

## ***Considerazioni sul consequenzialismo***

Pierpaolo Marrone

*Università di Trieste*

*Dipartimento di Studi Umanistici*

*marrone@units.it*

### **ABSTRACT**

In this paper I explore the idea of expected utility as a description of how we actually act and how we reason. This theory has both normative side and descriptive side. I argue that as idealized normativity it does not recognize the role of moral principle. It is possible that a coherentist strategy can have better chances to balance both our expectations on consequences and our preferences for states.

### **KEYWORDS**

Utilitarianism, expected utility, consequentialism.

1. L'utilità attesa è una buona descrizione di come effettivamente agiamo e di come ragioniamo, e una teoria della ragione nell'azione deve essere, dal punto di vista normativo, una teoria strumentale? Per rispondere a questa domanda ritengo sia opportuno cercare di precisare rapidamente quali forme abbia assunto, anche a livello di un determinato senso comune, una concezione strumentale della ragione nell'azione. Si tratta di una visione che tiene più cose assieme, poiché si è configurata sia come teoria del comportamento sia come teoria delle procedure di pensiero corrette in vista dell'azione<sup>1</sup>. In realtà, non è del tutto chiaro che una teoria del comportamento possa anche essere una valida descrizione dei processi di pensiero. Il motivo, a mio modo di vedere, è presto detto: l'idea di perseguire l'utilità attesa ha, come teoria dei processi di pensiero, una dimensione marcatamente normativa più che descrittiva, nel senso che prescrive quale insieme di preferenze, desideri, atteggiamenti intenzionali, sarebbe razionale

---

<sup>1</sup> R. Jeffrey, *On Interpersonal Utility Theory*, in "Journal of Philosophy", 1971, pp. 647-656; J. Broom, *Weighing Goods*, Oxford, Blackwell, 1991,

avere per soddisfare i requisiti della teoria. È dubbio che questi insiemi possano essere determinati a priori con un ragionamento che non sia circolare e tautologico, senza, cioè, la presenza di qualche altro insieme di assunzioni. La mia convinzione è che debba essere assunto un vasto insieme di conoscenze che l'agente deve possedere piuttosto stabilmente, nell'ambito di quella che, imprecisamente, possiamo chiamare personalità matura. Tale visione, in quanto si occupa dei processi di pensiero, ha come ambito primo di applicazione le strategie dell'azione individuali. Viceversa, come teoria del comportamento ha un ambito di applicazione ben più vasto. Può riguardare il comportamento di istituzioni politiche, di associazioni sindacali, di aggregazioni occasionali di individui, ma può essere anche utilizzata per descrivere il comportamento animale e la riproduzione sessuale<sup>2</sup>. Entrambe queste caratterizzazioni dell'azione sono normative in due sensi sottilmente diversi. Quello che le distingue è che la prima è normativa *ex ante*, ossia si impegna a prescrivere quello che è descrivibile come comportamento razionale attraverso la valutazione di alternative ordinate cardinalmente; nel secondo caso, invece, la normatività spesso è, per così dire, *ex post*, nel senso che il comportamento che riesce ad avere successo, farsi strada, e ad avere rilevanza per la maggioranza di un gruppo o per i suoi membri più influenti e in grado di dirigere gli altri, dimostra così la sua razionalità e non è invece assunto sia frutto di quella che potrebbe risultare una semplice casualità. Questa ultima spiegazione vale anche, ad esempio, per le interpretazioni in termini di *fitness* nelle teorie evolutive. Oltre ad essere normative queste due caratterizzazioni dell'azione sono anche descrittive e predittive? Non è indispensabile, in realtà, che sia così perché non vi è un legame analiticamente necessario fra il successo dei piani di azione di un agente e la presenza nella sua mente di "processi mentali corretti" per raggiungere il risultato che garantisce la massimizzazione dell'utilità attesa<sup>3</sup>. In particolare, ci sono forti evidenze contrarie che mostrano come gli individui non siano in grado di valutare correttamente la probabilità di eventi indispensabili per predisporre la

---

<sup>2</sup> R. Baker, *Guerre sessuali* (1996), Baldini & Castoldi, Milano, 1997, offre un interessante esempio di riduzionismo nell'ambito della biologia evuzionista per quanto riguarda le strategie riproduttive fra gli esseri umani.

<sup>3</sup> Si veda D. Satz & J. Ferejohn, *Rational Choice and Social Theory*, in "Journal of Philosophy", 1994, pp. 71-87.

strategia d'azione migliore<sup>4</sup>. Il successo della predizione non implica il successo della descrizione dei processi mentali dell'agente.

2. Rimane, tuttavia, vero che, quando si agisce, una determinata strumentalità è sempre presente, anche se non siamo disposti a ridurre tutte le relazioni fra uomini a relazioni di scambio economico<sup>5</sup>. Questa strumentalità può essere descritta come la presenza nell'agente di un insieme di credenze. Fra queste credenze si ritiene che almeno tre siano particolarmente importanti: 1) l'agente ritiene che un'azione sia razionale se consente di raggiungere un fine che si è proposto; 2) la ragione determina i mezzi adeguati al raggiungimento del fine; 3) la determinazione dei fini riguarda qualcos'altro dalla ragione, la quale opera come strumento di determinazione dei mezzi. Come è noto, alcuni si spingono ben oltre e pensano che la ragione non può *mai* determinare i fini, ossia che la ragione è *sempre* strumentale. In questa prospettiva, non ha quindi senso parlare di fine razionale, se non in quanto sia strumentalmente scelto, così come non ha molto senso parlare di ragion pratica. Seguendo alcuni autori<sup>6</sup>, si può, inoltre, introdurre un'utile tassonomia che distingue fra fini di primo ordine (ciò che motiva l'azione) e fini di secondo ordine (a loro volta mezzi per il raggiungimento dei fini di primo ordine). Se ho una specifica attrazione per le donne bionde non riempirò la mia casa di poster di Fernanda Lessa; se amo i cani di piccola taglia, e non possiedo animali domestici, non mi prenderò in casa un'iguana gigante; se desidero fare un'adozione a distanza e possiedo risorse limitate, non accenderò un mutuo per acquistare una Porsche (a meno che non sia vittima di compulsioni di un qualche genere<sup>7</sup>). Si noti che l'azione di un agente può conservare tutte le caratteristiche della razionalità strumentale,

---

<sup>4</sup> Recentemente è stato proposto che ciò dipenda in larga misura dal mancato uso di frequenze naturali nel calcolo degli agenti. Cfr. G. Gigerenzer, *Quando i numeri ingannano* (2002), Cortina, Milano, 2003.

<sup>5</sup> Sui presupposti storico-antropologici di questa operazione di riduzione, si veda l'indagine di Louis Dumont, *Saggi sull'individualismo* (1983), Adelphi, Milano, 1993.

<sup>6</sup> D. Dennett, *Brainstorms* (1978), Adelphi, Milano, 1991, pp. 404-429, che discute a sua volta H. Frankfurt, *Freedom of the Will and the Concept of a Person*, in "The Journal of Philosophy", 1971, pp. 5-20. Si tratta di una discussione che riguarda la filosofia della mente e l'etica, ma ritengo che alcuni suoi risultati, quali la tassonomia citata, possano essere utili nel quadro della presente discussione.

