem rządu. Zatem, jeżeli tylko kilka osób wraz ze mną zużyje nieco więcej gazu i elektryczności, by utrzymać swe mieszkania w cieple, działania na rzecz ukończenia wojny nie poniosą najmniejszego uszczerbku, a osobom tym będzie bez porównania przyjemniej. Jeżeli zatem skorzystam z gazu i elektryczności, zwiększy się dobro ogółu — stan, w którym kilku osobom jest cieplej, jest bez wątpienia lepszy niż stan, w którym nikomu nie jest ciepło”. Zdaniem Brandta, jest to rozumowanie całkowicie zgodne z teorią utylitaryzmu czynów i wskazuje ono na poważny błąd w tej teorii. Utylitaryzm czynów nie może mianowicie uznać za nasz obowiązek przeświadczenia o równym podziale obciążeń w sytuacji, kiedy dobro ogółu wymaga zgodnego poświęcenia od wszystkich. Proponuję nazwać ten argument „Brandtowskim argumentem z maksymalizacji dobra”.

Brandt idzie dalej, twierdząc, że przykład ten nasuwa dodatkowe trudności. Przypuśćmy, że Francuz będzie kontynuował swe rozumowanie następująco: „Gdyby wystarczająca liczba innych osób zdecydowała się skorzystać z gazu i elektryczności, wówczas wojna zostałaby przegrana, a moje powstrzymanie się od użycia gazu i elektryczności nie miałoby najmniejszego wpływu na przebieg wojny; miałyby to jednak istotną różnicę dla mej wygody. Zło ogółu zmniejszy się zatem, jeżeli zużyję nieco więcej gazu i elektryczności, żeby ogrzać swe mieszkanie. Powinienem więc skorzystać z gazu i elektryczności”. Brandt zdaje się być przekonany, że argument ten jest również prawomocny i że również ujawnia poważną lukę w utylitaryzmie czynów; utylitaryzm czynów dopuszcza mianowicie uzasadnienie zachowań egoistycznych zawsze, ilekroć sprawca czynów uzyskuje jakąś istotną korzyść z danego zachowania. Proponuję określić ten argument mianem „Brandtowskiego argumentu z minimalizacji zła”.

Istotą obiekcji Brandta zdaje się być przekonanie, że w przypadkach,

<sup>8</sup> R. B. Brandt, *Ethical Theory*, Englewood Cliffs 1959.

gdzie dla osiągnięcia pożądanego celu niezbędne jest powszechne współdziałanie, zwolennik utylitaryzmu czynów występuje z uzasadnieniem zachowań egoistycznych bez względu na cele tych zachowań. W przypadku Francuza, gdyby wojna została wygrana, przyczyniłby się on do zwiększenia dobra ogółu zwiększając swoje własne dobro, gdyby zaś wojna została przegrana, zmniejszyłby powszechne zło. Zatem niezależnie od wyników wojny — zwycięstwa lub klęski — powinien skorzystać z gazu lub elektryczności.

Argumentacja ta jest niewątpliwie godna zastanowienia. Nie jest jednak jasne, czy rzeczywiście jest ona przypadkiem rozważanego przeze mnie typu argumentacji. Na pewno jednak stanowi podstawę dla tego, co wydaje mi się bezsprzecznie trzecią wersją rozważanego typu argumentacji.

Załóżmy — pisze Brandt — że każdy Anglik będzie rozumował tak samo jak Francuz, wówczas każdy będzie mógł uzasadnić zachowanie egoistyczne i wojna zostanie przegrana. Wynika z tego wniosek, że „gdyby każdy respektował zasady rozumowania właściwego utylitaryzmowi czynów, wówczas wojna zostałaby przegrana i miałyby to przykre skutki dla każdego. Tak więc powszechne respektowanie dyrektywy utylitaryzmu czynów nakazującej dążyć do dobra ogółu łatwo może przynieść ogółowi ogromne zło”<sup>9</sup>. W tym wypadku istota obiekcji Brandta leży w przekonaniu, że należy odrzucić teorię utylitaryzmu czynów, ponieważ gdyby każdy rozumował w myśl zasad utylitaryzmu czynów, mogłoby to doprowadzić do zła i nieprawości. Argument ten będę nazywał „Brandtowskim argumentem z powszechnego rozumowania zgodnie z zasadami utylitaryzmu czynów”.

Kiedy J.J.C. Smart, broniąc utylitaryzmu czynów, analizuje Brandtowski argument z powszechnego rozumowania według utylitaryzmu czynów, zarzuca Brandtowi, iż „nie potrafi dostrzec, że Francuz posłużyłby się zapewne w swym rozumowaniu jako empiryczną przesłanką przypuszczeniem, że zaledwie kilka osób mogłoby myśleć tak samo jak on. Przepuszczalnie byłiby to zwolennicy tradycyjnej, nieutilitarystycznej moralności”<sup>10</sup>. Smart przyznaje jednak, że w społeczeństwie złożonym całkowicie ze zwolenników utylitaryzmu czynów Francuz znalazłby się w kłopotcie. Nie mógłby nadal liczyć na swych towarzyszy, iż będą ślepo posłuszni zaleceniom rządu, a ponieważ wiedza o tym, co uczynią inni, jest dlań niezbędną przesłanką, by rozważyć prawdopodobne konsekwencje możliwych dlań czynów alternatywnych, nie byłby on w stanie przewidzieć przebiegu swojego postępowania. Ponadto każdy spośród jego towarzyszy-utilitarystów jedzie na tym samym wózku: każdy, by móc

<sup>9</sup> Ibid., s. 43.

