

KLEMENS SZANIAWSKI

Teoria decyzji a etyka

W artykule niniejszym rozważona zostanie przydatność aparatury pojęciowej oraz twierdzeń teorii decyzji do analizy zagadnień etycznych. Podjęcie takiego tematu daje się uzasadnić bez szczególnych trudności. Decyzja etyczna (tzn. taka, która motywowana jest — m. in. przynajmniej — względami etycznymi) jest szczególnym przypadkiem decyzji w ogóle. Można więc oczekiwać, iż ogólne własności racjonalnego podejmowania decyzji przysługiwać będą tym przypadkom wyboru działania, które etyk wyróżnia jako specyficznie go interesujące.

Wydaje się ponadto, iż język teorii decyzji szczególnie dobrze się nadaje do konstruowania prostych, laboratoryjnych niejako, modeli reguł działania. Reguły te mogą być z kolei oceniane w świetle postulatów mających treść etyczną, a formę dostatecznie precyzyjną, by ich niesprzeczność (*resp.* sprzeczność) podlegała ścisłemu dowodowi. Uzyskane w ten sposób wyniki przyczyniają się, jak sędzę, do klaryfikacji pojęć, których dręcząca nieostrość jest faktem dobrze znanym.

Termin „teoria decyzji” wymaga paru słów wyjaśnienia. W sensie węższym jest to dyscyplina zajmująca się metodami optymalizacji, czyli wyznaczania działań, które przynoszą skutki najwyżej cenie pod jakimś ustalonym względem. Niewiele upraszczając, można by powiedzieć, iż jest to prakseologia matematyczna. W tym artykule jednak będę teorię decyzji pojmował szerzej, zaliczając do niej również takie dziedziny badań, jak teoria gier oraz dyscypliny pokrewne. Wspólną cechą tych dyscyplin jest język, w którym zawsze występuje element wartościowania oraz wyboru. Niektóre działy teorii decyzji dają się interpretować dwojako: opisowo lub normatywnie. Czym innym jest, oczywiście, twierdzenie, że ludzie podejmują decyzję wedle zasady maksymalizowania przeciętnej wartości, czym innym zaś — norma zalecająca im decydowanie wedle takiej właśnie zasady. W poniższych rozważaniach będę miał z reguły na uwadze interpretację normatywną.

Kończąc te wstępne wyjaśnienia pragnąłbym wreszcie zaznaczyć, iż głównym moim celem jest wskazanie, co etyk może uzyskać dzięki aparaturze pojęciowej teorii decyzji, a nie — czego uzyskać nie może. Rejestr doniosłych problemów etycznych nie poddających się sformułowaniu „decyzyjnemu” byłby zapewne długi, nie będę go tu jednak zestawiać.

Indywidualne podejmowanie decyzji

Wyniki uzyskane w tym zakresie odnoszą się do sytuacji następującej. Osobnik podejmujący decyzję (nazwiemy go decydentem) uprzytamnia sobie możliwe konsekwencje swych działań i potrafi je ocenić, nie wie jednak na pewno, która z tych konsekwencji się zrealizuje. Powstaje kwestia: jakie kryterium wyboru działania zasługuje na miano racjonalnego? Odpowiedź zależy od pewnych warunków dodatkowych.

Jeśli decydent potrafi ocenić względne szanse realizacji następstw poszczególnych działań, powinien wybrać to działanie, które przynosi skutki o najwyższej wartości przeciętnej. Zalecenie powyższe ma charakter — w pewnym sensie — tautologiczny, można bowiem wykazać, iż działanie w ten sposób wyróżnione jest w rozumieniu decydenta najwartościowsze.

Jeśli natomiast decydent nie dysponuje wiedzą, która pozwoliłaby mu przypisać prawdopodobieństwa konsekwencjom poszczególnych działań, to nie istnieje kryterium, o którym dałoby się bezspornie udowodnić, że jest najlepsze. Proponowane kryteria (Laplace, Wald, Hurwicz, Savage, Szaniawski) można oceniać w świetle postulatów, które łącznie konstytuowałyby sens pojęcia „racjonalnej decyzji”. Okazuje się jednak, iż owe postulaty są sprzeczne. Dowodzi to niespójności potocznych intuicji na temat racjonalności i zmusza do takiej rewizji znaczeniowej tego pojęcia, która polegałaby w każdym razie na jego treściowym ograniczeniu. Problem, z jakich postulatów racjonalności zrezygnować, musi być z konieczności rozstrzygnięty arbitralnie. Arbitralność ta przenosi się — rzecz jasna — na kryterium, które uznamy za racjonalne w zmodyfikowanym sensie tego słowa.

