

Identità personale, preferenze, narritività

Pierpaolo Marrone
Dipartimento di Filosofia
Università di Trieste
marrone@units.it

ABSTRACT

The paper explores various conceptual formats that are designed to help us organize our thoughts concerning what we do, especially personal identity, intertemporal preferences, and maximization. I discuss some suggestions from Parfit on personal identity and from Hare on interpersonal preference.

1. Indubbiamente, quando agisci hai intenzionalmente di mira il raggiungimento di un qualche obiettivo. Questo obiettivo sarà sempre specifico e singolare, anche se tu difficilmente ti accontenteresti di limitare le tue prospettive a un risultato particolare. In fin dei conti, la tua vita, come quella di chiunque altro, proprio perché è tua, soddisfa alcuni requisiti minimi di integrità personale e non ti appare come una rapsodia recitata da un attore diverso ad ogni istante, bensì come qualcosa che cerchi di comporre e di riconoscere come un'unità.

Gli esiti di questa operazione sono per lo più imperfetti, ma ciò che qui importa non è tanto la soddisfacibilità del risultato e nemmeno la sua praticabilità, quanto una certa costrizione, per non dire necessità, che ci guida a questo sforzo di coerenza narrativa, forse sempre incompiuta. Tutto questo vale per i tuoi atti intenzionali per lo meno, che per quanti pochi siano, sono quelli che ti vedono effettivamente agire come agente in prima persona. Non sto dicendo che siano solo o soprattutto gli atti intenzionali che danno sapore alla tua vita. Molto spesso sono le cose totalmente inattese a indirizzare nel bene e nel male le nostre vite e a dare loro il sapore specifico che hanno, e non invece i nostri atti intenzionali. Rimane però il fatto che quando agisci intenzionalmente, lo fai nella presunzione di essere proprio tu ad agire e non qualcun altro.

Le azioni intenzionali sono spesso compiute sulla base di quella che tu ritieni essere la migliore informazione disponibile al momento. Puoi naturalmente essere in errore e sbagliarti anche grossolanamente, ma è importante distinguere questa classe di azioni intenzionali da un'altra classe di azioni intenzionali, quelle dove decidi di ignorare la migliore informazione disponibile. Fatti salvi alcuni casi speciali, quando decidi di ignorare la migliore in-

formazione disponibile agisci in base a una forma di incontinenza. Ti accendi la sigaretta anche se hai un'informazione sufficientemente dettagliata dei danni che il fumo provoca. Una signora isterica ti passa platealmente davanti alla fila del supermercato. Reagisci in maniera fisicamente aggressiva anche se sai che la tua azione potrà avere conseguenze spiacevoli, ad esempio quelle contemplate dal codice penale o da un marito particolarmente prestante. Hai quindi agito in maniera sicuramente intenzionale, ma tuttavia con modalità incontinenti, le quali fanno sì che le tue azioni devono essere catalogate sotto il segno dell'irrazionalità. Tu stesso, se potessi fare una ripresa alla moviola della tua azione incontinente riusciresti ad indicare con una buona approssimazione il punto in cui le cose hanno cominciato ad andare per il verso sbagliato. È un segno questo alquanto dirimente del fatto che, secondo la tua stessa opinione, le cose non sarebbero dovute andare in quel modo, ossia del fatto che tu ritieni effettivamente di aver agito irrazionalmente e come non avresti dovuto agire sulla base del miglior giudizio disponibile per te. Si è stabilito quindi che le tue azioni intenzionali si possono raggruppare in almeno due categorie: quelle razionali e quelle irrazionali e incontinenti. A meno che non sia patologicamente portato ad azioni incontinenti che ti creano danni ripetuti, sono quelle che appartengono al primo insieme che costituiscono numericamente la maggior parte delle tue azioni intenzionali.

Quando tu agisci intenzionalmente e razionalmente che cosa ti prefiggi in effetti di ottenere? Al di là del risultato specifico, esiste uno scopo generale che è intenzionato dalla tua azione, ossia una sorta di oggetto ideale o una sorta di schema ricorrente, dotato di qualità specifiche che ricorrendo in atti appartenenti a quella determinata intenzionalità razionale li rende categorizzabili in un unico insieme? Io penso che le cose stiano effettivamente in questo modo, ma ritengo anche che l'analisi degli atti intenzionali razionali possa gettare una qualche luce non banale sul problema dell'utilità e su quello dell'intertemporalità, ossia su quella narratività che parrebbe essere un tratto costante della nostra descrizione come agenti morali.