<sup>10</sup> J. J. C. Smart, op. cit., s. 42 i n.

przeprowadzić własne rozumowanie, musi mieć przesłankę dotyczącą zachowania innych (w tym także Francuza). „W sytuacji tej istnieje błędne koło — pisze Smart — które rzeczywiście wymaga odwołania się do metod teorii gier”<sup>11</sup>.

W dalszej części swych wywodów Smart przedstawia zapożyczone z teorii gier rozwiązanie problemu Francuza znajdującego się w społeczeństwie utylitarystów czynu. Sądzi on, że „rzecz nie ma zbyt wielkiego znaczenia praktycznego [...], jest jednak interesująca dla teoretycznego zrozumienia etyki”<sup>12</sup>.

Wyznać muszę, że niewiele oczekuję, przedstawiając interesujące i pomysłowe odwołanie się Smarta do teorii gier; nie widzę bowiem żadnej istotnej różnicy, w przypadku rozważanego problemu, pomiędzy odwołaniem się do teorii gier a odwołaniem się do zasad praktycznych. W obu przypadkach próbuje się przewyciężyć zarzut, że „rozumowanie według zasad utylitaryzmu czynów” jest sprzeczne samo w sobie. Jestem przekonany, że dociekliwy krytyk, gdyby mu na tym zależało, mógłby przedłożyć argument okazujący, iż próba zastosowania teorii gier na użytek utylitaryzmu czynów jest tak samo niepraktyczna, wewnątrznie sprzeczna i paradoksalna, jak próba odwołania się do zasad praktycznych.

*Istota problemu: czego oczekujemy od teorii etycznej?*

Twierdzę, że wszystkie przedstawione argumenty, płynące z tzw. rozumowania właściwego utylitaryzmowi czynów, są błędne — nie tyle w szczegółach, co w założeniu. Jestem przekonany, że zasadzają się one na powszechnym pomieszaniu pojęć, które wymaga wyjaśnienia. Proponuję więc przeciąć ów gordyjski węzeł, stawiając proste i jednoznaczne pytanie: co w ogóle uznajemy za cel teorii etycznej, a w szczególności co uznajemy za cel utylitaryzmu czynów? Czego właściwie oczekujemy od teorii etycznej? Jaką rolę powinna pełnić?

Kiedy Platon kazał Sokratesowi pytać Eutyfrona — „Co to jest zbożność?” — pierwsza odpowiedź Eutyfrona, jak można było oczekiwać, polegała na wyliczeniu czynów i rodzajów czynów, uznawanych przezeń za zbożne. Sokrates odrzuca tę odpowiedź, ponieważ, jak powiada, nie pytał o przykłady zbożności, lecz o ogólną ideę. Odrzuca pierwszą odpowiedź, ponieważ jest (przynajmniej w części) niepełna. Nie mówi bowiem nic o tym, co jest wspólne wszystkim rzeczom zbożnym. Nie jest odpowiedzią ogólną.

<sup>11</sup> Ibid., s. 43.

<sup>12</sup> Ibid.

Wyłaniające się z drugiej odpowiedzi Eutyfrona pojęcie zbożności jest bardziej zbliżone do oczekiwań Sokratesa. „Cokolwiek jest miłe bogom, to jest zbożne, a co niemile, to bezbożne”. Stwierdzenie to zdaje się korygować przynajmniej jedną z wad poprzedniego określenia. Nowa formuła zdaje się być całkowicie ogólna. Wskazuje na cechę lub bardzo złożony kompleks cech przysługujących tylko i wyłącznie rzeczom zbożnym. Jedynie zbożne rzeczy są umiłowane przez wszystkich bogów. Jednakże Sokrates stawia teraz bardziej docieklive pytanie: „Czy rzeczy zbożne są umiłowane przez wszystkich bogów, ponieważ są zbożne, czy też dlatego są zbożne, ponieważ są umiłowane przez wszystkich bogów?” Waga pytania jest oczywista. Gdyby rzeczy zbożne były umiłowane przez wszystkich bogów tylko dlatego, że są zbożne, wówczas zbożność ich byłaby logicznie wcześniejsza niż bycie umiłowanym przez wszystkich bogów, a bycie umiłowanym przez wszystkich bogów nie byłoby istotną cechą zbożności. Jeżeli jednak rzeczy zbożne są zbożne, ponieważ są umiłowane przez bogów, wówczas bycie umiłowanym przez wszystkich bogów jest istotą zbożności. Godząc się, iż rzeczy zbożne są umiłowane przez wszystkich bogów, ponieważ są zbożne, Sokrates i Eutyfron odrzucają drugie rozwiązanie jako nieprawdziwe. Sądzą, iż jest ono nieprawdziwe, ponieważ mimo iż może wskazywać na cechę lub zbiór cech, które mogą przysługiwać tylko rzeczom zbożnym, nie wskazuje jednak na cechę lub zbiór cech, które przysługują tylko rzeczom zbożnym dzięki temu, iż są one zbożne. Słowem, Sokrates i Eutyfron nie wierzą, że jest to wskazanie cech decydujących o zbożności czynu.

Przykład ten stanowi znakomitą ilustrację zadań, jakie niektórzy filozofowie przypisują teorii etycznej. Ma ona określić cechy decydujące o zbożności, dobroci czy słuszności czynu. Filozofowie wiążący z teorią etyczną takie oczekiwania będą traktować utylitaryzm czynów jako koncepcję cech decydujących o słuszności czynu.