<sup>7</sup> Sarebbero, in questo caso, esemplificazioni che rientrano nel problema dell'*akrasia*, per le quali cfr. le ricerche di A. Mele, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford University Press, Oxford, 1992.

anche se l'esito dell'azione non rappresenta affatto il risultato atteso, magari perché sono intervenute circostanze avverse o perché, come spesso accade, l'informazione era incompleta. Ossia: la rilevanza del processo va tutta sul lato della procedura decisionale dell'azione e non sul suo inizio – il fine desiderato – o sulla sua conclusione – l'esito –. La credenza dell'agente di poter raggiungere un certo fine attraverso un determinato insieme di mezzi, non obbliga affatto, in realtà, ad assumere l'intero insieme delle premesse del ragionamento strumentale. Di fatto, esistono versioni della razionalità nell'azione che accettano le prime due premesse, ma negano la terza. Un esempio ne sono varie forme di deontologismo kantiano<sup>8</sup> oppure la teoria delle capacità della Nussbaum<sup>9</sup>. In queste teorie, il fine non è scelto dalla ragione strumentale e, nello stesso tempo, si nega che la scelta del fine non abbia alcun fondamento in altre forme di razionalità. Nelle versioni kantiane, questo è ben visibile nella nozione degli individui come fini in sé, ossia come enti capaci di prospettarsi la nozione pura di dovere come guida dell'azione, per cui per adempiere alla legge morale è necessario trattare l'umanità in se stessi, così come negli altri, sempre come fine *e mai soltanto* come mezzo. Nella versione della Nussbaum, questo è illustrato dalla naturalizzazione della teoria dei diritti attraverso la teoria delle capacità. In queste versioni, la ragione strumentale è una delle manifestazioni, e non la più importante, di una più generale razionalità. Ad essa, ad ogni modo, non spetta la determinazione dei fini, non perché questi siano extrarazionali, ma perché la loro scelta è di pertinenza di una forma di razionalità specifica.

Questa forma di razionalità – indirizzata verso il compimento morale del regno dei fini o verso la vita buona – discrimina fra fini che sono giudicati razionali e fini che sono ritenuti irrazionali (a seconda che corrispondano a un qualche test di universalizzabilità o alla promozione delle *capabilities*). Queste versioni individuano, quindi, una razionalità più complessa e più profondamente impegnata nell'azione, perché il suo raggio d'azione è più vasto, comprendendo la scelta dei fini e la determinazione dei mezzi adeguati alla natura di questi fini. Questa versione della razionalità non è uno strumento immediatamente predittivo, ma poiché possiede una dimensione sia descrittiva sia normativa sia ascrivibile, può essere predittiva in maniera

---

<sup>8</sup> O. O'Neil, *Bounds of Justice*, Cambridge University Press, Cambridge, 2000.

<sup>9</sup> M. Nussbaum, *Diventare persone. Donne e universalità dei diritti*, (2000), Il Mulino, Bologna, 2001, e E. Greblo, *Dai diritti alle capacità. L'universalismo contestuale di Martha Nussbaum*, in "Filosofia politica", 2002, pp. 251-273.

mediata. In questo senso, dopo che si è accertato il fine e in presenza di un'informazione empirica rilevante, si può anche, poniamo, sostenere che la scelta di determinati mezzi strumentali può essere erronea<sup>10</sup>. L'idea che l'utilità attesa sia la versione più coerente e più semplice di una ragione strumentale è, in definitiva, un'idea non del tutto chiara perché non sono chiare le dimensioni predittive, descrittive, normative dell'utilità. Come teoria del ragionamento può trovare spazio accanto o al di sotto di una ragione che sceglie anche i fini; come teoria generale del comportamento appare sia come una teoria descrittiva sia come una teoria normativa, e talvolta solo in campi ristretti.

3. In questo modo, tuttavia, la teoria dell'utilità attesa paga la sua elegante semplicità con una illusione di totalità, ignorando che le norme sono contestuali così come lo sono i fini, e che la stessa distinzione mezzi-fini è un prodotto largamente culturale, che varia in ragione dei diversi vocabolari che usiamo per rendere ragione della nostra presenza in un contesto per noi rilevante<sup>11</sup>. Le norme forniscono all'agente ragioni per agire, ma la stessa ragione strumentalmente considerata può divenire parte importante della motivazione a scegliere un determinato corso d'azione anziché un altro<sup>12</sup>. In effetti, può essere che le norme siano parte importante della determinazione a seguire uno specifico corso d'azione o che siano addirittura dirimenti per agire in un senso anziché in un altro, anche se sostenere che un agente ha delle ragioni per fare una determinata cosa è diverso dal dire che sia motivato a fare proprio quella cosa. Le norme sono di assoluta rilevanza nello strutturare il nostro comportamento e possono essere variamente vincolanti e formalizzate. Le norme per la lettura di un quotidiano, ad esempio, prescrivono che di solito il quotidiano vada letto dalla prima pagina e non dall'ultima. Si tratta di norme che è ragionevole pensare siano emerse in parte spontaneamente, in quanto generate anche come effetti secondari da altre norme – relative alla scrittura, alla possibilità offerte dai mezzi tecnici, e così via –. Le norme che regolano il fuorigioco nel gioco del calcio sono invece più strutturate e meno

---

<sup>10</sup> È ciò che motiva di solito la disobbedienza civile. Sulla razionalità strumentale della disobbedienza civile, cfr. H. Zinn, *Democrazia e disobbedienza* (1996), Il Saggiatore, Milano, 2003.

<sup>11</sup> G. Carlson, *Plans, Expectations, and Act-Utilitarian Distrust*, in "Philosophical Studies", 1979, pp. 295-300.

<sup>12</sup> Idea che mi pare stia alla base del suggestivo G. Alvi, *Il secolo americano* (1996), Adelphi, Milano, 1996.

soggette a variabili – anche se queste possono esserci e, difatti, rivestono un ruolo importante nella determinazione finale del risultato di una partita –. Entrambe queste classi di norme – gli esempi potrebbero essere moltiplicati a piacere, dalle norme relative alla circolazione stradale alle norme sul galateo alimentare – sono norme culturali, la cui forza motivazionale è spiegabile ricorrendo a spiegazioni storiche, sociologiche, psicologiche. Sono norme la cui convenzionalità stipulativa è esplicitamente assodata. Vi sono, poi, anche altri generi di norme che riguardano la convergenza su *punti focali* del comportamento, che si sono dimostrate efficaci proprio dal punto di vista dell'utilità attesa, ma l'aderenza alle quali non può semplicemente essere interpretata come il risultato di un ragionamento formale<sup>13</sup>. Esistono, inoltre, altre norme la cui culturalità non è invece così evidente, ed è, anzi, fortemente dubbia. La norma che proibisce *prima facie* l'omicidio sembra essere una di queste. La norma che proibisce di torturare bambini per divertimento pare essere un'altra norma di questo genere. Si tratta di norme che appaiono transculturali, nel senso che in ogni mondo possibile dove ci è permesso immaginare la presenza del genere umano queste norme dovrebbero valere. L'idea in questo caso è che la presenza di un genere naturale – il genere umano – svolge quella funzione autoritativa che è invece svolta dalla pressione sociale o psicologica o dall'inerzia storica nei contesti richiamati precedentemente<sup>14</sup>. In questo ultimo caso, le norme possono bensì essere violate o contemplare delle eccezioni – anche se non vi sono eccezioni, se non patologiche, alla norma che proibisce di torturare bambini per divertimento –, come ad esempio quando si consente l'omicidio per legittima difesa oppure in guerra, ma la loro forza autoritativa *prima facie* non è affatto diminuita dalle eccezioni. Anche in virtù di questa considerazione alcuni si proclamano oggettivisti in relazione all'esistenza di queste regole – che possiamo chiamare “valori” –, nel senso che essere parte del genere umano significa per lo meno possedere delle disposizioni mentali ad essere motivati “così e così” nelle circostanze appropriate. In questo senso, si tratta di norme che hanno una validità extraculturale. Non affronto il problema se queste norme esprimano