Pod względem formalnym problem indywidualnej decyzji reprezentowany jest przez układ (tzw. macierz) liczb: $\|u_{ij}\|$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$. Wiersze reprezentują możliwe do podjęcia działania, kolumny zaś — stany rzeczy, od których zależą skutki owych działań. Liczba u_{ij} jest wyrazem oceny (z punktu widzenia decydenta) sytuacji zrealizowanej przez działanie i przy zajęciu stanu rzeczy j .

Jeśli decydent potrafi przypisać poszczególnym stanom rzeczy prawdopodobieństwa ich zajęcia, powiedzmy p_j , to dla każdego działania można

ustalić przeciętną wartość, jaką ono realizuje. Ogólnie, dla działania i wartość ta wynosi: $\sum_{j=1}^n u_{ij}p_j$. Racjonalne jest wówczas podjęcie działania, które tę wartość przeciętną maksymalizuje. Mówimy w odniesieniu do takiej sytuacji o podejmowaniu decyzji w warunkach ryzyka. Decyzją w warunkach niepewności nazywa się wybór działania, gdy informacja probabilistyczna nie jest dostępna. Rozumowanie przebiega wówczas mniej więcej tak.

Kryterium decyzji przyporządkowuje każdemu problemowi jego rozwiązanie. Rozwiązaniem tym jest (niepusty) podzbiór zbioru wszystkich działań, zwany optymalnym w sensie tego kryterium. Oto przykład kryterium decyzji (zwanego „maximin” i pochodzącego z teorii gier): każde działanie scharakteryzować można przez najmniejszą wartość dającą się za pomocą tego działania realizować, czyli $\min_{1 \leq j \leq n} u_{ij}$; optymalne w sensie „maximin” jest to działanie, dla którego ów dolny kres wartości jest największy.

A oto przykład postulatu pod adresem wszelkiego kryterium decyzji (postulat taki stanowi częściową charakterystykę pojęcia racjonalności). Jeśli jakieś działanie jest nieoptymalne, to nie stanie się optymalnym wówczas, gdy do zbioru dostępnych działań dołączone zostanie dowolne działanie nieoptymalne. Łatwo zilustrować konkretnym problemem decyzji intuicyjność tego postulatu. Łatwo też sformułować szereg innych, równie intuicyjnych, warunków koniecznych racjonalności kryterium decyzji.

Opis powyższy jest nader ogólnikowy. Precyzyjniejsze ujęcie sprawy byłoby równoważne wykładowi odpowiedniego fragmentu teorii decyzji, który znaleźć można w innych źródłach¹. Jasne powinno być w każdym razie to, iż omawiana teoria dotyczy takiej sytuacji, w której możliwe konsekwencje alternatywnych działań są znane i ocenione, nie wiadomo natomiast, która z nich zajdzie. Teoria stara się odpowiedzieć na pytanie: co to znaczy „zdecydować racjonalnie”? Okazuje się wprawdzie, że nie istnieje jednoznaczna i zadowalająca wszystkie nasze intuicje odpowiedź na powyższe pytanie, ale dzięki dokonanej analizie wiemy przynajmniej, czego możemy od pojęcia racjonalności niesprzecznie wymagać, a także potrafimy w świetle tych wymagań oceniać rozmaite koncepcje racjonalności.

Jakie znaczenie mają naszkicowane wyżej rezultaty dla kogoś, kto interesuje się zagadnieniami etycznymi? Sądzę, że następujące.

¹ Zob. np. K. Szaniawski, *Some Remarks Concerning the Criterion of Rational Decision Making*, „*Studia Logica*” 1960, IX.