Ale są także tacy filozofowie, którzy oczekują od teorii etycznej innych, znacznie ważniejszych rzeczy. Jeśli np. Smart przywołuje w sukurs teorie gier, to wydaje się, iż poszukuje on takich metod, które, gdyby je stosować, dostarczyłyby nam na użytek praktyki właściwych i użytecznych odpowiedzi na pytania typu: „Czy powinienem w tym szczególnym przypadku skorzystać z dodatkowej racji gazu i elektryczności, by ogrzać swe mieszkanie, czy też nie powinienem?” Kluczem jest w tym wypadku słowo „użytecznych”: bo jeśli utylitaryzm czynu przedstawi adekwatną koncepcję cech decydujących o słuszności czynu, dostarczy tym samym właściwej odpowiedzi na pytanie tego typu. Właściwą odpowiedzią na postawione pytanie będzie więc: „Wtedy i tylko wtedy, jeżeli czyniąc to przyczynisz się do stworzenia wewnętrznie dobrych układów spraw i sytuacji nie mniej, niż gdybyś spełnił jakikolwiek inny czyn alternatyw-

ny”, bądź też — jako że jest to raczej niewygodny sposób wysłowienia — „Wtedy i tylko wtedy, jeżeli czyniąc to zwiększysz użyteczność”. Niestety, w przypadkach indywidualnych odpowiedź ta może się wydać bezużyteczna, jako że ktoś mógłby odrzec: „Tak, lecz nie to chcę wiedzieć. Chcę wiedzieć, który spośród czynów alternatywnych powinienem spełnić, a tego mi nie powiedziano”. W istocie jednak utilitaryzm czynów mówi nam, jaki czyn należy spełnić. Nie mówi natomiast, czy rzeczywiście użyteczność skorzystania w tym wypadku z gazu i elektryczności będzie mniejsza niż w wypadku nieskorzystania, jednakże stwierdza, którą z obu rozważanych ewentualności wybrać — tę, która zwiększy użyteczność. To, czy odpowiedź ta jest prawdziwa, nie ma w tym wypadku najmniejszego znaczenia. Ale nawet gdyby tak było, powinniśmy znać metodę, dzięki której moglibyśmy określić w poszczególnym przypadku, w warunkach bezpośrednio użytecznej wiedzy o sytuacji, która z ewentualności rzeczywiście zwiększy użyteczność. Słowem, winniśmy znać zasady podejmowania decyzji. Jest to jeszcze inny rodzaj oczekiwań, jakie wiążą się z teorią etyczną — winna ona dostarczać zasad podejmowania decyzji.

Nie są to jedyne możliwości, które się przed nami rozpościerają. Wolno nam oczekiwać, by teoria etyczna dostarczała pewnego rozeznania zasad racjonalnego sądu moralnego, wypowiedzanego bezinteresownie i świadomie przez osobę dojrzałą i dobrze poinformowaną. Brandt w swym szkicu *Some Merits of One Form of Rule-Utilitarianism*<sup>13</sup> zdaje się mieć nadzieję, że jego koncepcja „idealnego kodeksu moralnego” będzie taką właśnie teorią. Inni filozofowie zdają się oczekiwać, że teoria etyczna poinformuje nas o sposobach uzasadniania sądów moralnych stosowanych przez zwykłych ludzi bądź też opiszem moralne przekonania zwykłych ludzi albo zanalizuje znaczenia najistotniejszych terminów moralnych.

Przyznaję więc, że być może nie ma *a priori* żadnych racji, dla których teoria etyczna nie mogłaby spełnić wszystkich tych oczekiwań, chociaż, o ile mi wiadomo, w rzeczywistości żadnej teorii nie udało się tego osiągnąć. Jestem przekonany, że istnieją dobre racje, ażeby uznać, że żadna teoria etyczna nie może spełnić jednocześnie wszystkich wiązanych z nią oczekiwań. Możliwe jednak, że mogłaby to uczynić teoria uwzględniająca wiele aspektów; teoria, która w jednym ze swych wymiarów byłaby koncepcją cech decydujących o słuszności czynu, w innym metodą podejmowania decyzji, w jeszcze innym — opisem uzasadniania sądów moralnych przez zwyczajnych ludzi itd. Skłonny jestem zgodzić się nadto, że przy pewnych możliwych założeniach moglibyśmy wywodzić np. koncepcje

<sup>13</sup> R. B. Brandt, *Some Merits of One Form of Rule-Utilitarianism*, w: *The Concept of Morality*, „University of Colorado Studies, Series in Philosophy” 1967, nr 3.

cech decydujących o słuszności czynu z wiedzy o zasadach racjonalnych sądów moralnych, wypowiedzianych bezinteresownie i świadomie przez osoby dojrzałe i dobrze poinformowane. Mimo to przedstawienie koncepcji cech decydujących o słuszności czynu nie jest t y m s a m y m, co przedstawienie zasad racjonalnych sądów moralnych, wypowiedzianych bezinteresownie i świadomie przez osobę dojrzałą i dobrze poinformowaną, ani też nie jest tym samym, co zbudowanie teorii podejmowania decyzji.

*Utylitaryzm czynów: teoria cech decydujących o słuszności czynu  
czy teoria podejmowania decyzji?*

Chciałbym podkreślić, i stanowi to główne założenie mojego wywodu, że teoria typu utilitaryzmu czynów może spełnić tylko jedno z wymienionych oczekiwań; może to uczynić, nie będąc jednak w stanie spełnić żadnego z pozostałych oczekiwań. Najjaskrawszym chyba tego przykładem jest zdanie analityczne, że te i tylko te czyny są słuszne, którym przysługuje cecha słuszności. Jest to zdanie wyróżniające cechę przysługującą tylko i jedynie czynom słusznym i dzięki której są one słuszne. Nie jest to chyba najużyteczniejsza teoria podejmowania decyzji, którą mógłby kierować się nasz Francuz podejmując decyzję, czy użyć gazu i elektryczności do ogrzania swego mieszkania, czy też nie. Nie jest to także, godzi się zauważyć, najciekawsza koncepcja cech decydujących o słuszności czynu.