---

<sup>13</sup> T. Schelling, *The Strategy of Conflict*, Harvard University Press, Cambridge (Mass.), 1980<sup>2</sup>, sui punti focali e, inoltre, D. Lewis, *La convenzione* (1969), Bompiani, Milano, 1973 e M. Gilbert, *Rationality, Coordination, and Convention*, in “Synthese”, 1990, pp. 1-21.

<sup>14</sup> J.J. Thompson, *The Realm of Rights*, Harvard University Press, Cambridge (Mass.), 1990, utilizza questa idea per svolgere il suo attacco alla distinzione fatti-valori. In sintesi, il suo ragionamento è che se esistono dei giudizi morali necessari, allora la distinzione fatti-valori deve essere rifiutata. Per una critica si veda J. Bennett, *The Necessity of Moral Judgement*, in “Ethics”, 1993, 103, pp. 458-472.

proprietà del mondo naturale, e, se sì, del modo in cui tali proprietà possano essere accertate. Mi limito a notare che coloro che negano ci siano tali proprietà sulla base del fatto che noi non possediamo gli adeguati strumenti epistemologici per accertarne l'esistenza dovrebbero anche sostenere che la ragione strumentale, per gli stessi identici motivi, non può avere la forza autoritativa che alcuni sono disposti a concederle.

Io credo che esista almeno una possibilità di intendere la forza normativa della ragione strumentale senza bisogno di fare assunzioni particolarmente impegnative, nel senso che non pare essere particolarmente impegnativo sostenere che in determinate occasioni le conclusioni della ragione strumentale possano avere una persuasività notevole a motivare sulla base della nostra psicologia. È, ad esempio, importante tener conto che in certi processi decisionali una persona può trarre motivazioni ulteriori per agire dal fatto che già in passato si è comportata secondo un determinato schema motivazionale. Si tratta di motivazioni incrementali, nelle quali si afferma una connessione causale senza assumere alcuna precisa connessione giustificativa necessaria. Non c'è nulla di particolarmente impegnativo nel sostenere che, data la nostra costruzione psicologica, alcune forme di motivazioni siano per noi più attraenti di altre. In questo senso, la ragione strumentale è motivazionale poiché genera informazioni per decidere fra corsi d'azione alternativi. Questa posizione era, ad esempio, sostenuta da Hume assieme a un disposizionalismo morale che, coniugato al suo sentimentalismo, ne fa un fautore moderato di una teoria della ragione strumentale<sup>15</sup>. Questo genere di sentimentalismo pensa alla motivazione come a ciò che realizza delle disposizioni umane, non si compromette con una qualche forma di metafisica inverificabile, accetta la sfida della spiegazione più comprensiva e rigorosa – la sfida della spiegazione scientifica – naturalizzando l'etica. Hume è anche l'autore che ci spiega perché l'utilità piace<sup>16</sup> ed è giustamente considerato un protoutilitarista, ma per abbracciare o fondare una qualche forma di utilitarismo che si mostrasse adeguata a quelle che considerava caratteristiche dell'azione individuale e della socialità non aveva bisogno di alcuna teoria monistica dell'utilità attesa. Non nego che la sua concezione della ragione nell'azione sia scevra di problemi e nodi irrisolti. Ad esempio, nel momento in cui accorda un peso notevole alla ragione, la quale accerta ciò che è alla

---

<sup>15</sup> Ma si vedano le critiche di J. Broome, *Can a Humean be Moderate?*, in R. G. Frey & C. Morris (eds.), *Value, Welfare, Morality*, Cambridge University Press, Cambridge, 1993, pp. 51-73.

<sup>16</sup> D. Hume, *Ricerca sui principi della morale* (1751), Laterza, Roma-Bari, 1987, pp. 224-245.

portata dei nostri desideri, attraverso il ragionamento logico e attraverso l'accertamento causale, non dovrebbe rendere necessaria anche la naturalizzazione di questa stessa facoltà rendendola un meccanismo psico-sociale? Sia lo scetticismo sull'io sostanza sia anche le notazioni scettiche sul contratto originale paiono inclinare in questa direzione.

Se ci si può chiedere se il genere di obiettività resa possibile dal disposizionalismo non andrebbe perso perché sarebbe incluso in una circolarità priva di punti archimedei esterni, è necessario però affiancare a questa domanda quest'altra: dobbiamo necessariamente ridurre la normatività a elementi non normativi? Se il circolo al quale ho accennato è in qualche modo sussistente, sembrerebbe lo si debba negare, dal momento che negli asserti imperativi ciò che esprime la direzione normativa – una qualche forma verbale di “dovere” – deve essere ridotto a espressione psicologica con molte cautele e precauzioni<sup>17</sup>. Il successo della formalizzazione presso gli scienziati sociali<sup>18</sup> introduce due significativi problemi a una concezione dell'utilità che, naturalmente, sembrerebbe avere a che fare con desideri e stati dell'esperienza. Il primo è apparentemente un semplice spostamento semantico. Le utilità non si riferiscono tanto a desideri o a valori, quanto piuttosto alle preferenze dell'agente. Questo comporta un diverso approccio al problema del rischio. Tale approccio sottolinea che l'identificazione dell'utilità con valori monetari è completamente *ad hoc*, poiché esistono una infinità di funzioni che si incrementano in ragione decrescente, ed indubbiamente l'associazione con un valore monetario varia da persona a persona, anche se non sempre è del tutto chiaro come<sup>19</sup>. Il secondo riguarda il fondamento di questa decisione: perché una decisione dovrebbe essere basata sul valore atteso di queste utilità? Sembrerebbe che l'unico argomento che viene esibito per usare il valore atteso sia che questo si palesa sul lungo periodo, quando un gioco viene ripetuto più volte. È da sottolineare il merito di una interpretazione del valore atteso basata sulla frequenza, ma non è affatto chiaro perché dovrebbe applicarsi a un agente che partecipa a un gioco una sola volta. Esprimendosi con una certa approssimazione ciò che per alcuni von Neumann e Morgenstern hanno dimostrato è che se un agente è capace di

---

<sup>17</sup> A. Gibbard, *Wise Choices, Apt Feelings*, Harvard University Press, Cambridge (Mass.), 1991 offre, appunto, una compiuta sintesi di questo genere di espressivismo.