Dunque: tu hai agito intenzionalmente e razionalmente. Che cosa hai fatto? Hai agito in maniera tale da soddisfare un qualche aspetto della tua volontà, ossia hai agito nella convinzione che con la tua azione avresti realizzato nella maniera migliore e nelle migliori condizioni conoscitive disponibili al momento per te un tuo desiderio. Tutto questo si può sintetizzare nella proposizione: (1) "agisco razionalmente quando faccio ciò che realizza un mio desiderio presente". Si noti che, sulla base di quanto finora si è detto, sarebbe pleonastico dire che agisco anche intenzionalmente. Se agisci razionalmente (ossia in maniera giustificata) allora agisci anche intenzionalmente. Non vale

invece la converso, ossia, come si è visto, non è sufficiente, sebbene sia necessario, agire intenzionalmente per agire razionalmente.

2. Il problema è che (1) copre condizioni di razionalità che potrebbero apparirci troppo modeste, sulla base di quella narratività che la maggior parte di noi considera una caratteristica importante e forse essenziale della nostra autocomprensione come soggetti morali. Una soddisfazione di un tuo desiderio presente non è affatto detto sia ciò che tu razionalmente dovresti perseguire, ossia non è affatto detto che ciò che tu ora desideri sia anche ciò che ti darà la soddisfazione più intensa in un futuro che dovresti tenere in conto. Puoi desiderare intensamente al momento attuale di possedere una Porsche cabrio e, ammesso rientri nelle tue disponibilità finanziarie, andare dal concessionario a comprarla. Potrebbe però essere che la soddisfazione che evidentemente tu provi per il possesso di una bella macchina sia meglio soddisfatta da un altro modello che magari è maggiormente durevole, più semplice da guidare, meno impegnativo nella manutenzione e così via. In altre parole, è possibile che quello che tu desideri al momento attuale rappresenti su una scala una intensità massima, relativamente a un gruppo di desideri omogenei, ma non sia, invece, ciò che ti procurerà nel futuro la soddisfazione maggiore. Per questo motivo, se ti ritieni in grado di fornirti un'opinione razionale sul perché hai deciso di fare quello che stai per fare, dovresti compiere quell'azione che pensi massimizzerà la tua soddisfazione relativamente al desiderio che hai anche in un segmento temporale maggiormente esteso del presente.

Se (2) “per un agente agire razionalmente significa fare ciò che crede massimizzerà la soddisfazione di un suo desiderio attuale”, allora noi introduciamo una dimensione temporale nella pianificazione delle tue azioni – ossia nella semplice tua volontà di essere razionale nell'azione – che in (1) non era precedentemente così chiara. Ma il passaggio da (1) a (2) è tutt'altro che innocente, dal momento che la dimensione temporale può diventare del tutto qualificante nella tua percezione di te stesso come agente razionale. Infatti, quali ragioni tu pensi di poter convincentemente esibire per dover preferire un determinato momento della tua vita anziché un altro? Perché privilegiare $t1$ a $t2$, domani a dopodomani al 19 maggio del 2015? Che cosa possiede di così speciale domani rispetto ad innumerevoli altri momenti della tua vita futura?

Notoriamente, tanto Machiavelli quanto Hume erano profondamente pessimisti sulle capacità previsionali dell'essere umano a partire da una constatazione gnoseologica, che per loro si imponeva con la forza dell'evidenza empirica: noi siamo troppo ancorati a tutto ciò che percepiamo come presen-

te e prossimo per essere capaci di svolgere piani che vadano oltre la vicinanza temporale e spaziale. La maggior parte di noi è costretta nella gabbia del presente, e solo alcuni, talvolta – il buon politico, ad esempio, secondo Machiavelli, che deve però essere sorretto tanto dalla virtù quanto dalla fortuna -, ne possono evadere, per lo più momentaneamente. Tuttavia, le mie preoccupazioni non sono qui essere di natura descrittiva. Mi interessa piuttosto il versante normativo della questione e in quale modo questo aspetto si intrecci con quello della inevitabile temporalità dell'azione razionale. Un primo intreccio potrebbe proprio essere questo: non ci sono ragioni né a priori né universali e forse nemmeno generali per preferire un momento del futuro a un altro. Ovviamente, ci possono essere molte ragioni particolari perché nello specifico corso di azione che tu intraprendi queste ragioni vengano preferite. Se sei uno speculatore che opera in borsa, agirai diversamente a seconda che tu abbia una strategia rialzista o ribassista, nel senso preciso che valuterai in maniera molto diversa segmenti temporali diversi. Se sei una giovane donna in carriera potrebbe non essere indifferente il fatto che tu scopra di essere incinta. Gli esempi possono essere moltiplicati *ad libitum*, ma riguardano sempre occasioni specifiche di azione per le quali si adducono sempre ragioni particolari e non la struttura dell'azione razionale. All'interno di questa struttura sembra ragionevole non accordare preferenze particolari a segmenti temporali specifici. In questo senso, (3) “se tu agisci razionalmente farai proprio ciò che ritieni massimizzerà la soddisfazione di quei desideri compatibili con la tua razionalità, che pensi in qualche modo si estenderanno al corso della tua esistenza”.