Znany jest fakt, iż trudność, jaką sprawia wybór najlepszego w sensie moralnym działania, pochodzi co najmniej z dwóch źródeł: jednym z nich jest ocena skutków działania, drugim — niepewność tych skutków. Byłoby rzeczą pożądaną — z teoretycznego przynajmniej punktu widzenia — wyodrębnienie tych dwóch typów trudności w analizie wszelkiej decyzji. W tym zakresie wyniki uzyskane przez teorię decyzji są wyraźnym krokiem naprzód — mimo zastrzeżeń, o których będzie dalej mowa. Ów krok naprzód polega na wskazaniu, w jaki sposób postępować w warunkach niepewności lub ryzyka co do skutków działań — przy założeniu, że problem wartościowania tych skutków został pomyślnie rozstrzygnięty. Otóż temu założeniu trzeba obecnie poświęcić nieco uwagi, ponieważ ono właśnie wydaje się szczególnie wątpliwe w odniesieniu do wartościowań moralnych, skutkiem czego teoria decyzji w obecnej swej postaci nie posiada dla etyka takiej doniosłości, jaką by posiadać mogła.

Najpierw parę słów wyjaśnienia co do sposobu reprezentowania wartościowań przez liczby. Po pierwsze: przyporządkowanie liczb wartościowanym przedmiotom (w abstrakcyjnym sensie tego słowa) jest w znacznych granicach dowolne. Pod względem formalnym liczby te mają taki status, jak miary temperatury określanej za pomocą zwykłego termometru (a nie na skali absolutnej). Dowolny jest zatem wybór jednostki i początku układu. Wynika stąd, że nie wszelkie relacje między tymi liczbami reprezentują związki między odpowiednimi wartościowaniami. Jeśli np. przedmiotowi *A* przyporządkowana jest na czyjejś skali wartościowań liczba 3, a przedmiotowi *B* — liczba 6, to fakt, że $6 = 2 \times 3$, nie upoważnia nas do twierdzenia, iż przedmiot *B* jest przez ową osobę dwa razy wyżej ceniony niż przedmiot *A*. (Z tych samych przyczyn, dla których nie ma sensu powiedzenie, że wczoraj była dwa razy wyższa temperatura niż dzisiaj, chociaż odpowiednie wskazania termometru w stopniach Celsjusza wynosiły 6 i 3.)

Po drugie: sens reprezentowania wartościowań przez liczby jest behawiorystyczny, odwołuje się mianowicie do decyzji danego osobnika w szczególnego typu sytuacjach. Trudno tu przedstawiać tę kwestię w całej rozciągłości, ograniczę się więc do przykładu. Jeśli przedmiotom *A*, *B*, *C*, przyporządkowane są na czyjejś skali wartościowań odpowiednio liczby 1, 2, 4, to ma to mniej więcej taki sens: osobnik ów jednakowo ceni otrzymanie przedmiotu *B*, jak udział w loterii przyznającej mu *A* z prawdopodobieństwem $\frac{2}{3}$, zaś *C* z prawdopodobieństwem $\frac{1}{3}$. Wzmiankowaną tu równocześnie można stwierdzać empirycznie przez stawianie danego osobnika w odpowiednio skonstruowanych sytuacjach i obserwowanie dokonanego przezeń wyboru.

Zwracano niejednokrotnie uwagę na fakt, iż tak pojęty pomiar intensywności preferencji trudny jest do przeprowadzenia nawet w wa-

runkach laboratoryjnych, a więc w odniesieniu do sytuacji szczególnie prostych. Tym większe zapewne byłyby trudności w przypadku wartościowań etycznych, które — jak wiadomo — cechują się raczej złożonością niż prostotą. Wydaje się, iż w najlepszym razie wartościowaniom etycznym można by było przypisywać reprezentację liczbową na podstawie autorefleksji osobnika wartościującego, który starałby się rzetelnie określić własną decyzję w odpowiednio dobranych sytuacjach fikcyjnych. Pomiar doświadczalny w ścisłym tego słowa znaczeniu nie dałby się chyba przeprowadzić. Zatem i literalna interpretacja modelu decyzyjnego w kategoriach etycznych nie byłaby możliwa. Nie likwiduje to oczywiście przybliżonej stosowalności tego modelu, który zresztą i w innych dziedzinach badań nad postępowaniem ludzkim pełni taką właśnie rolę, funkcjonuje mianowicie jako schemat, którego parametrom nadaje się wartości liczbowe w trybie oszacowania dalekiego od dokładności.