Biorę przede wszystkim pod uwagę rozróżnienie między teoriami etycznymi, pojmowanymi jako pewne koncepcje cech decydujących o słuszności czynu itd., a teoriami rozumianymi jako pewne koncepcje podejmowania decyzji, ponieważ rozważane przeze mnie argumenty stają się bardziej wiarygodne, jeżeli rozpatrywać je łącznie. Twierdzę, że utilitaryzm czynów mógłby przedstawić pewną koncepcję cech decydujących o słuszności czynu bez jednoczesnej rezygnacji z metody, dzięki której można byłoby bez trudu uzyskiwać w praktyce prawidłowe i bezpośrednio użyteczne odpowiedzi na pytania typu: „Czy powinienem w tej sytuacji użyć gazu i elektryczności do ogrzania mieszkania, czy też nie?”

W tym miejscu należałoby zadać pytanie: „Czego właściwie oczekuje sam utilitarysta czynów od swej teorii etycznej?” Czy ma ona dostarczyć pewnej koncepcji cech decydujących o słuszności czynu, zasad podejmowania decyzji, obydwo naraz, czy czegokolwiek innego? Niestety, nie jest jasne, jaka powinna być prawidłowa odpowiedź na to pytanie. Każde ze znanych mi sformułowań utilitaryzmu czynów stwierdza, że czyn jest słuszny wtedy i tylko wtedy, jeżeli przyczyni się nie mniej niż jakikolwiek inny czyn alternatywny do zwiększenia wewnętrznie dobrych ukła-

dów zdarzeń, tzn. zwiększy użyteczność lub coś w tym rodzaju. Sprawia to na mnie rzeczywiście wrażenie koncepcji cech decydujących o słuszności czynu. Nie mogę sobie wyobrazić, aby stanowisko to mogło pretendować do tego, by informować nas w praktyce, jakież to czyny przyczynią się rzeczywiście do zwiększenia użyteczności. Jednak zwolennicy utilitaryzmu czynów niejednokrotnie zalecają pewną szczególną koncepcję podejmowania decyzji — koncepcję oceny i porównywania prawdopodobnych konsekwencji poszczególnych czynów alternatywnych. Jest to w istocie pewna koncepcja podejmowania decyzji, przypisywana najczęściej utilitaryzmowi czynów zarówno przez jego zwolenników, jak krytyków. Smart utrzymuje np., że utilitaryzm czynów jest pomyślny, przynajmniej po części, jako pewna koncepcja podejmowania decyzji. Piśze on: „Kryterium utilitarystyczne ma zatem pomagać osobie, która mogłaby uczynić różne rzeczy, gdyby zdecydowała się je uczynić, w wyborze, które spośród owych rzeczy powinna uczynić”<sup>14</sup>. Gdzie indziej zaś pisze: „utilitaryzm czynów ma dostarczyć metody podejmowania decyzji, jak należy postąpić w tych przypadkach, gdy rzeczywiście decydujemy, jak postąpić”<sup>15</sup>.

Nie roszczę sobie prawa do tego, aby stanowić, co się powinno uznać za cel utilitaryzmu czynów bądź czego powinno się oczekiwać od utilitaryzmu czynów. Rzecz jasna, zarówno ci filozofowie, którzy wysuwali roztrząsane przez mnie argumenty, jak i ci, którzy próbowali je kwestionować, skłonni są przypuszczać, iż utilitaryzm czynów stanowi zarazem koncepcję cech decydujących o słuszności czynu i teorię podejmowania decyzji, a przynajmniej są przekonani, że do tego pretenduje. Słabsze wersje wysuwanych argumentów mają okazać, że przynajmniej jeden rodzaj teorii podejmowania decyzji — koncepcja oceny i porównywania prawdopodobnych konsekwencji czynów alternatywnych — jest niepraktyczny lub wewnętrznie sprzeczny, jeśli stosować go do ustalenia, który (lub które) spośród alternatywnych czynów, możliwych dla nas w danej chwili, w warunkach bezpośrednio użytecznej wiedzy o sytuacji, winny być nazwane słusznymi, zgodnie z właściwą utilitaryzmowi czynów teorią cech decydujących o słuszności czynu. Mocniejsze wersje tego argumentu mają ukazać, że każda poważna próba stosowania takiej teorii prowadzi do błędnego koła.

Trudniej zinterpretować trzy obiekcje Brandta. Nie jest jasne, czy należy je traktować jako krytykę właściwej utilitaryzmowi czynów koncepcji cech decydujących o słuszności czynu, czy jako krytykę koncepcji oceny i porównywania prawdopodobnych konsekwencji alternatywnych, czy obu koncepcji zarazem. Rozróżnienie to może być bez znaczenia dla

<sup>14</sup> J. J. C. Smart, op. cit., s. 33 (podkr. — R. E. B.).

<sup>15</sup> Ibid., s. 31 (podkr. — R. E. B.).

naszej oceny Brandtowskich argumentów z maksymalizacji dobra i minimalizacji zła. Jest ono jednak istotne dla oceny Brandtowskiego argumentu „z powszechnego rozumowania w myśl zasad utilitaryzmu czynów”. W zależności od interpretacji argument ten może być rozumiany rozmaicie. Przy obu jednak interpretacjach, pamiętając o wprowadzonym rozróżnieniu, argument ten nie jest bynajmniej najmocniejszy.