<sup>18</sup> Per una posizione anti-utilitaristica nelle scienze sociali, si veda A. Caillé, *Critica della ragione utilitaria* (1983), Torino, Boringhieri, 1991.

<sup>19</sup> D. Bernoulli, *Exposition of a New Theory of Measurement of Risk* (1738), in “Econometrica”, 1954, pp. 23-36.

esprimere le sue preferenze relativamente a ogni possibile coppia di giochi costruiti a partire da alcune semplici alternative ed è guidato unicamente dal valore della sua utilità attesa, allora egli agisce in accordo con le sue preferenze *reali*, posto che ci sia un elemento di consistenza minimo nelle sue preferenze.<sup>20</sup>

Una volta che l'agente abbia assegnato delle preferenze numeriche a stati probabili del mondo, noi non abbiamo bisogno di null'altro per predire la sua scala di preferenze, se non di verificare che gli assiomi siano stati rispettati.<sup>21</sup> Questa teoria è puramente predittiva, nel senso che predice il comportamento dell'agente, se l'agente segue l'ordinamento delle preferenze, ordinamento che viene messo in forma dall'assunzione di quegli assiomi. Non dice nulla di ciò che accade nella mente degli agenti. Questa asserzione cautelativa era anche in accordo con il generale *milieu* culturale comportamentistico all'interno del quale von Neumann e Morgenstern fornirono la loro assiomatizzazione. Tuttavia, le interpretazioni mentalistiche del teorema dell'utilità attesa sono piuttosto diffuse, nel senso che è sembrato consequenziale assumere che, se l'agente soddisfa gli assiomi nell'ordinamento delle sue preferenze, è allora un agente che agisce in accordo con la teoria dell'utilità attesa, come se nella sua mente accadesse una interpretazione della prova matematica in un senso normativo. Io ritengo, invece, che ciò che la formalizzazione indica è che è possibile rappresentare delle preferenze e un loro ordinamento in base a determinati assiomi; viceversa, si tratta di un equivoco pensare che *ogni* rappresentazione delle preferenze e del loro ordinamento debba seguire quell'ordinamento assiomatico. Questo è stato un equivoco fortunato nelle scienze sociali perché apriva la strada maestra al progetto dei confronti intrapersonali ed interpersonali delle utilità. Contro questo equivoco, ad esempio, polemizza Arrow<sup>22</sup>, che non crede affatto che la comparazione intrapersonale delle preferenze sia possibile in linea di principio ed analogo spirito polemico si ritrova in coloro<sup>23</sup> che negano la pretesa di porre a

---

<sup>20</sup> Questa posizione è talvolta espressa da R. D. Luce & H. Raiffa, *Games and Decisions*, cit., p. 21, che indica in maniera significativa quanto forte sia stata la tentazione di stabilire un nesso sostanziale fra preferenze e un modello determinato di funzione di utilità pensato per rappresentarle.

<sup>21</sup> P. Hammond, *Consequentialism and the Independence Axiom*, in B. R. Munier (ed.), *Decision and Rationality*, Dordrecht, Reidel, 1988, e A. Sen, *Evaluator Relativity and Consequential Evaluation*, in *Philosophy and Public Affairs*, 1983, pp. 113-132.

<sup>22</sup> K. Arrow, *Social Choice and Individual Values*, Wiley, New York, 1951.

<sup>23</sup> Nello specifico, Broome, Hammond, Sen.

fondamento dell'utilitarismo quella specifica versione della teoria dell'utilità attesa.

In definitiva, la teoria dell'utilità attesa non è affatto una modalità di misurare le preferenze per ordinare dei valori e costruire un'immagine rappresentativa di ciò che un agente razionale dovrebbe fare. Nell'apparato dimostrativo non vi è alcuna indicazione su come stabilire legami analitici con preferenze in quanto rappresentino valori o con preferenze in quanto rappresentino funzioni di benessere, sebbene l'idea che questo legame ci sia o sia quanto meno implicito e facilmente deducibile abbia trovato sostenitori entusiasti fra molti non economisti<sup>24</sup>. Piuttosto, indica come rappresentare delle preferenze che si ordinano in un modo determinato perché soddisfano determinati assiomi, in maniera tale da poter predire come dovrebbe scegliere un agente in condizioni di rischio. La conclusione che poiché la prova è valida, allora gli agenti sono dei massimizzatori di utilità, è indebita, così come lo è quest'altra: che le preferenze degli agenti siano generate da una qualche struttura psicologica profonda che è formalizzata dalla prova. I numeri generati da una funzione di von Neumann e Morgenstern non rivelano nulla sull'"utilità in generale" e nulla sulla misura del "benessere"<sup>25</sup>. Un legame fra la funzione di utilità generata dall'aderenza di una determinata rappresentazione dell'utilità attesa per l'agente e determinati valori dovrebbe fornire prove circostanziate della necessità di questo legame, eppure la consapevolezza della necessità di queste prove è alquanto scarsa. Questo legame è difficile possa essere mai provato in un senso analitico, perché la teoria dell'utilità attesa è normativa, e non ci fornisce affatto un modello predittivo sulla base del quale giudicare del comportamento umano. Questo significa che nulla ci autorizza a pensare che, ammesso che le nostre preferenze soddisfino gli assiomi iniziali, nella nostra mente ci siano delle funzioni matematiche che guidano e indirizzano il nostro comportamento o anche soltanto che il nostro comportamento dovrebbe conformarsi a queste funzioni quando operiamo delle scelte. Si tratta, in realtà, di uno strumento per rappresentare un ordine e nemmeno l'unico. Non è invece una teoria della ragione, o una teoria del ragionamento nell'azione.

---

<sup>24</sup> È il caso di D. Gauthier, *Morals by Agreement*, cit.

<sup>25</sup> G. Loomes & R. Sugden, *Regret Theory: An Alternative Theory of Rational Choice under Uncertainty*, in "Economic Journal", 1982, pp. 805-824.

4. Si diceva che ci sono anche delle ragioni legate alla temperie culturale del momento in cui la teoria è stata elaborata che supportano una interpretazione dell'utilità attesa come rappresentazione specifica e non unica e non invece come una funzione psicologica. Infatti, l'idea che l'utilità attesa rappresenti degli stati interni alla mente dei soggetti non avrebbe passato nessuno dei test neopositivisti e behavioristici sulla rilevanza e sul significato di come noi valutiamo qualcosa attraverso la funzione di utilità, in quanto questa valutazione è qualcosa di distinto da una semplice disposizione a scegliere.