Najpoważniejszą jednak przeszkodę w bezpośrednim stosowaniu rozważanego modelu do decyzji etycznej stanowi, jak się zdaje, ten fakt, iż zakłada się tu rozwiązanie problemu, który należy w etyce do najistotniejszych i najtrudniejszych. Zakłada się mianowicie, że konsekwencje czynów są pod względem swej wartości etycznej porównywalne jednowymiarowo (tzn. dają się rozmieścić na jednowymiarowej skali wartości) czyli — mówiąc analitycznie — ich oceny są reprezentowalne (przynajmniej w zasadzie) przez pojedyncze liczby. Otóż to właśnie wydaje się szczególnie wątpliwe.

Wolno mianowicie sądzić, iż trudność decyzji etycznej pochodzi m.in. stąd, że porównawcza ocena konsekwencji czynów różnie na ogół wypada w świetle różnych norm postępowania, akceptowanych przez decydenta. Normy te wyznaczają więcej niż jedną skalę wartościowań i jest otwartym problemem, czy można je zredukować do jednego wymiaru. Taką właśnie sytuację ma się na myśli mówiąc o konflikcie wartości. W języku analitycznym sytuacja daje się opisać przez przyporządkowanie każdej konsekwencji dwóch (lub więcej) liczb reprezentujących różne wartościowania. W ten sposób założenia modelu ulegają istotnej zmianie i proponowane rozwiązania nie dają się już zastosować. Powstaje pytanie, czy dałoby się coś powiedzieć na temat racjonalnego podejmowania decyzji przy tak zmienionych założeniach, tzn. przy zastąpieniu wartościowania jednowymiarowego przez wielowymiarowe.

Zauważmy przede wszystkim, że jeśli nie istnieje niepewność co do skutków działań, to problem redukcji wartościowań wielowymiarowych do jednego wymiaru daje się wyraźnie sformułować i dopuszcza możliwość takich rozwiązań, które nie wymagałyby utopijnego poszukiwania wspólnego miernika dla wartości różnego typu. Zwrócił na to uwagę

Milnor w swym znanym artykule o *Grach z przyrodą*². Można mianowicie wspomnianą wyżej macierz $\|u_{ij}\|$ poddać nowej interpretacji, wedle której kolumny reprezentowałyby wartościowania — pod różnymi względami — skutków działań, jednoznacznie tym działaniom przyporządkowanych. Problem (Milnor nazywa go „problem of amalgamation”, mając na myśli amalgamację wielu wartościowań w jedną ocenę sumaryczną) polegałby nadal na wyróżnieniu działań optymalnych i nie jest wykluczone, iż nadawałoby się do tego celu któreś ze znanych kryteriów decyzji, np. maximin. Trudno jednak tę kwestię rozstrzygnąć, dopóki nie sformułuje się jakichś wymagań w stosunku do takiej procedury; nie wydaje się uzasadnione automatyczne przeniesienie na ten teren postulatów charakteryzujących racjonalną decyzję w warunkach niepewności. Jest to, o ile mi wiadomo, zagadnienie, które nie zostało dotychczas podjęte, odnotowuję je zatem jako zasługujące na uwagę.

Warto może przy okazji zauważyć, że gdyby powyższy problem (kryterium decyzji w warunkach pewności, ze względu na wartościowanie wielowymiarowe) został zadowolająco rozstrzygnięty, to rozwiązanie to dałoby się przenieść również na decyzję w warunkach ryzyka. Mając bowiem — w przypadku np. dwuwymiarowości — macierz podwójnych ocen: $\|u_{ij}, v_{ij}\|$ oraz prawdopodobieństwa p_j przyporządkowane jej kolumnom można by było zredukować te dane do dwukolumnowej macierzy przeciętnych wartości: $\|\sum_{j=1}^n u_{ij}p_j, \sum_{j=1}^n v_{ij}p_j\|$. Z kolei do tej macierzy można byłoby zastosować przyjęte kryterium amalgamacji ocen różnych typów.

Natomiast problem decyzji w warunkach niepewności, ze względu na wartościowanie wielowymiarowe, nie daje się przy obecnym stanie wiedzy rozwiązać i nie wiadomo, czy jest w ogóle rozstrzygalny.

Na zakończenie tych uwag chciałbym dać wyraz przekonaniu, że nawet przy daleko posuniętym sceptycyzmie co do praktycznej stosowalności wyników teorii decyzji do rozstrzygania problemów etycznych nie sposób odmówić tej teorii pewnej wartości kształcącej. Wydaje się mianowicie, że dzięki zapoznaniu się z tą teorią lepiej sobie uprzytamniamy, na czym polega akt decyzji i które jego elementy wymagają dokładniejszego zbadania.