3. I tuoi stessi desideri presenti si presenteranno come degni di soddisfazione solo se saranno razionalmente compatibili con (3). Questo può causare, come si vedrà, dei problemi, perché quando si sostiene che il segmento temporale dove accade che i tuoi desideri siano soddisfatti non ha importanza, si sostiene che la narratività che costituisce la tua vita etica, narratività che naturalmente non può non essere temporale, prescinde da una determinata sequenza di segmenti temporali, ossia non può razionalmente farsi carico di preferire una data sequenza a un'altra. Questo in effetti pare essere alquanto controintuitivo e si cercherà di mostrare in queste pagine come il passaggio a (3) non sia affatto necessario.

Ammettiamo, tuttavia, per il momento che si tratti di una prospettiva invece plausibile per un agente razionale. Se lo facciamo non ci sorprenderà più di tanto veder sostenuta la posizione di chi pensa che (4) “non soltanto tu non dovresti prendere in considerazione la temporalità nella soddisfazione

dei tuoi desideri, ma non dovresti nemmeno a rigore preoccuparti di chi siano questi desideri". È questa la posizione di Parfit, per il quale l'identità personale, fondata sull'unità della coscienza, non è rilevante.¹ In effetti, sia alcuni degli esperimenti mentali immaginati da Parfit, sia descrizioni cliniche di soggetti affetti da lesioni cerebrali post-traumatiche sostengono la visione di Parfit dell'unità della coscienza come un'illusione. Non mi addentrerò in questo aspetto della filosofia di Parfit, anche se avrò modo di discutere alcune sue posizioni sulla neutralità temporale. Quello che mi preme invece suggerire con la presente discussione è che il passaggio a (3) non è inevitabile e, se questo sarà ritenuto ragionevole, allora forse anche il passaggio a (4) sarà ritenuto meno attraente. Naturalmente ci si potrà chiedere quale sia mai l'elemento attraente in (4). Penso sia piuttosto semplice dirlo: (4) aggiunge qualcosa di importante rispetto a (3) sul problema della neutralità. Nei nostri momenti migliori potremmo essere molto attratti dal pensare che è la struttura stessa dell'azione razionale a richiedere che tanto il tempo quanto l'identità personale dell'agente siano messi tra parentesi.

C'è un'accusa ricorrente che viene mossa all'utilitarismo e che è stata resa celebre da J. Rawls, da A. Sen e B. Williams,² ossia quella di non tenere in sufficiente conto dell'identità personale degli agenti. A parte il fatto che alcuni potrebbero non ritenerla affatto un'accusa e una carenza teorica, lo stesso potrebbe essere detto del contrattualismo o del kantismo. Lo si potrebbe dire di larghissima parte del codice civile e penale, eppure in questi casi nessuno pensa si tratti di una manchevolezza. Per quale motivo? Io credo perché dove è necessario che siano formulate regole di condotta e di limitazione dei comportamenti la prospettiva generale dovrebbe essere quella della clausola lockiana, ossia del non recare danno ad altri. La generalità della norma sembra escludere la narratività, almeno a un primo livello, ma non è invece così ovvio che la stessa esclusione debba essere richiesta quando parliamo di come sarebbe razionale che noi agiamo. In questo caso ne va della nostra capacità di concepirci, ognuno di noi, come centro intenzionale di azione.

Penso che questo ultimo punto non sia affatto non controverso, ma possa essere difeso con delle argomentazioni efficaci. Si prenda, ad esempio, la ben nota posizione che i desideri non esauriscono affatto il campo della motivazione. Si tratta della posizione kantiana, la quale è stata a più riprese soste-

¹ D. Parfit, *Ragione e persone* (1984), Milano, Il Saggiatore, 1989.