Quando si interpreta l'utilità attesa come un modello del comportamento che è generato dalla nostra mente, l'operazione che si compie può essere di duplice natura: o si tratta di un'interpretazione di norme di comportamento o si stanno interpretando processi di ragionamento relativi all'azione. Adottando l'una o l'altra di queste due interpretazioni – talvolta entrambe – il minimo che si possa dire è che la teoria viene forzata in direzione di un risultato etico-pratico – “si deve agire in questa maniera” – o gnoseologico-metafisico – “queste sono alcune delle caratteristiche della mente” –, nessuna delle quali si ritrovano negli assiomi. Ma fino a qui, nulla di male, naturalmente: se viene ritenuta un'attività ermeneutica legittima pensare quello che un pensatore non ha effettivamente pensato, ma avrebbe dovuto pensare in base alle sue premesse, allora non c'è motivo per ritenere a priori un fraintendimento ogni estensione di una teoria in campi, che originariamente non entravano nella sua formulazione iniziale. Tuttavia, questa estensione dell'utilità attesa verso una dimensione conoscitiva e strutturale, rappresenta almeno in parte un fraintendimento, per quanto non ne manchino le ragioni. Ne segnalo almeno una, ossia l'idea che l'ordinamento delle preferenze che dovrebbe guidare determinate scelte rimandi a un insieme di leggi psicologiche che sono identiche per tutti gli esseri umani<sup>26</sup>. Ma è necessario che una teoria puramente predittiva debba basarsi su un'assunzione tanto impegnativa quanto indimostrata? Non è forse vero che la propensione soggettiva – e culturale – al rischio è un fatto che pare per lo meno indebito escludere dalla ricerca dell'utilità? Le stesse difficoltà sulla raccolta delle informazioni rilevanti – sulla difficoltà di decidere quale sia l'informazione rilevante e quale non lo sia – potrebbero consigliare di argomentare, a voler essere prudenti, che molti agenti non agiscono solo dopo aver attribuito valori probabilistici intersoggettivi definiti agli eventi, anche senza concludere che si

---

<sup>26</sup> J. Harsanyi, *L'utilitarismo*, (a cura di S. Morini), Il Saggiatore, Milano, 1988 fa ripetutamente propria questa posizione.

comportino irrazionalmente<sup>27</sup>. Se la teoria dell'utilità attesa non era stata pensata per funzionare come una teoria della ragione o come una teoria del ragionamento corretto si potrebbe però sostenere che aggiungere corollari normativi e psicologici è operazione legittima e indispensabile per la completezza della teoria stessa. Per questo motivo, affiancare la sovrastruttura psicologica, che considera le preferenze dell'agente come ciò che definisce i fini della sua azione, all'interpretazione formale dovrebbe fornire un senso molto meno indefinito all'intera faccenda, poiché se ne è concluso che non qualsiasi preferenza possa svolgere questa funzione, ma unicamente quelle che soddisfano gli assiomi.

Questo punto è decisivo, perché se esiste un vaglio normativo attraverso il quale devono passare le nostre preferenze per poter essere utilmente considerate come formatrici di fini, questo significa che gli assiomi non sono strumentali e che perciò la stessa teoria dell'utilità attesa, in quanto non è altro che una derivazione da questi teoremi, avrebbe una portata normativa. Anche, quindi, utilizzare le forme di ragionamento, che su questi teoremi si basano, avrebbe una portata normativa e non invece strumentale. Questa posizione, ad esempio, è stata fatta propria da David Gauthier, che ha considerato l'assioma di transitività come una sorta di test di consistenza delle nostre preferenze, interpretandolo quindi in senso normativo<sup>28</sup>. Il passo successivo è leggere gli assiomi come regole della logica, suggerendo quindi che si tratta di regole non strumentali, che vincolano la coerenza di un insieme determinato delle nostre preferenze. Tutto questo sembra violare il *mantra* humaneo che la ragione è solo schiava delle passioni – anche se Hume diceva che la ragione è e deve essere schiava delle passioni, mischiando opportunamente il livello descrittivo con quello prescrittivo –. Una volta fatta questa piccola deviazione dal solco di Hume, come si fa ad essere sicuri che altri principi normativi non debbano essere invocati? Come si fa ad essere certi che, ad esempio, considerazioni etico-morali non debbano essere messe in gioco per assicurare una qualche forma di stabilità nel corso del tempo alle nostre preferenze? Siamo d'accordo che la stabilità nel corso del tempo non è la stessa cosa della coerenza logica, ma perché la prima dovrebbe essere meno importante della seconda? La stabilità della cooperazione si basa, infatti, su un assetto delle preferenze di ciascuno degli agenti coinvolti, che si presume

---

<sup>27</sup> Si vedano le ricerche contenute in D. Kahneman, P. Slovic, A. Tversky (eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, 1982.

<sup>28</sup> D. Gauthier, *Moral Dealing*, Cornell University Press, Ithaca, 1990, pp. 234-273.

non cambierà per un periodo di tempo rilevante. La strategia *tit for tat* è precisamente una scommessa su questo genere di stabilità<sup>29</sup>.

Questa stabilità è almeno altrettanto importante della coerenza formale delle preferenze di ciascun agente. Analogo discorso vale per altri standard normativi che potrebbero essere chiamati in causa; ad esempio, standard di carattere estetico o di carattere religioso. Questo significa che, poiché una concezione non strumentale ovvero normativa della ragione è alla base di una concezione strumentale della ragione, l'insieme di questa visione è, in definitiva, non strumentale, a meno che non si ritenga che gli assiomi che dovrebbero dare forma all'insieme delle nostre preferenze in gioco siano così evidenti da non richiedere alcuna discussione. Questo però è tutt'altro che ovvio. Esistono varie strategie per evitare questa commistione del normativo con lo strumentale che ora occorre esaminare brevemente. Una prima mossa è questa: sostenere che pur non essendo immediatamente evidente che gli assiomi siano dei vincoli strumentali alle nostre preferenze, rimarrebbe vero che abbiamo delle ragioni strumentali per seguirli. Ossia, dal momento che il risultato della messa in forma delle nostre preferenze, che si strutturano attraverso gli assiomi, è la consistenza, si potrebbe dire che un agente che desidera raggiungere i propri scopi massimizzando la propria utilità farebbe meglio ad essere coerente nelle proprie preferenze. Questo è un consiglio e una prescrizione che si traduce in una difesa consequenzialistica degli assiomi. In altre parole: è una difesa la quale non si esprime in nulla sulla natura di quegli assiomi e, in particolare, non si esprime sul fatto se questi assiomi rappresentino o meno una formalizzazione di regole naturali e fondamentali nel ragionamento della mente umana. A questo si deve aggiungere che tale posizione, d'altra parte, non sembra suggerire una difesa coerentista della opportunità di adottare il sistema di assiomi, poiché non si afferma che noi abbiamo buone ragioni per accettarli dal momento che si dimostrano in accordo con altri principi della mente. Quello che si dice è che l'esperienza pare suggerire che adottare una sistemazione coerente delle proprie preferenze consente di raggiungere in maniera relativamente più diretta – con un avverbio e un aggettivo anch'essi intrisi di normatività – i propri scopi. Ne consegue che quando si afferma che un agente ha un insieme di preferenze che non è strutturato coerentemente, perché l'ordine di questo insieme non gli consente di massimizzare l'utilità nella versione standard della teoria, ciò che si dice è in realtà una tautologia, poiché l'utilità attesa è definita nei termini

---

<sup>29</sup> R. Axelrod, *Giochi di reciprocità* (1984), Feltrinelli, Milano, 1987.

di strutturare le proprie preferenze secondo quegli assiomi. Si potrebbe però sostenere che se la teoria non indica il modo preciso di specificare come si determina il mezzo per raggiungere un fine specifico, questa non è affatto una buona ragione per negare la base intuitiva della teoria, precisamente quella che si richiamava in apertura di queste pagine, ossia che ciò che l'agente fa è quello che lui ritiene debba effettivamente essere fatto per soddisfare un insieme di desideri. Del resto, proprio una considerazione intuitiva di questo genere sembrerebbe essere alla base dell'enfasi sulla coerenza delle preferenze fra di loro.