Zbiorowe podejmowanie decyzji

Prace na ten temat, inspirowane przez praktykę głosowań, mają tradycję sięgającą co najmniej czasów Condorceta. Główny wynik uzyskany

² J. W. Milnor, *Games against Nature*, w zbiorze *Decision Processes*, New York 1954.

został przez Arrowa, któremu zawdzięczamy sformułowanie problemu w języku umożliwiającym ściśle wnioskowania.

Chodzi, mówiąc ogólnie, o przyporządkowanie preferencjom indywidualnym — preferencji społecznej.

Użycie słowa „preferencja” zamiast słowa „wartościowanie” nie jest przypadkowe. O intensywnościach preferencji nic się w tym modelu nie mówi, zakładając jedynie uporządkowanie preferencyjne ocenianych przedmiotów. Takie osłabienie pojęcia wartościowania ma swój powód; ten mianowicie, iż wykorzystanie intensywności preferencji mogłoby tu polegać jedynie na międzyosobowym ich porównywaniu, czyli — na czymś w rodzaju Benthamowskiego rachunku przyjemności i przykrości. A to jest procedura nasuwająca poważne wątpliwości, etycznej właśnie natury. Nie chcą jej więc stosować twórcy rozważanej teorii. Sądzę, że postępują słusznie.

Zagadnienie przedstawia się w najogólniejszym ujęciu tak. Dany jest zbiór osób, który nazwiemy społeczeństwem, oraz zbiór przedmiotów (sposobów postępowania, kandydatów na określone stanowisko, lub tp.). Wiadome jest, jak każdy członek społeczeństwa porządkuje preferencyjnie owe przedmioty; łącznie nazwiemy te oceny profilem uporządkowań preferencyjnych (dla danego społeczeństwa). A oto problem: jak ma wyglądać uporządkowanie preferencyjne reprezentujące ocenę dokonaną przez społeczeństwo jako całość?

Formalnie rzecz biorąc, poszukuje się „dobrej” reguły (zwanej czasem funkcją dobrobytu społecznego), która z każdym możliwym profilem wiązałaby dokładnie jedno uporządkowanie. Trzeba wszakże ustalić najpierw sens słowa „dobry” w odniesieniu do takiej reguły.

Dokonuje się tego przez postulaty. Nie będę ich tu wyliczał, ponieważ łatwo dostępne źródła³ przedstawiają tę kwestię dokładnie. Zaznaczę tylko, że niektóre z tych postulatów mają treść etyczną, zresztą nieskomplikowaną i przekonywającą. Jeden z nich np. domaga się, aby reguła zbiorowej decyzji nie była dyktatorska, tzn. aby nie istniał taki osobnik, którego wartościowanie jest rozstrzygające społecznie — bez względu na oceny innych członków zbiorowości. Inny postulat żąda, by reguła nie była narzucona, tzn. — aby liczyła się w pewnym sensie z preferencjami jednostek.

Znane — i paradoksalne — twierdzenie Arrowa pokazuje, iż postulaty te nie zawsze (tzn. nie przy każdym profilu uporządkowań preferencyjnych) dadzą się ze sobą pogodzić.

³ K. J. Arrow, *Social Choice and Individual Values*, New York 1951; R. D. Luce, H. Raiffa, *Gry i decyzje*, Warszawa 1964, p. 304-337.

Wobec tego faktu trzeba albo zmodyfikować postulaty, albo tak ograniczyć zmienność preferencji jednostkowych, by sprzeczność postulatów była wykluczona. Obie te drogi zostały z dobrym skutkiem wypróbowane.

Teoria zbiorowego podejmowania decyzji dostarcza najwcześniejszego bodaj przykładu postępowania, które na tym polega, że w odniesieniu do prostego (krytycy powiedzą: zbyt prostego) modelu behawioralnego formułuje się w języku matematyki postulaty o treści etycznej. Z tego względu wydaje się ona pouczająca. Praktyczne znaczenie wykazanej przez Arrowa sprzeczności jest zapewne niewielkie, ponieważ w społeczeństwach rzeczywistych tak znaczna rozbieżność preferencji indywidualnych należy do rzadkości. Jednakże fakt, iż bardzo intuicyjne wymogi etyczne mogą okazać się sprzeczne, nie jest pozbawiony teoretycznego znaczenia⁴. Należy podkreślić, iż rezultat ten został uzyskany właśnie dzięki zastosowaniu języka matematycznego, który umożliwia ścisłą dedukcję.