² A. Sen A., & B. Williams B. (a cura di), *Utilitarismo e oltre*, (1982), Milano, Il Saggiatore, 1984.

nuta anche recentemente.³ Secondo la posizione kantiana si può essere motivati all'azione senza che intervengano desideri. Anzi, la motivazione ad agire non basata sui desideri rappresenta tanto l'agire etico quanto l'agire razionale nella loro purezza. Kant, come è noto, era talmente convinto della cosa da ritenere l'espressione 'ragion pura pratica' un puro pleonasma rispetto a 'ragion pratica'. L'azione motivata da desideri è impura in quanto vi entrano come fattori causali le inclinazioni personali. Le inclinazioni personali sono soggettive e non permettono di rintracciare quella struttura motivazionale universale dell'azione che è il segno della moralità; al contrario, quando la motivazione ad agire è costituita da un'obbligazione, i desideri non vi hanno parte. L'azione etica è universale precisamente perché motivata dal dovere morale che è universale e perciò impersonale.

Questa posizione rifiuta radicalmente la plausibilità di (1) ed ha avuto grande fortuna, ma a me sembra che si avvolga in aporie inestricabili e che non resista a una confutazione del genere seguente. Abbiamo visto che agire razionalmente significa agire intenzionalmente, sebbene non valga la converso. Se agisci intenzionalmente ciò significa che alcuni tuoi atti mentali (credenze e progetti, e così via) hanno un potere causale sulla tua azione, ad esempio causano alcuni movimenti del tuo corpo che costituiranno proprio quell'azione che intendevi svolgere. Tutto questo potrebbe essere descritto anche in una maniera leggermente diversa e egualmente plausibile. I tuoi stati mentali, infatti, che influenzano anche la tua azione in maniera tale che questa si presenti come intenzionale, che cosa altro sono se non un desiderio di compiere proprio quell'azione? Sembra perciò difficile concepire che esistano ragioni che non diano luogo a desideri di compiere qualcosa. Se questo è plausibile, allora questa parte della filosofia morale di Kant deve essere rigettata. Infatti, l'azione intenzionale deve essere spiegabile, per il fatto stesso che è intenzionale, per mezzo di una relazione causale tra credenze, progetti e così via del soggetto. Questa spiegazione non è altro che una forma di causalità e, dal momento che credenze, progetti e così via che danno luogo a un'azione, danno luogo anche al corrispondente desiderio, sono i desideri ad avere forza motivazionale assieme alle credenze, ai progetti e così via che li generano.

Ciò che vale per l'azione intenzionale in genere, vale anche per quella razionale. La differenza più evidente è che nell'azione razionale dovrebbe essere l'informazione disponibile ad accordarsi con il desiderio. In altre parole, è la volontà di tener conto della migliore informazione disponibile che è dispo-

³ O. O'Neill, *Acting on Principle*, New York: Columbia University Press, 1975; C.M. Korsgaard, *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press 1996.

zionale rispetto al pensiero proposizionale che fa di un desiderio anche un desiderio razionale. Esempifichiamo. Io posso avere un intenso desiderio di diventare molto ricco, a partire da una situazione che rende questo desiderio estremamente improbabile, per la mancanza di capitali iniziali, di doti speculative, e di molte altre cose ancora. Se io acquisto un biglietto della lotteria, mi comporto irrazionalmente rispetto al mio desiderio e alle mie condizioni attuali? Dipende da quali altre informazioni io ho disponibili. Se io so che le mie probabilità di vincere una somma enorme, realizzando il punteggio massimo, sono all'incirca di una su seicentoventi milioni, e non ho altre chances ragionate di arricchirmi, e, allo stesso tempo, il costo della giocata ha un'influenza minima sul mio reddito, allora l'azione che si conclude nell'acquisto del biglietto della lotteria è generata da motivazioni e da un desiderio razionali. Se, viceversa, io non ho nemmeno la forza di recarmi a fare la giocata e preferisco fantasticare ad occhi aperti, non si può nemmeno dire che abbia delle intenzioni rispetto al mio desiderio – per quanto di improbabile realizzazione –, perché non ne esiste una traduzione in un'azione e il mio desiderio di arricchirmi rimane completamente vuoto. Rimarrebbe irrazionale anche se io dovessi giocare con la convinzione che le probabilità siano a mio favore per una qualche ragione.