Questo ci riporta alla considerazione svolta più in alto: non è tanto strumentalmente razionale seguire gli assiomi per soddisfare le proprie preferenze; piuttosto, è l'uso di questi assiomi che rende coerente l'insieme delle preferenze e perciò soddisfare queste preferenze è parte di ciò in cui consiste un ragionamento e un comportamento strumentale. Può ben essere che una qualche prospettiva sulla correlazione mezzi-fini sia necessaria, così come può ben essere che sia necessaria una considerazione globale sul ragionamento che determina i mezzi in vista di questi fini; tuttavia, questa è palesemente un'altra cosa dalla caratterizzazione della teoria come struttura razionale che prescinde da considerazioni relative al valore. Se gli assiomi forniscono una sorta di vaglio attraverso il quale far passare le nostre preferenze *prima* che noi possiamo affermare di aver agito razionalmente in base a quelle stesse preferenze, ciò significa che definiscono che cosa significa essere un fine per l'azione in maniera indiretta. Allora la conclusione che se ne deve trarre è che questi assiomi non possono essere difesi sulla base di considerazioni consequenzialistiche. Esistono, inoltre, numerose evidenze empiriche che mostrano che gli agenti reali violano continuamente gli assiomi della teoria. Alcune di queste violazioni sembrerebbero intuitivamente ragionevoli, cosicché la teoria classica dell'utilità attesa non sembrerebbe potersi proporre nemmeno come una teoria plausibile del comportamento umano. Mi limiterò a due assiomi della teoria, l'assioma di transitività e l'assioma di riduzione delle lotterie composte. L'assioma di transitività afferma che se un agente preferisce  $a$  a  $b$  e  $b$  a  $c$ , allora preferisce anche  $a$  a  $c$ . Immaginiamo ora un agente che possa compiere tre azioni differenti. Ognuna di queste tre azioni ha tre possibili esiti, tutti e tre equiprobabili. Possiamo immaginare una situazione perfettamente coerente dove  $a$  viene preferito a  $b$  e  $b$  viene preferito a  $c$ , ma dove  $c$  viene preferito ad  $a$ . Se preferisco il caffè ( $a$ ) alla cioccolata ( $b$ ) e la cioccolata ( $b$ ) al tè ( $c$ ), posso in tutta coerenza preferire

il tè (c) al caffè (a). L'assunzione che sta alle spalle del teorema delle lotterie composte è che se è possibile ridurre una lotteria composta a una lotteria semplice, l'agente sarà indifferente alle diverse lotterie. Il significato di questo assioma è dubbio, perché non è ovvia quale ne sia la sua rilevanza e alcuni dubitano che possa essere realmente considerato un assioma, ma necessiti di ulteriori assunzioni<sup>30</sup>.

In realtà, il comportamento umano sembra essere una violazione piuttosto frequente di questo assioma. Di norma, infatti, gli individui non sono indifferenti al numero di passi necessari a determinare gli esiti delle loro azioni. Se ho desiderio di comprare una nuova automobile e adotto la soluzione di chiedere un finanziamento, anziché quella di risparmiare per quattro anni, non posso certo dire che la seconda soluzione sia equivalente alla prima. Il desiderio di evitare determinati passi intermedi – quattro anni di risparmi – potrebbe guidarmi a preferire una determinata risoluzione anziché un'altra. Si può anche dare il caso di una persona che sia sensibile a una lotteria che contempla diversi risultati intermedi verso l'esito finale e si accontenta di raggiungere alcuni di questi risultati, anche se il loro raggiungimento non aumenta la probabilità di raggiungere l'esito finale. Questa sembra essere l'esperienza dei giocatori di azzardo. Se io offro una somma di denaro a un incallito giocatore, somma che lui spera di guadagnare giocando, a patto che non giochi, che cosa accadrà? Si comporterà molto probabilmente come il marito di Doña Flor nel romanzo di Jorge Amado. Mi darà le più ampie assicurazioni che non andrà a giocare e, intascata la somma, si dirigerà verso la bisca più vicina. Non si tratta solo di una notazione psicologica, quanto di segnalare il fatto che per il giocatore è importante *come* raggiungere il risultato che desidera. Questa è un'esperienza analoga a quella che sperimentano agenti che desiderano raggiungere un risultato, che richiede una lunga preparazione professionale. Esistono, ovviamente, dei casi in cui il vertice professionale viene raggiunto ad inizio carriera in maniera del tutto meritata, ma in genere si tratta di casi piuttosto rari. Ci sono, invece, casi in cui questo vertice viene raggiunto rapidamente per circostanze estrinseche alla professione (amicizie, appoggi politici, e cose del genere). Molti rifiuterebbero una soluzione così rapida delle loro carriere, in assenza di circostanze basate sul merito. Il perché è abbastanza chiaro: fa parte integrante delle utilità della lotteria che si preferisce che questa si svolga in un modo determinato.

---

<sup>30</sup> D. Kreps, *Notes on a Theory of Choice*, Westview Press, Boulder, 1988, pp. 50-55.

Alcuni agenti amano il rischio e la *suspense* e daranno un valore diverso e più alto a una lotteria, che contempra diversi stadi di ansia verso il raggiungimento dell'obiettivo finale, rispetto a una lotteria semplice dove questa *suspense* sia ridotta. Non c'è nessuna ragione evidente di pensare che un'attitudine verso il rischio non possa essere contemplata in un calcolo di utilità per l'agente, perché la modalità in cui si svolge una lotteria può avere per lui un valore del tutto rilevante. Per questo una lotteria con più passi può essere preferita a una più semplice. Un resoconto descrittivo della prassi umana in situazioni di rischio sarebbe incompleta se ignorasse questa possibilità. In molte lotterie il piacere di stare giocando in un determinato modo, anziché in un altro, deve essere incluso nella stessa funzione di utilità. Se io voglio conquistarmi l'attrazione di una persona e posso scegliere se invitarla fuori a cena nella speranza di iniziare una relazione formale oppure farle ingerire inconsapevolmente una sostanza psicotropa, che la farà innamorare di me, non è indifferente come io raggiungo il mio obiettivo. In realtà, non esiste nessuna buona ragione normativa per negare una legittima propensione al rischio. In altre parole, l'assioma non consente un'attitudine positiva del giocatore verso il fatto stesso di partecipare a una lotteria. Si tratta, però, di un'avversione moralistica, che nega che nelle scelte dell'agente possa entrare in gioco la considerazione di fonti di utilità dirette al fatto stesso di giocare proprio un determinato gioco, ad esempio, basato sopra un qualche ordine lessicografico. L'idea che le preferenze siano solo sui risultati delle lotterie e non possano anche essere sulla qualità delle lotterie pare, quindi, essere molto controintuitiva.