Przypuśćmy, że będziemy rozumieli argument z maksymalizacji dobra następująco. Przyjmijmy jako przesłankę zdanie, że rzeczywiście zdecydowana większość Anglików postąpi zgodnie z zaleceniem rządu, a więc zdanie, że działania na rzecz szybkiego zakończenia wojny rzeczywiście nie ucierpią, jeżeli kilka osób ogrzeje swe mieszkania używając gazu i elektryczności; przyjmijmy też, że sytuacja, w której kilka osób ma ciepło w mieszkaniu, jest lepsza niż sytuacja, w której nikt nie ma ciepła w mieszkaniu. Zgodnie z utilitarystyczną koncepcją cech decydujących o słuszności czynu, twierdzenia te nie wydają się implikować, że byłoby rzeczą słuszną, gdyby kilka osób użyło gazu i elektryczności, a Francuz byłby jedną z nich. Jeśli jednak właściwa utilitaryzmowi czynów koncepcja cech decydujących o słuszności czynu ma takie konsekwencje, wypada się zgodzić z Brandtem, iż ujawnia on poważną lukę w tej teorii. Analogiczna interpretacja Brandtowskiego argumentu z minimalizacji zła ujawnia — jak mi się wydaje — podobny błąd we właściwej utilitaryzmowi czynów koncepcji cech decydujących o słuszności czynu. Ale przyjąwszy taką interpretację, traci rację Brandtowski argument „z powszechnego rozumowania w myśl zasad utilitaryzmu czynów”. Bo nie może być tak — jest to po prostu logicznie niemożliwe — by działania na rzecz szybszego ukończenia wojny nie poniosły żadnego uszczerbku, gdyby każdy Anglik skorzystał z gazu i elektryczności dla ogrzania swego mieszkania; gdyby rzeczywiście każdy Anglik użył odpowiednio więcej gazu i elektryczności dla ogrzania swego mieszkania, przyczyniłoby się to do przegrania wojny.

Ale możliwa jest również inna interpretacja tych argumentów. Zaczniemy od Brandtowskiego argumentu z maksymalizacji dobra. Przypuśćmy, że Francuz stoi w obliczu decyzji, czy użyć gazu bądź elektryczności, czy też nie. Pragnie on zastosować jako teorię podejmowania decyzji metodę oceny i porównywania prawdopodobnych konsekwencji możliwych dlań alternatyw działania i wobec tego rozumuje następująco: „Wydaje mi się rzeczą rozumną, by założyć, że biorąc pod uwagę charakter Anglików, zdecydowana ich większość postąpi zgodnie z zaleceniem rządu. Wydaje się również czymś rozsądnym przypuszczać, że wojna nie przedłuży się, jeżeli tylko kilka osób, włączając w to mnie, użyje nieco więcej gazu i elektryczności dla ogrzania swych mieszkań. Wydaje się rzeczą rozsądną przypuszczać, że dobro ogółu zwiększy się, jeżeli użyję gazu i elektrycz-



ności. Przyjąwszy te rozsądne założenia i właściwą utylitaryzmowi czynów koncepcję cech decydujących o słuszności czynu wnoszę więc, że byłoby rzeczą słuszną, gdybym skorzystał z gazu i elektryczności". Jeżeli tak należy rozumieć Brandtowski argument, wówczas wypada zgodzić się, że stosowana przez Francuza metoda oceny i porównywania prawdopodobnych konsekwencji alternatywnych czynów pozwala mu, na mocy przyjętych założeń, uzasadnić uchylenie się od uczestnictwa w tym, co powinno być wspólną ofiarą. Godzę się więc, że przy podobnej interpretacji Brandtowski argument z minimalizacji zła wskazuje, iż stosowana przez Francuza teoria podejmowania decyzji pozwala mu, na mocy przyjętych założeń, uzasadnić zachowanie egoistyczne. Zgadzam się także, że każdy Anglik mógłby przeprowadzić podobne jak Francuz rozumowanie z równie fatalnymi skutkami jak te, o których pisze Brandt. Jednakże przy tej interpretacji Brandtowski argument z powszechnego rozumowania w myśl zasad utylitaryzmu czynów okazuje się uogólnioną wersją krytykowanego przeze mnie argumentu w jego słabszym wydaniu<sup>16</sup>.

Trzy wersje krytykowanego przeze mnie argumentu mogą być przekonywujące wobec przypisywanej powszechnie utylitaryzmowi czynów koncepcji podejmowania decyzji. Nie sądzę jednak, aby miały one jakiegokolwiek znaczenie w sprawie prawdziwości utylitaryzmu czynów w jego potocznym rozumieniu. Bez wątpienia może być tak, że zarazem te i tylko te czyny są słuszne, które przyczyniają się do zwiększenia wewnętrznie dobrych układów spraw i sytuacji nie mniej niż jakiegokolwiek inne czyny alternatywne, i że metoda oceny i porównywania możliwych konsekwencji możliwych czynów alternatywnych jest niepraktyczna lub wewnętrznie sprzeczna dla ustalenia scharakteryzowanych w ten sposób, w warunkach bezpośrednio użytecznej wiedzy o sytuacji, czynów słusznych.