Wydaje się, że problematyka decydowania zbiorowego daleka jest od wyczerpania. Istnieje mnogość procedur (zwłaszcza wyborczych), które są w praktycznym użyciu. Zbadanie ich formalnych własności — w świetle wymogów, jakie można by wobec nich sformułować — to temat zasługujący na podjęcie.

Sprawiedliwy podział

W tej dziedzinie zajmowano się kwestiami różnymi. Pewna liczba prac poświęcona jest tzw. grom sprawiedliwego podziału⁵. Chodzi tutaj o skonstruowanie takiej n -osobowej gry strategicznej, która niejako „automatycznie” realizowałaby równy — i w tym sensie sprawiedliwy — podział jakiegoś dobra o charakterze ciągłym. Automatyzm ów polegałby na tym, iż uczestnicy tej gry, działając w dobrze pojętym interesie własnym, musieliby wybierać takie strategie, które łącznie dają wynik sprawiedliwy.

Wartość (znów: niemal wyłącznie teoretyczna) uzyskanych wyników polega zapewne na laboratoryjnym pokazaniu, iż narzucając ludziom odpowiednio dobrane reguły gry oraz zakładając motywację rozumnie egoistyczną można ich skłonić do zachowań dających pożądany etycznie rezultat. Samo jednak pojęcie „sprawiedliwego podziału” wydaje się — wskutek swego prymitywizmu — mało interesujące dla etyka.

⁴ Suppes w artykule *The Philosophical Relevance of Decision Theory* ("The Journal of Philosophy" 1961, s. 612) przypisuje wynikom Milnora i Arrowa taką rolę, jaką swego czasu odegrało na terenie logiki odkrycie antynomii teorio-mno-gościowych.

⁵ R. D. Luce i H. Raiffa, op. cit., s. 337-340.

Nieco ciekawszy pod tym względem jest taki przypadek, gdy podziałowi podlega skończony zbiór przedmiotów (pozytywnie ocenianych, a więc zasługujących na miano dóbr)⁶. Zakłada się tu, iż znane są — dla każdego uczestnika podziału — uporządkowania preferencyjne owych przedmiotów. Wprowadzone zostaje pojęcie podziału, uogólnione probabilistycznie w stosunku do potocznego pojmowania tego słowa. Reguła dystrybucji przyporządkowuje każdemu profilowi uporządkowań preferencyjnych — dokładnie jeden podział (w uogólnionym sensie tego terminu). Następnie formułuje się pod adresem takiej reguły postulaty o prostej treści etycznej.

Głównym wynikiem jest twierdzenie, że dwa takie postulaty (postulat równości szans oraz optymalności w sensie Pareta) nie dla każdego profilu uporządkowań preferencyjnych dają się zarazem spełnić.

Tak określony problem sprawiedliwego podziału sformułowano stosunkowo niedawno. Wydaje się, iż jest to teren etycznie obiecujący, trudno jednak powiedzieć w tej sprawie coś stanowczego, dopóki badania nie zostaną dalej posunięte.

Sprawiedliwy arbitraż

Ta problematyka wyrosła na gruncie teorii gier o sumie niezerowej, która to teoria bada częściowy konflikt interesów. „Częściowy”, a więc taki, który dopuszcza rozwiązania preferowane przez obu graczy⁷ w stosunku do pewnych innych rozwiązań. Jasną jest rzeczą, że w takiej sytuacji opłacalne bywa porozumienie — chociażby dla uniknięcia wyniku zgodnie uznanego za gorszy. Stąd możliwość akceptowanego przez obie strony arbitrażu, który w przypadku konfliktu zupełnego (gra o sumie zerowej) byłby albo niepotrzebny, albo nie do przyjęcia.