5. Le preferenze possono essere molto diverse<sup>31</sup>. Posso volere il mio piacere immediato, la soddisfazione di un vicino di casa particolarmente simpatico, la pace nel mondo, la riduzione del debito internazionale e così via. In questo senso, la funzione di utilità dovrebbe essere uno strumento per misurare e ordinare delle preferenze nella mente dell'agente, ma non è anche uno strumento che valuta e decide a che cosa servono le preferenze. Questa interpretazione della funzione di utilità non è nemmeno a rigore una teoria del

---

<sup>31</sup> Molti ritengono che il padre del contemporaneo utilitarismo, H. Sidgwick, abbia sostenuto una posizione che relaziona il bene all'informazione completa, sebbene un'affermazione del genere non si ritrova nella sua riflessione. Si veda R. Shaver, *Sidgwick's False Friends*, in "Ethics", 1997, 107, pp. 314-320

ragionamento. Prende semplicemente atto che l'agente ha delle preferenze le cui fonti motivazionali possono essere molto diverse e che queste preferenze possono essere rappresentate in una determinata maniera. Probabilmente, questa interpretazione più debole è quella preferita dalla maggior parte dei teorici perché pare essere particolarmente parca nelle sue assunzioni, mentre è molto più problematica la classica interpretazione monistica di Bentham, che insiste sul fatto che, in definitiva, noi abbiamo nelle nostre strutture mentali solo una preferenza, quella per il piacere, per la sua massimizzazione complessiva, cosicché se alcune preferenze sembrano essere incoerenti fra di loro, la rappresentazione permessa dall'applicazione degli assiomi garantisce che il risultato complessivo atteso dall'ordine delle preferenze riformato non lo sia. In entrambe le interpretazioni, quella pluralistica e quella monistica, gli assiomi sarebbero perciò una maniera di generalizzare un modo di ragionare che informalmente è presente nella struttura psicologica degli esseri umani, la formalizzazione di una funzione mentale se non, addirittura, di una funzione biologica. Anche questo ritratto si espone a considerazioni non consequenzialistiche nella misura in cui la mappatura delle preferenze che l'agente disegna è almeno in parte funzione del mondo in cui ci si attende che le conseguenze avranno luogo, come il risultato di lotterie, di stati di certezza, di ciò che lui interpreta come il contesto stesso nel quale avviene la scelta.

Per questo la psicologia del decisore deve essere una psicologia più complessa di quella prospettata dalla utilità attesa. A meno che non si ritenga che le preferenze siano generate in noi in maniera tale da essere spontaneamente coerenti, gli assiomi della teoria non possono non avere un aspetto normativo. Da questo punto di vista, la loro funzione dovrebbe essere più quella di renderci consapevoli che non possiamo volere realmente cose contraddittorie, che non quella di descrivere un modo di ragionare. L'idea che l'agente debba essere capace di distinguere non solo le proprie preferenze relativamente ai risultati, ma anche le proprie preferenze relativamente al modo in cui i risultati vengono generati, non può affatto essere messa da parte, a meno di non introdurre appositamente un assioma. Ammettiamo che tu debba scegliere in una lotteria fra 200 euro certi e una determinata somma  $S$  incerta, ma superiore ai 200 euro. Mentre i 200 euro non accresceranno la tua utilità marginale, se non di poco, la vincita di  $S$  potrebbe, forse, cambiare la tua vita. Si può argomentare la tua preferenza per i 200 euro sostenendo che tu non ami i rischi, che le lotterie con esiti incerti, anche se estremamente vantaggiosi, ti rendono nervoso e ti procurano una sofferenza che può essere

lenita solo dalla conoscenza del risultato, anche se a te sfavorevole. È, in questo esempio, il contesto stesso della scelta che determina la tua preferenza. Immaginiamo ora che un agente dica di preferire S perché è portato ad assumere rischi e perché l'utilità attesa da S è più alta della utilità rappresentata dalla certezza dai 200 euro. In questa versione, l'utilità di S non è affatto uguale a quella della prima versione, perché nel primo esempio S incorpora per l'agente prudente (forse eccessivamente prudente) anche la disutilità di procurarsi lo stress associato alla partecipazione alla lotteria stessa. Per questo motivo, l'utilità di S nella seconda situazione è superiore all'utilità di S nella prima, perché in questa ultima abbiamo rintracciato un certo sentimento di avversità dell'agente al rischio. In ogni scelta si possono trovare, indagando abbastanza approfonditamente, motivazioni che sono riconducibili a contesti della scelta leggermente differenti, che rendono situazioni apparentemente simili, in realtà, diverse. Questo può essere imbarazzante, perché potrebbe giustificare qualsiasi agente dall'accusa di aver violato qualsiasi teorema. Poiché differenze fra alternative anche estremamente simili possono essere sempre rintracciate con un po' di buona volontà, allora ogni accusa di irrazionalità a qualsiasi azione potrebbe dimostrarsi infondata e sostanzialmente inconfutabile. Una teoria di questo genere, poiché ricorre continuamente a giustificazioni *ad hoc*, sembra inutile come strumento descrittivo dell'azione e come strumento di previsione; e, tuttavia, ciò non autorizza a inferire che sussistano buone ragioni normative per escludere il contesto della scelta. Alcuni hanno sostenuto che in realtà non si tratta nient'affatto di una situazione problematica. Tutto ciò di cui abbiamo bisogno è di osservare come si comporta l'agente in una situazione di incertezza ossia come manifesta apertamente le sue preferenze.

Questa prescrizione, tuttavia, non ci dice affatto come discriminare una preferenza da uno stato di indifferenza. Se entro dal fruttivendolo, cerco di dare un'occhiata in giro e poi me ne esco con un chilo di mele, questo non informa affatto se io preferivo semplicemente acquistare qualcosa nel primo negozio che ho incontrato sulla mia strada, se intendevo comprare un chilo di pesche, ma poi ho ripiegato su uno di mele, perché le pesche erano troppo care, oppure se preferisco le mele alle pesche. Questa incertezza si verifica per il semplice fatto che l'attività di preferire è un'attività interpretativa e non un semplice evento da osservare disponibile a tutti "là fuori nel mondo". Di quale interpretazione allora dovremmo fare uso per identificare correttamente le preferenze e il loro ordine? Un modo potrebbe esser questo: individuare le

alternative attraverso le quali le preferenze vengono formate. Alcune di queste saranno indifferenti per l'agente, ossia preferire l'una o l'altra non cambierà sostanzialmente nulla nella sua utilità attesa. A questa procedura si dà il nome di principio di indifferenza razionale<sup>32</sup>. Si noti, tuttavia, che anche questo principio è, ancora una volta, marcatamente normativo, poiché ci dice che cosa dovrebbe fare un agente di fronte a due situazioni ragionevolmente simili in base a un qualche principio di individuazione, ossia gli si dice sostanzialmente, di nuovo, quali preferenze dovrebbe avere in determinate situazioni, ricorrendo all'intuizione di due situazioni ragionevolmente simili, senza tenere in conto che l'idea di 'situazioni ragionevolmente simili' può essere potentemente influenzata da presupposti culturali e da differenze individuali. Ciò rende una parte importante della teoria dell'utilità dipendente da un qualche incerto appello all'intuizione. Tuttavia, forse non occorre imbarcarsi lungo il percorso suggerito dal principio di indifferenza razionale e accettare molto semplicemente che le preferenze sono definite soggettivamente, e che questa definizione fa parte degli stati psicologici dell'agente. Queste preferenze sono individuabili perché sono oggetto dell'esperienza dell'agente. In questo senso, se il rimorso o la paura o la gioia o l'esaltazione fanno parte per l'agente delle conseguenze attese dell'esperienza di un'azione allora dovremmo includere uno, alcuni, tutti questi sentimenti nel calcolo dell'utilità. Per questa prospettiva, il corso d'azione che l'agente sceglie è un dato della realtà dal quale non si può astrarre non più di quanto si possa astrarre dai sentimenti associati alla scelta di un particolare corso d'azione a preferenza di un altro. Che cosa può suggerire questo? Che non sembra essere un'operazione illuminante incorporare le preferenze per gli stati nelle preferenze per le conseguenze, quando effettivamente le preferenze sono per gli stati e non per le conseguenze.