Założeniem, które zdaje się leżeć u podstaw analizowanych argumentów, założeniem, które w równej mierze przyjmują zwolennicy utylitaryzmu czynów i ich krytycy, jest przeświadczenie, że jeżeli przyjmuje się właściwą utylitaryzmowi czynów koncepcję cech decydujących o słuszności czynu, tym samym *a priori* należy uznać pewną szczególną metodę podejmowania decyzji — metodę oceny i porównywania możliwych konsekwencji czynów alternatywnych. Uznawana przez utylitaryzm czynów koncepcja cech decydujących o słuszności czynu ma określać te cechy lub

<sup>16</sup> Można by mi zarzucać, że zniekształcam w tym wypadku intencje Brandta. Proponowana przezeń w *Ethical Theory* formuła utylitaryzmu czynów odwołuje się do *maximum net expectable utility*, a on sam wyraźnie stwierdza, że woli takie sformułowanie aniżeli sformułowanie odwołujące się do *maximum actual utility*. Jednak Brandt stwierdza również wyraźnie, że przedkładane przezeń argumenty krytyczne „są przekonujące bez względu na sposób ich interpretacji”; *Ethical Theory*, s. 381.

ten sam zbiór cech, które przysługują jedynie i wyłącznie czynom słusznym i mocą których są one słuszne. Spośród dostępnych dla nas w danej sytuacji możliwości działania pragnęlibyśmy ustalić, w warunkach bezpośrednio użytecznej wiedzy o sytuacji, tę możliwość, której przysługuje owa cecha lub zbiór cech, założywszy naturalnie, że jesteśmy przekonani o prawdziwości utilitaryzmu czynów i pragniemy spełnić czyn moralnie słuszny, a sama teoria nie stawia żadnych logicznych ograniczeń dotyczących sposobu ustalania poszukiwanej alternatywy działania.

„Jednakże — mógłby ktoś oponować — skąd zwolennik utilitaryzmu czynów wie, w tym i w każdym innym przypadku, który spośród możliwych dlań czynów zwiększy użyteczność, jeżeli nie oceni i nie porówna ich prawdopodobnych konsekwencji?” Istotnie, skąd wie? Być może, jak sugeruje Smart, powinien się zdać na los<sup>17</sup>. Być może, winien zawsze postępować zgodnie z zasadami — jak sądzi Moore<sup>18</sup>. David Gauthier pisze, że powinien on postawić sobie pasjansa<sup>19</sup>. A może powinien poradzić się swego horoskopu lub Cyganki? Przyznać muszę, że nie wiem, w jaki sposób może on określić we wszystkich poszczególnych przypadkach, który spośród możliwych dlań czynów rzeczywiście zwiększy użyteczność. Mogę jednak udzielić mu pewnej rady. Mogę powiedzieć: zastanów się, jaki tryb postępowania okazał się użyteczny w tego rodzaju sytuacjach. Jeżeli nie dysponuje on żadnym pewnym sposobem określenia dla wszystkich poszczególnych przypadków, który spośród możliwych dlań czynów zwiększy użyteczność, może przynajmniej wziąć pod uwagę te sposoby postępowania, które okazały się użyteczne w przeszłości; może także zwrócić się do teoretyków, prosząc o wiarygodne zasady podejmowania decyzji, które mogłyby być stosowane na przyszłość. Wydaje mi się, że odwołanie się Smarta do teorii gier i różne powoływanie się na zasady praktyczne są przykładami tego typu posunięć. Kroki te wyraźnie zmiierzają w kierunku teorii podejmowania decyzji, a ich sukcesy lub niepowodzenia wcale jeszcze nie dowodzą, że właściwa utilitaryzmowi czynów teoria cech słusznotwórczych jest prawdziwa lub fałszywa. W utilitaryzmie czynów samo pojęcie funkcjonalnej lub niefunkcjonalnej teorii podejmowania decyzji zakłada pewną koncepcję cech decydujących o słuszności czynu. Bo czymże byłaby proponowana dla utilitaryzmu czynów niefunkcjonalna teoria podejmowania decyzji, jeśli nie teorią istotnie uniemożliwiającą nam skuteczne stwierdzenie, który spośród możliwych dla nas czynów, w warunkach bezpośrednio użytecznej wiedzy o sytuacji, ma być uznany za słuszny w myśl właściwej utilitaryzmowi czynów teorii cech decydujących o słuszności czynu?

<sup>17</sup> J. J. C. Smart, op. cit., s. 43.

<sup>18</sup> G. E. Moore, *Principia Ethica*, Cambridge 1965, s. 164.

<sup>19</sup> D. Gauthier, *Rule-Utilitarianism and Randomization*, „Analysis” 1964-1965, nr 25, s. 68-69.

Mimo że proponowana przez utylitaryzm czynów koncepcja cech decydujących o słuszności czynu nie nakłada żadnych ograniczeń logicznych na uznawane przez nią metody podejmowania decyzji, można jednak w pewnym sensie mówić, że utylitarystyczna koncepcja cech decydujących o słuszności czynu przesądza o właściwych utylitaryzmowi czynów metodach podejmowania decyzji. Przesądza o tyle tylko, że metody te nie są wiarygodnymi metodami wyznaczania w warunkach bezpośrednio użytecznej wiedzy o sytuacji, które spośród możliwych w danym czasie czynów zwiększają użyteczność. Przypominam w ten sposób często wypowiediany w utylitaryzmie czynów postulat, że użycie zasad praktycznych, reguł teorii gier itp. może być uzasadnione w obrębie utylitaryzmu czynów. Jednakże dopóki nie uda się wykazać, że uznanie zajmowanego przez utylitaryzm czynów stanowiska w sprawie cech decydujących o słuszności czynu wcale nie pociąga za sobą *a priori* określonej metody oceny i porównywania możliwych konsekwencji czynów alternatywnych bardziej aniżeli jakiegokolwiek innej metody podejmowania decyzji, teoria utylitaryzmu czynów nadal będzie wystawiona na najróżniejsze zarzuty w stylu rozważanych przeze mnie argumentów.