Powstaje problem: jaka zasada arbitrażu zasługuje na miano sprawiedliwej? Również i w tym wypadku formalne ujęcie zagadnienia pozwala ściśle określić wymogi pod adresem sprawiedliwej reguły arbitrażu, czyli aksjomatycznie ukonstytuować sens tego pojęcia⁸. Wedle tych wymogów sprawiedliwa zasada arbitrażu powinna 1) przyznawać każdemu graczowi co najmniej tyle, ile mógłby uzyskać nie pytając partnera o zgodę, 2) być optymalna w sensie Pareta, 3) w sytuacjach symetrycznych przyznawać obu graczom tyle samo, 4) przy rozszerzeniu zbioru dopuszczalnych wyników zmieniać rozstrzygnięcie tylko wtedy, gdy nowy wynik

⁶ K. Szaniawski, *O pojęciu podziału dóbr*, „Studia Filozoficzne” 1966, 2.

⁷ Łatwo powyższe określenie uogólnić na grę n-osobową, dla prostoty jednak będzie tu mówić tylko o grach dwuosobowych.

⁸ R. D. Luce i H. Raiffa, *op. cit.*, rozdz. 6. Także R. B. Braithwaite *Theory of Games as a Tool for the Moral Philosopher*, London 1955.

jest lepszy od dotychczasowych. Można wykazać, że istnieje dokładnie jedna zasada arbitrażu spełniająca wszystkie te postulaty.

Zauważmy, że jeśli takie pojmowanie sprawiedliwego arbitrażu nie zaspokaja czyichś intuicji etycznych, to oczywiście może on zastąpić powyższe postulaty innymi. Wypada też zaznaczyć, że rozważany model zakłada, iż obaj gracze po swojemu wartościują możliwe wyniki gry. Wartościowanie jest tu rozumiane tak, jak w teorii indywidualnego podejmowania decyzji. Ale proponowane rozwiązanie nie polega na kompensowaniu straty jednego gracza zyskiem drugiego, czyli nie odwołuje się do międzyosobowych porównywań wartości.

Byłoby rzeczą ciekawą porównanie postulatów charakteryzujących sprawiedliwy arbitraż z postulatami dotyczącymi sprawiedliwego podziału. Jedno i drugie pojęcie wydaje się pewnym uszczegółowieniem sprawiedliwości dystrybtywnej, nie skonstruowano jednak dotychczas modelu wystarczająco ogólnego, by objął oba te przypadki. Na pierwszy rzut oka różnica polega na tym, iż arbitraż musi brać pod uwagę różnice w strategicznej sytuacji graczy — w tym mianowicie sensie, że przyznaje więcej graczowi będącemu w korzystniejszym położeniu. Z etycznego punktu widzenia nie zasługuje to chyba na aprobatę, ponieważ jest niejako respektowaniem „prawa silniejszego”. Uzasadnieniem takiego rozstrzygnięcia jest jego realizm: na inny sposób rozstrzygnięcia gracz „silniejszy” mógłby się po prostu nie zgodzić. Jest jednak właśnie rzeczą dyskusyjną, czy realizm (w sensie szansy zaakceptowania przez obu graczy) jest tym, czego trzeba koniecznie wymagać od sprawiedliwej zasady arbitrażu.

Istnieją również inne rozwiązania problemu arbitrażu. Różnią się one od przedstawionego wyżej tym, że ustalają jakąś wspólną dla obu graczy jednostkę wartości. Najciekawsza wydaje mi się propozycja Braithwaite'a, niestety zbyt skomplikowana, by ją w tym krótkim przeglądzie referować.

Również i ta problematyka, choć ma już dość bogatą literaturę, pozostawia sporo otwartych problemów. Badania dotychczasowe koncentrowały się raczej na strategicznych aspektach gier o sumie niezerowej. Przyjęcie wyraźnie etycznego punktu widzenia mogłoby przynieść nowe rezultaty.

Przy okazji zaznaczę, że wiele kwestii natury pojęciowej nie zostało dotychczas podjętych. Podział gier na kompletnie i częściowo tylko antagonistyczne wydaje się zbyt gruby: pożądane byłoby mieć jakiś miernik, który pozwalałby mówić o sytuacjach bardziej i mniej konfliktowych. Pojęcie „siły strategicznej” oraz spokrewnione z nim pojęcie szantażu nadal pozostają w sferze niesprecyzowanych intuicji. A nie są to pojęcia obojętne z punktu widzenia etyka.