Da questo non segue che due soggetti con un eguale bagaglio di informazioni avranno anche gli stessi desideri razionali. L'eguaglianza epistemica nelle conoscenze non porta nell'azione all'eguaglianza dei risultati, poiché, ad esempio, deve essere dato il peso adeguato a una eventuale differente propensione al rischio dei soggetti. Tuttavia, una volta accertata, si genereranno insieme differenti di desideri razionali, che potrebbero essere simili, se simile è la propensione al rischio dei soggetti. In altre parole, il potere causale di un identico bagaglio epistemico è una faccenda che deve essere risolta empiricamente, mentre rimane del tutto plausibile a priori che credenze, opinioni, progetti e così via siano necessariamente intrecciati a desideri che possono essere compatibili con questi e perciò razionali. In sintesi, sono questi i motivi per cui l'etica ha a che fare principalmente con desideri e deve essere rigettata la prospettiva etica kantiana che li concepisce come area dell'antropologia pragmatica.

4. Si è visto che il requisito temporale che si è enunciato in (1) presenta delle difficoltà che inducono a pensare che lo schema generale dell'azione razionale debba avere caratteristiche di marcata neutralità temporale. La neutralità temporale richiede, tra le altre cose, che l'agente razionale individui un desiderio che sia dominante intertemporalmente, non solo nella prospettiva di

evitare errori di calcolo morale, ma anche per evitare i paradossi dell'azione akratica. Parfit si è occupato estesamente delle difficoltà legate alla dimensione temporale delle varie teorie della razionalità pratica. Per illustrare i dilemmi intertemporali Parfit introduce la teoria degli obiettivi attuali (*present aim theory*) o P. “Supponiamo che in un dilemma del prigioniero il mio obiettivo sia quello di realizzare l'esito migliore per me. Secondo P in questo caso è razionale la scelta che arreca un beneficio a chi la compie. Se il mio obiettivo è quello di arrecare benefici agli altri o di superare il test kantiano, a essere razionale è la scelta altruistica. Se il mio obiettivo è di fare quello che fanno gli altri – magari perché non voglio essere un *free rider* – quale sia la scelta razionale è dubbio. Tutto dipende da quel che credo facciano gli altri”.⁴ Questa teoria è rigettata da Parfit, poiché P può essere in contrasto con i miei obiettivi di più lungo periodo. Il punto non è, tuttavia, solo questo. Parfit, infatti, segnala un contrasto ulteriore e maggiormente rilevante, vale a dire il contrasto tra P e la *teoria dell'interesse personale* o S, che è anch'essa una teoria della razionalità pratica. “S assegna a ciascuna persona questo obiettivo: conseguire quelli che per lei sarebbero gli esiti migliori e che consentirebbero alla sua vita di andare nel miglior modo possibile”.⁵ Posto questo obiettivo è lecito chiedere che cosa o quali condizioni lo realizzino. Parfit individua tre teorie che vi rispondono in maniera adeguata: a) l'edonismo; b) l'appagamento dei desideri; c) l'oggettivismo dei valori. “Tutte queste teorie sostengono inoltre che, nel decidere che cosa sia meglio per qualcuno, si dovrebbe assegnare lo stesso peso a tutte le parti del suo futuro. Gli eventi futuri possono essere meno prevedibili; e un evento prevedibile dovrebbe contare di meno se è meno probabile che accada. Non dovrebbe, invece, contare di meno per il solo fatto che, se accadrà, accadrà più tardi”.⁶ Questo requisito di neutralità temporale riguarda il carattere di razionalità pratica di cui è investita (S), ossia la sua portata deliberativa. In questo senso, è possibile riformulare (S) con “(S1) Per ogni persona c'è un fine ultimo sommamente razionale: che la sua vita sia, per lei, la migliore possibile”.⁷ Che il fine razionale ultimo sia la soddisfazione dei desideri dell'agente è un requisito il quale, unito a quello di neutralità temporale, porta a sostenere che l'agente dovrebbe fare ciò che comporta il soddisfacimento dei suoi desideri avendo come punto di riferimento l'intera sua vita. Si tratta di un requisito impegnativo, ma che riguarda unicamente la neutralità temporale non quella personale. Natural-

⁴ D. Parfit, *Ragioni e persone*, cit., p. 119.

⁵ D. Parfit, *Ragioni e persone*, cit., p. 6.

⁶ D. Parfit, *Ragioni e persone*, cit., p. 6.

⁷ D. Parfit, *Ragioni e persone*, cit., p. 7.

mente, c'è un contrasto tra (1) e (S1), dal momento che (1) si limita ad affermare che l'agente dovrebbe massimizzare i desideri che ha al tempo t . Sembra perciò rispondere all'evidenza che mentre (S1) è temporalmente neutrale, (1) invece non lo è.