#### *Uwagi końcowe*

Kończąc, chciałbym uchylić dwie możliwe obiekcje. Można by sądzić, że wprowadzone przeze mnie rozróżnienie pomiędzy właściwą utylitaryzmowi czynów teorią cech decydujących o słuszności czynu i — opartą na zasadzie oceny i porównywania możliwych konsekwencji czynów alternatywnych — teorią podejmowania decyzji jest nieuprawnione ze względów historycznych. Sami zwolennicy utylitaryzmu czynów powszechnie utrzymywali, że teoria ta dostarcza nam racjonalnego sposobu rozstrzygnięcia, jak powinniśmy postępować w poszczególnych przypadkach. Oni też nalegali, by raczej brać pod uwagę fakty poszczególnych sytuacji, aniżeli „powołać” zasady. Odpowiadam: jeżeli zwolennicy utylitaryzmu czynów nie widzą różnicy pomiędzy logicznie niezależnymi teoriami — teorią cech decydujących o słuszności czynu i teorią podejmowania decyzji, to należy im to ukazać. Nie widzę powodu, dla którego współczesny utylitaryzm czynów miałby być obciążony błędami przeszłości.

Inna możliwa obiekcja może dotyczyć faktu, że nie wspomniałem nawet o tym, czego się najbardziej oczekuje od teorii etycznej, mianowicie pomocy w podejmowaniu decyzji moralnych.

Gdyby tak było, skłonni bylibyśmy mniemać, że koncepcja cech decydujących o słuszności czynu jest zarazem koncepcją podejmowania de-

czyż. Wprowadzone przeze mnie rozróżnienie byłoby więc nieuprawnione, ponieważ jest sztuczne. Innymi słowy, odróżniłem rzeczy, których w ogóle niepodobna odróżnić.

W odpowiedzi na pierwszą część obiekcji przyznaję, że istotnie oczekujemy od teorii etycznych pomocy w podejmowaniu decyzji moralnych; jednakże nie widzę żadnych racji, aby przypuszczać, że pomoc, jakiej można rozumnie oczekiwać od teorii etycznej, jest tak bezpośrednia, jaką w myśl obiekcji powinna być. Mamy nadzieję, że teorie naukowe pomogą nam w opanowaniu otaczającego nas świata, jednakże, o ile mi wiadomo, nikt dotychczas nie twierdził, że gdybyśmy w płonącym budynku zaczęli obliczać masę ciała, by dowiedzieć się, czy zabijemy się, skacząc z pierwszego piętra, z pewnością zginęlibyśmy w płomieniach; zatem prawa Newtona nie są w ogóle żadną pomocą w opanowaniu otaczającego nas świata. Nie chciałbym twierdzić, że nie ma żadnych różnic pomiędzy teorią naukową a teorią etyczną. Nie chciałbym również twierdzić, że wszelkie teorie są zbyt ogólne, aby mogły być bezpośrednio praktycznie stosowane, tak jak w myśl obiekcji winna być bezpośrednio praktycznie stosowalna teoria etyczna. Mimo to, gdybyśmy dysponowali prawdziwą teorią cech decydujących o słuszności czynu, mogłoby to być znaczną pomocą praktyczną nawet wówczas, gdyby teoria ta była ogólna. Chociaż więc teoria ta nie powie nam zapewne, jak ustalić, które spośród możliwych dla nas w danej chwili czynów są tymi, którym rzeczywiście przysługują cechy moralnie znaczące, mogłaby ona przynajmniej powiedzieć nam, czego należy poszukiwać. Mogłaby przedstawić kryterium pomiaru skuteczności lub nieskuteczności zasad praktycznych oraz polityki społecznej i prawnej — zasady zaś byłyby wtedy skuteczne, jeżeli ich akceptacja zwiększyłaby liczbę czynów uznanych przez teorię za słuszne.

Co się zaś tyczy drugiej części obiekcji, mogę powiedzieć tylko tyle: wydaje mi się rzeczą oczywistą, że postulat, iż to, co decyduje o słuszności czynów, polega na ich użyteczności, nie jest równoznaczny z postulatem, że właściwym sposobem ustalania słuszności danego czynu w sytuacji różnych możliwych dla nas w tej chwili czynów alternatywnych jest ocena i porównanie możliwych konsekwencji wszystkich czynów. Jeżeli mam rację i owe dwa postulaty nie są równoznaczne, wówczas wprowadzone przeze mnie rozróżnienie nie jest ani sztuczne, ani też nieuprawnione. A jeśli, jak sądzę, argumenty z rozumowania w myśl zasad utilitaryzmu czynów czerpią swą wiarygodność głównie z faktu, iż jedne wygrywa się przeciwko drugiemu, rozróżnienie to wydaje mi się godne wyraźnego uświadomienia.

Р. Эугене Бейльс

## ЧЕГО НАДО ОЖИДАТЬ ОТ УТИЛИТАРИЗМА ПОСТУПКОВ?

В последнее время принято отличать утилитаризм поступков (*act-utilitarianism*) от утилитаризма принципов (*rule-utilitarianism*). Можно сказать, что согласно утилитаризму поступков конкретный поступок (в противоположность типу или классу поступков) является правильным тогда и только тогда, если его полезность — т. е. вклад в состав внутреннее добрых предприятий (*contribution towards intrinsically good states of affairs*) — оказывается не меньше какого-нибудь другого возможного в этом плане поступка. По утилитаризму принципов поступок является правильным тогда и только тогда, если он согласован с принципом основанным на полезности. Целью статьи является попытка оценки аргументов, которые в последнее время даются в поддержку тезиса, что утилитаризм поступка не может быть адекватной этической теорией. Общей чертой этих аргументов является указание на практические затруднения и парадоксы, связанные со всякими попытками применения утилитаризма поступков к решению конкретных моральных проблем.