Podsumowanie

Omówiłem pokrótce te działy dyscypliny zwanej teorią decyzji, których znajomość może się na coś przydać etykowi jako takiemu. Niektóre badania (nad zbiorowym podejmowaniem decyzji lub pojęciem sprawiedliwego podziału) podjęte zostały z inspiracji etycznej i mogą być nazywane — z niewielką dozą przesady — uprawianiem etyki środkami matematycznymi. Inne okazują się pożyteczne pośrednio — dzięki precyzacji takich pojęć jak „racjonalność”, które to pojęcia nie należą wprawdzie do swoiście etycznych, odgrywają jednak w refleksji etycznej dość ważną rolę.

Nie jest wykluczone, że wyniki, które tu szkicowo przedstawiłem, wydadzą się rozczarowujące wskutek odniesienia do sytuacji sztucznie schematyzowanych, raczej dalekich od tego stopnia złożoności, jaki cechuje badane przez etyka realne sytuacje życiowe. Sądzę, że nie jest to powód, aby je lekką ręką odrzucać.

Po pierwsze: za cenę prostoty zyskuje się tu możliwość udowodnienia twierdzeń, które nie są na pierwszy rzut oka oczywiste. Po drugie: lepsza aproksymacja rzeczywistości jest — być może — do uzyskania; badania tego typu prowadzi się przecież od niedawna. Po trzecie: wolno sądzić, że zaznajomienie się z aparaturą pojęciową teorii decyzji jest pouczające dla badaczy ludzkich zachowań, niezależnie nawet od możliwości praktycznego zastosowania tych pojęć. A etycy badają m.in. właśnie postępowanie ludzkie.

Клеменс Шанявски

TEORIA ПРИНЯТИЯ РЕШЕНИЯ И ЭТИКА

Целью настоящей статьи является представление и обсуждение возможности применения теории решения для выяснения этических концепций.

1. Индивидуальные решения. Результаты касающиеся концепции правильности могут, в принципе, применяться к специальным случаям когда последствия действий оцениваются с этической точки зрения. Главное затруднение это неизмеримость оценок постулированных теорией решения. Этические оценки как правило многозначны в отражении противоречия оценок.

2. Коллективные решения. Хорошо известный результат, полученный Апро освещает в некоторой степени идею справедливого общественного решения. Многие критерии, применяемые актуально должны быть исследованы по этой линии.

3. Понятие справедливого раздела. Это важно для выяснения смысла справедливости. Последний результат показывает противоречие между постулатом равных шансов и постулатом оптимальности (в его самой слабой форме).

4. Схемы арбитража. Здесь вопрос в том, чтобы найти этически удовлетворяющие критерии для решения (частично) конфликтов. Возможны этические возражения к решениям предлагаемым до сих пор; они освещают стратегические аспекты конфликта, поддерживая, в некотором смысле, более сильного оппонента.

Klemens Szaniawski

DECISION THEORY AND ETHICS

The purpose of this paper is to present and discuss possible uses of decision theory in clarifying ethical concepts.

1. Individual decision making. The results concerning the concept of rationality could, in principle, be applied to the special case when consequences of available actions are appraised from the ethical viewpoint. The main obstacle: unidimensionality of valuations, postulated by decision theory. Ethical valuations are typically many-dimensional, reflecting conflicts of values.

2. Collective decision making. The well-known result by Arrow throws some light on the concept of „fair” social decision. Many criteria that are actually used remain to be investigated along these lines.

3. The notion of fair division. This is important for elucidating the meaning of justice. A recent result shows the incompatibility of the postulate of equal chances with the postulate of optimality (in its weakest form).

4. Arbitration schemes. Here the problem is to find an ethically satisfactory criterion of resolving (partial) conflicts. A possible ethical objection to the solutions proposed so far: they reflect the strategic aspects of conflict, favorizing — in a sense — the stronger opponent.

Dyskusja nad referatem doc. dr K. Szaniawskiego

Dyskusja miała w zasadzie charakter dialogu, w którym prelegent wyjaśniał niektóre zawiłości teorii decyzji oraz wzajemne związki pewnych pojęć etyki i teorii decyzji. Jedynie prof. Fritzhand podjął próbę ogólnej oceny przydatności teorii decyzji dla etyki, podkreślając przy tym wartość wysiłków wprowadzenia do etyki większej precyzji i racjonalności. Mówca widzi w teorii decyzji substytut skompromitowanego „rachunku” Benthama, podkreślając, iż w tym kierunku zmierza także D. Braybrook w książce *Strategy of Decision*.

W dyskusji zabierali również głos: dr H. Jankowski, doc. dr A. Podgórecki i mgr Z. Szawarski.