Questo contrasto è realmente così forte? Non sembrerebbe, al contrario di quanto si sarebbe indotti di primo acchito a pensare. Quando si prescrive la massimizzazione dei desideri che attualmente si ha, si prescrive anche di non perseguire la massimizzazione dei desideri che si avranno presumibilmente nel corso della propria vita? Il conflitto tra (1) e (S1) è, suggerisco, un conflitto contingente, ma non necessario. È molto spesso del tutto razionale pensare che i desideri che l'agente cerca di massimizzare nel tempo t , potrebbero, almeno alcuni di loro, rientrare in uno specifico sottoinsieme. Questo sottoinsieme comprenderebbe desideri reiterati nel corso del tempo. Lasciando da parte i desideri akratici, è possibile ipotizzare che ci siano desideri che, massimizzati nel presente, estendono il loro potere causale nel futuro, rispondendo così esattamente al requisito richiesto da (S1). Tuttavia, c'è anche un'altra ragione per pensare che il conflitto tra (S1) e (1) non sia un conflitto tra due versioni profondamente differenti della razionalità pratica. Si prenda la notazione precedente sulla motivazione. Se è vero che pensieri, credenze, progetti hanno un potere causale sull'azione, in maniera tale che è ragionevole che l'agente abbia il desiderio di agire in conformità a questi stessi pensieri, credenze, progetti che attualmente ha, allora anche quando l'agente compie qualcosa in conformità a quanto richiesto da (S1), agisce sulla base di un desiderio presente massimizzandolo. Si noti però che può essere avanzata anche un'altra considerazione di un certo interessere. Si tratta di questo: poiché agisco in base a pensieri, credenze, progetti che hanno un potere causale sui desideri, esiste un desiderio soverchiante, che forse sarebbe meglio chiamare meta-desiderio, di agire in conformità ai propri progetti, credenze, progetti. Questo meta-desiderio non può essere negato se non generando contraddizione. Mi spiego con un esempio. Ammettiamo che tu sia preda in periodo particolarmente difficile della tua vita di un determinato automatismo di pensiero, che genera credenze errate e sofferenze. Tutto ciò ha un potere causale sul corso delle tue azioni, in primo luogo perché queste credenze necessariamente generano un desiderio di agire in conformità ad esse. Ammettiamo ora che tu voglia liberarti da quello che percepisci come un meccanismo generale perverso, in maniera tale che i tuoi pensieri, credenze, progetti precisamente non generino il desiderio di agire conformemente ad essi. L'unica risorsa a te accessibile pare essere che tu ti formi altre credenze, pensieri, progetti, ossia che tu intraprenda un processo di revisione critica delle tue credenze attuali al fine di sostituirle con altre credenze. Non vi è maniera di sfuggire a questa

connessione tra credenze e desiderio soverchiante. Anche se tu dovessi abbracciare una soluzione estrema, quale porre fine alla tua vita cosciente, non vi saresti in effetti sfuggito, perché non avresti fatto altro che confermarlo con il tuo gesto. Dopo tale atto estremo non potresti più essere semplicemente definito un soggetto agente.

5. Il legame tra pensieri, credenze, progetti e il desiderio di agire in conformità a questi non può, cioè, essere negato se non sostituendo a determinati progetti, credenze, pensieri altri specifici progetti, credenze, pensieri. Io penso che siamo in presenza di una struttura trascendentale, nel senso che il legame così individuato tra pensiero e desiderio costituisce una struttura necessaria per l'azione. Questa notazione può essere interessante proprio per il requisito della neutralità temporale richiesto da Parfit. Questa struttura dell'azione è temporalmente neutrale, anche se non forse nel senso normativo che vorrebbe Parfit. Per Parfit, infatti, tu non hai nessuna ragione a priori per preferire un determinato segmento temporale a un altro; quindi, non dovresti farlo. Nel caso che illustravo siamo in presenza di qualcosa che è invece strutturale e fondamentale. Il legame pensiero-desiderio è temporalmente neutro sia che tu lo decida sia che tu non lo decida. È semplicemente una descrizione di qualcosa che accade sempre nell'azione. In questo senso, penso che (S1) debba essere incorporata in (1), ossia (S1) deve essere considerata un caso specifico di (1). Si potrebbe sostenere che, in realtà, le cose stanno in maniera esattamente contraria a quanto si sta affermando, e che è (S1) a incorporare (1), sebbene in una versione leggermente modificata di (1). Il suggerimento è stato, in effetti, ancora una volta, avanzato da Parfit. “Secondo tale teoria certi tipi di obiettivi, pur sopravvivendo a tale processo di deliberazione, sono intrinsecamente irrazionali e non forniscono ragioni per l'azione. Ciascuna persona ha più ragione di fare ciò che meglio realizzerà, degli obiettivi attuali, quelli che non sono irrazionali. È la *teoria critica degli obiettivi attuali*”.⁸ Qui si presenta un problema, ossia quello di determinare se esistano effettivamente desideri che siano intrinsecamente irrazionali. ‘Intrinsecamente irrazionali’ significa ‘irrazionali a prescindere dal contesto’. Il contesto comprende (a) sia le informazioni migliori disponibili all'agente, (b) sia lo sfondo precedente delle sue valutazioni pregresse, (c) sia il suo sistema di valori, (d) sia la capacità di rappresentarsi in maniera adeguata i tre elementi precedenti. Trovo effettivamente difficile immaginare un esempio che possa sostenere l'esistenza non contestuale di desideri intrinsecamente irrazionali. Certo, tu