Можно выделить три типа этих аргументов: «слабший», «сильнейший» и «обобщённый». Слабую версию даёт с целью показания, что имея в виду трудности связанные с определением поступка, который среди возможных в данный момент поступков, способствует образованию внутреннее добрых составов предприятий, но не меньше какого-нибудь другого (из возможных) поступка, деятельность согласно утилитаризму поступков является практически невозможной и внутреннее противоречивой. Задачей сильнейшей версии является доказательство, что каждая более серьёзная попытка рассуждения согласно принципам утилитаризма поступков неухронно ведёт к *regressus ad infinitum*. Версия обобщённая заключается в ссылке к тексту: «что было бы, если бы каждый рассуждал согласно принципам утилитаризма поступков?» Автор подробно анализирует так самые аргументы, как и типичные контраргументы. В самой аргументации, так и в контраргументации имеем дело — по его мнению — с определённым характеристическим смещением понятий, которое причисляется к тому, что рассматриваемые аргументы являются необоснованными.

В статье выделяется несколько типов логически независимых концепции, которые — как считает автор — могут найти применение в этике. Это концепция правильнообразующих примет (*right-making characteristics*), концепция принятия решений (*decision-making procedure*) и концепция распознавания мотивов рационального морального суждения, высказываемого бескорыстно и сознательно лицом зрелым и хорошо поинформированным (*account of the considered moral judgements of informed, matured persons in their disinterested, reflective moments*).

Факт, что эти концепции являются логически независимыми, ведёт к заключению, по мнению автора, что данная этическая теория может быть истинной теорией правильнообразующих примет, не будучи в то же время теорией принятия решений. Далее, автор считает, что все рассматриваемые им аргументы критиков, стремящихся доказать, что приписываемая утилитаризму поступков теория принятия решений (т. е. процедура оценки и сравнения возможных следствий альтернативных поступков), понимаемая как метод определения, который из поступков в данный момент для нас возможных должен считаться правильным согласно свойственной утилитаризму поступков концепции правильнообразующих примет, является непрактической и внутреннее противоречивой. Такие аргументы, по мнению автора, если даже являются обоснованными, не доказывают ещё, что предлагаемую утилитаризмом поступков концепцию правильнообразующих примет надо считать неверной. Поскольку в большинстве формулировок утилитаризм поступков явится как теория правильнообразующих примет и поскольку признание свойственной утилитаризму

поступков концепции правильнообразующих примет не ведёт за собой *a priori* необходимости признания теории принятия решений, то отсюда нельзя быть уверенным, имеют ли эти аргументы существенное значение для определения истинности утилитаризма поступков в его типичных формулировках.

## R. Eugene Bales

### WHAT DO WE EXPECT OF ACT-UTILITARIANISM?

In recent years, act-utilitarianism has been distinguished from rule-utilitarianism. We may say (roughly) that act-utilitarianism is the thesis that a particular act (as opposed to a type of act or class of acts) is right if and only if its utility — that is contribution towards intrinsically good states of affairs — is no less than that of some alternative act. Rule-utilitarianism is (roughly) the thesis that an act is right if and only if it conforms to a rule somehow grounded in utility. The present paper concerns one type of argument sometimes used as an attempt to show that act-utilitarianism cannot be an adequate ethical theory. Arguments of this type are characterized by an emphasis on practical difficulties involved in, or paradoxes arising out of, the attempt to apply act-utilitarianism theory to concrete moral situations.

Three forms of this type of argument may be distinguished: "weak", "strong", and "generalized". "Weak" forms are used to attempt to establish that, because of difficulties involved in determining which of the acts open to us at a given time would contribute towards intrinsically good states of affairs no less than would any alternative, the attempt to act as an act-utilitarian is impracticable or self-defeating. "Strong" forms are used to attempt to establish that any serious attempt to deliberate along act-utilitarian lines forces the deliberator into a vicious regress. "Generalized" forms employ a generalization test, "What would happen if everyone deliberated along act-utilitarian lines?" Arguments of these three forms are developed in the paper in some detail, and some standard replies are noted. The thesis set forth is that a systematic confusion runs throughout all the arguments and replies, and that the arguments do not establish what they were intended to establish.

Several logically independent procedures, all of which can and do find their way into ethics, are distinguished: providing accounts of right-making characteristics, providing decision-making procedures, providing accounts of the considered moral judgements of informed, mature persons in their disinterested, reflective moments, and some others. The author argues that, because the procedures are logically independent; a proposed ethical theory could provide, for example, a true account of right-making characteristics without providing, for example, a decision-making procedure. He then notes that all the arguments of the type he has been considering are used to attempt to discredit act-utilitarianism by showing that a decision-making procedure frequently associated with act-utilitarianism, *viz.*, the procedure of estimating and comparing probable consequences of alternative acts, is impracticable or self-defeating as a method for picking out, from the alternatives open to us at a given time, the one or ones (under practically

helpful descriptions) which would be pronounced right by the act-utilitarian account of right-making characteristics. The author points out that such arguments, even if valid, do not show that act-utilitarian account of right-making characteristics is false. Moreover, because most formulations of act-utilitarianism appear to be accounts of right-making characteristics, and because acceptance of the act-utilitarian account of right-making characteristics does not commit one *a priori* to the decision-making procedure of estimating and comparing probable consequences of alternative acts, if it is not clear that such arguments are even relevant to the question of whether act-utilitarianism, as it is usually formulated, is true.