L'agente può agire indipendentemente dalle informazioni sulle probabilità che ha acquisito sulle conseguenze delle sue azioni, privilegiando gli stati del mondo nei quali si svolgono differenti lotterie composte che per lui rimangono irriducibili. Qual è la morale che possiamo trarre da queste ultime riflessioni? Che la teoria dell'utilità attesa nella sua interpretazione consequenzialistica come teoria *tout court* della ragione è inadeguata. La nostra ragione, il nostro ragionamento è spesso sensibile ad altre cose che alle conseguenze. Questo è un semplice dato che ricaviamo dall'esperienza e non c'è nessuna ragione né

---

<sup>32</sup> J. Broom, *Weighing Goods*, cit., 1991, pp.89-95.

descrittiva né prescrittiva per sottovalutarlo e non prenderlo in considerazione. Ovviamente, tutto questo non significa affatto che la teoria dell'utilità attesa non sia utile in contesti economici per descrivere comportamenti che accadono nel mercato, e nemmeno per sostenere che questi specifici comportamenti siano forse quelli maggiormente rilevanti in quel contesto. Gli economisti, quindi, hanno tutte le ragioni di usare della teoria per effettuare delle previsioni. Il punto è non confondere la capacità di fare delle previsioni in contesti determinati (ammesso che questo sia possibile, cosa che certamente non è di mia competenza giudicare) con una qualche teoria monistica delle capacità, inclinazioni, propensioni della mente. Se la teoria è realmente predittiva in determinati campi ristretti o, più probabilmente, in determinate circostanze all'interno di questi campi, questo non accade perché si tratta di una teoria capace di fornire un'immagine adeguata della maniera nella quale ragioniamo né del modo nel quale dovremmo ragionare. Ora, dal momento che ci sono forme di ragionamento e di comportamento che palesano violazioni degli assiomi della teoria dell'utilità attesa, e che non sembrano affatto essere irrazionali, se sono necessarie delle modificazioni, queste non sembra si debbano situare dalla parte del comportamento e dell'azione e delle sue giustificazioni, ma dalla parte della teoria. Le giustificazioni dell'azione dovranno essere situate in un contesto che tenga conto che non tutto il nostro agire è consequenzialista, ossia che spesso il nostro ragionamento è contestuale in un senso profondo e non episodico. Se il consequenzialismo è una teoria falsificabile e non una qualche sorta di teoria mitica sull'agire, l'idea della dipendenza contestuale dovrebbe essere considerata come una sorta di test delle concezioni del ragionamento sul quale è basato. Evidenze che ognuno di noi può trarre dalla propria vita, ma che ci vengono anche dalle scienze empiriche, suggeriscono fortemente che noi non siamo *solo* degli agenti consequenzialisti, sebbene siamo *anche* degli agenti che operano in base ad assunzioni sulle conseguenze. Questo può aprire la strada a una teoria del comportamento che integri gli assiomi della teoria dell'utilità attesa, mantenendo il suo spessore strumentale? Io penso che occorra essere fortemente scettici su una prospettiva del genere. Le fonti motivazionali dell'azione umana sono molteplici, e mosse riduzionistiche, sebbene non siano affatto illegittime a priori – e siano, per altro, storicamente ricorrenti – oltre che, probabilmente, necessarie ai fini della modellizzazione, perdono l'essenziale del comportamento umano, ossia la sua varietà nell'azione e nelle fonti della

motivazione. Ma vi è un'altra conseguenza che penso si possa trarre e che è di maggiore rilevanza per la filosofia morale ed è questa: se il nostro ragionamento per l'azione presenta delle caratteristiche importanti che non sono sussumibili nel modulo consequenzialista, perché dovremmo pensare che il nostro agire morale invece lo sia?

Per quanto ritenere che consequenzialismo e strumentalismo siano sinonimi possa essere un errore, e un agente consequenzialista non sia vincolato all'idea che le sue azioni siano *tutte* strumentali, tale agente rimane, tuttavia, vincolato all'idea che il bilanciamento e la scelta delle azioni da compiere siano inquadrati nel modulo consequenzialista. Ma questo modulo può assumere una varietà plurale di forme e, perciò, una prospettiva strumentale implica il consequenzialismo, ma non vale l'inverso. Le preferenze relative agli stati e quelle relative alle conseguenze sono fundamentalmente differenti e non possono essere ridotte le une alle altre, il che equivale a dire che non esiste affatto un unico genere di cosa che richiede di essere massimizzato né un'unica procedura per farlo, né un'unica procedura per massimizzare cose diverse; ne consegue che le prescrizioni su come soddisfare un insieme di preferenze non possono affatto evitare di tenere conto della integrazione e della modificazione reciproca di differenti generi di preferenze. Questa necessità è, credo, particolarmente chiara se immaginiamo che alcune preferenze per gli stati siano di genere morale o abbiano una base morale.

Detto questo, si potrebbe ancora sostenere che la teoria dell'utilità attesa è una forma di idealizzazione, che come tutte le modellizzazioni lascia da parte qualcosa per concentrarsi su ciò che viene ritenuto maggiormente rilevante? Penso che occorra dire che se noi abbiamo certamente necessità di modellizzazioni, abbiamo però bisogno di modelli che siano adeguati all'oggetto. Se l'agire pratico non si conforma al modello dell'utilità attesa, non ha evidentemente senso dire che è l'agire a dover essere riformato. Siamo invece nella necessità di formulare un modello migliore. Una rappresentazione errata delle ragioni nell'azione è particolarmente infelice in ambito normativo, sia perché introduce surrettiziamente l'idea di una normatività idealizzata che sarebbe derivata dalle teorie scientifiche sia perché non riconosce che, se normatività deve esserci, questa andrà situata anche dal lato di principi che governino l'integrazione delle preferenze. È possibile che una difesa di questa prospettiva debba assumere una strategia coerentista. Questa strategia dovrà bilanciare sia le nostre attese sulle conseguenze sia le nostre preferenze per gli stati. Probabilmente risulterà ancora inaccettabile a chi

PIERPAOLO MARRONE

coltiva sogni di ingegneria sociale, ma nelle sue inevitabili approssimazioni sarà più vicina al pluralismo dell'azione umana.