⁸ D. Parfit, *Ragioni e persone*, cit., p. 122.

potresti voler ripetere le imprese di Alessandro Magno o di qualche altro grande condottiero del passato perché pensi di esserne un discendente in linea diretta ed averne così acquisito le abilità strategico-militari incorporandole nel tuo codice genetico. Il desiderio è irrazionale perché il contesto dal quale nasce ha a che fare con la patologia mentale. È molto facile escogitare innumerevoli altri esempi di questo genere ed il caso non è, perciò, particolarmente interessante.

Ammettiamo, tuttavia, che tu sia a conoscenza della struttura generale del dilemma del prigioniero e del risultato sub-ottimale che si ottiene scegliendo la strategia dominante. Ammettiamo che tu scelga ora effettivamente questa strategia. Siamo effettivamente in presenza degli elementi necessari per poter sostenere che tu hai fatto la tua scelta perché preda di un desiderio intrinsecamente irrazionale? Potresti essere stato mosso da considerazioni di sfondo che basterebbero a giustificare la tua azione e a salvaguardare, dal non essere intrinsecamente irrazionale, il tuo desiderio di compiere proprio quella scelta che ti conduce all'esito sub-ottimale. Qualora queste altre considerazioni non fossero presenti e tu semplicemente ti trovassi in una situazione descrivibile formalmente secondo lo schema del dilemma del prigioniero, allora è chiaro che basterebbe la consueta carenza informativa prevista dal dilemma stesso per non rendere la tua azione irrazionale, ma semplicemente aderente alla strategia dominante. Quello che voglio suggerire è che nella procedure deliberative non esistono condizioni oggettive indipendenti dagli agenti che consentano di etichettare a priori un tuo desiderio come 'intrinsecamente irrazionale'. Tutto questo non deve essere confuso con una posizione relativista rispetto ai valori, piuttosto si intende semplicemente affermare il valore strumentale dei desideri. Rispetto a questo aspetto strumentale i valori sono come degli oggetti ideali, e la loro idealità, ossia la loro relativa non contestualità, non esclude affatto un ordine gerarchico interno ai valori.

Si è visto come Parfit invochi la neutralità temporale come un elemento di superiorità di (S1) rispetto ad altre concezioni, ad esempio, a quella della razionalità deliberativa espressa da (1). Si è visto anche la concezione deliberativa e causale di (1) non supporta queste conclusioni di Parfit. È utile qualche ulteriore precisazione riguardo alla concezione della neutralità temporale che qui è in gioco. Infatti, io penso se ne possano distinguere per lo meno due: da una parte, c'è (a) "la nozione di neutralità temporale relativa ai desideri di colui che compie l'azione". Quello che la neutralità temporale prescrive è che i desideri dell'agente nel corso del tempo, per quanto possibile, siano conosciuti e, per quanto è possibile, ne sia conosciuta la loro intensità. La loro importanza dovrebbe essere proporzionale alla loro intensità e

non alla loro collocazione nel segmento temporale della vita dell'agente. Questa concezione ritiene sia importante unicamente l'intensità del desiderio e che questa sia la sola base per considerarlo come desiderio (relativamente) dominante per l'agente. Da un'altra parte, c'è una concezione ancora più esigente della neutralità temporale. Secondo quest'altra concezione, (b) "non ha alcuna rilevanza che un desiderio sia proprio di un agente oppure di un altro, se la loro unica differenza consiste in un differente posizionamento nel continuo temporale". In questo senso, se io ho la possibilità di soddisfare un mio desiderio al tempo $t1$ che è identico al desiderio che un altro avrebbe al tempo $t2$, e io giustifico l'azione che realizza il mio desiderio sostenendo non tanto l'appartenenza di tale soddisfazione a una biografia, quanto la maggiore prossimità al presente, allora agisco in violazione della neutralità temporale. Se io invece giustificassi la mia scelta sostenendo la superiorità valoriale del mio desiderio, non avrei effettuato nessuna violazione della neutralità temporale: in definitiva, poiché si farebbe riferimento a due differenti strumentalità in vista del raggiungimento di due diversi ordini di valori, se ne potrebbe legittimamente concludere che siamo in presenza semplicemente di due desideri diversi.

6. Entrambe queste concezioni sono visioni egualitarie della soddisfazione dei desideri. La seconda concezione è però radicalmente egualitaria, dal momento che mette tra parentesi l'identità personale dell'agente. Questa seconda idea di neutralità temporale e interpersonale presenta molteplici problemi. Il principale tra questi a me sembra essere che ogni deliberazione rischierebbe di richiedere qualcosa di molto prossimo all'etica supererogatoria. Ci si allontanerebbe in tal modo in maniera irrimediabile da qualsiasi cosa possa essere considerata una descrizione plausibile della nostra personale esperienza morale. Inoltre, si appiattirebbe l'esperienza morale su una dimensione sorprendentemente deontologica e irrealisticamente esigente. Di questa interpretazione radicale della neutralità temporale e interpersonale, quindi, non mi occuperò oltre. La prima interpretazione mi sembra invece maggiormente interessante e suscettibile di analisi. Una volta che l'identità personale non viene più considerata come un tratto problematico, ciò che rimane e viene espresso dall'idea di neutralità temporale è semplicemente lo scopo dell'azione razionale, a mio modo di vedere. Tale scopo è la massimizzazione intertemporale della soddisfazione dei desideri dell'agente. Si tratta di una tesi plausibile e sostenibile? Cominciamo da una questione riguardante i desideri passati. Ogni agente è in grado di distinguere i desideri che ha al presente da quelli che ha avuto nel passato. Tra questi ultimi, ve ne sono altri che ha avuto e che

attualmente ha ancora, o in forme identiche o in forme comparabilmente simili a quelli del passato. L'agente dovrebbe tenere presenti questi desideri nel suo calcolo? Da un lato, sembrerebbe fin troppo ovvio sostenere che i desideri passati possono contribuire potentemente a formare le nostre preferenze attuali. Non sembrerebbe quindi esserci motivo per escluderli da una considerazione presente. Questo però ci condurrebbe direttamente a un paradosso. Il peso dei desideri è valutabile dal grado di soddisfazione che io posso assegnare loro. I desideri passati non possono per definizione essere soddisfatti. Si potrebbe argomentare che effettivamente non esiste una soddisfazione positiva dei desideri passati, ma che questo non significa che noi non possiamo collocarli lungo una scala di soddisfazione. Precisamente la loro collocazione in questa scala è dalla parte dei valori negativi. Questa, però, a me pare essere un'inutile complicazione. Il fatto che un desiderio non ci sia più non equivale alla sua frustrazione. Una frustrazione non è tanto una soddisfazione negativa, quanto una soddisfazione negata. Dal momento che soddisfare un desiderio significa pianificare le condizioni materiali che condurranno alla sua realizzazione, allora soltanto i desideri presenti e futuri possono avere soddisfazione, ma non quelli passati, che, quindi, a rigore non possono nemmeno essere frustrati.

Semberebbero considerazioni banali, ma non nel senso che riflettere sulle condizioni di realizzabilità dei desideri comporta anche una sorta di terapia cognitiva per l'agente. Un risultato di questa terapia, per così dire, è che il passato è irredimibile e non può essere scontato. Tu non puoi, quindi, avere obiettivi rispetto al tuo passato, poiché i desideri che avevi non ci sono più. Se tu pensi altrimenti sei chiaramente vittima di un errore cognitivo e il tuo modello deliberativo dovrebbe essere sottoposto a revisione, la quale, come si diceva, avrebbe anche un qualche valore terapeutico. Quello che conta è perciò il futuro. Ma rispetto al futuro quali desideri dovresti avere? Ognuno ne ha molti che possono essere raggruppati in insiemi diversi, ma per i nostri fini è utile distinguere soprattutto due gruppi. Il primo comprende quei desideri che, sulla base delle tue informazioni attuali, se realizzati, condurranno a massimizzare una qualche tua soddisfazione futura. Il secondo gruppo comprende un altro genere di desideri. Si tratta di quei desideri che in qualche modo l'agente ritiene di poter plasmare. Questi desideri giocano un ruolo molto importante nelle nostre vite, dal momento che altro non sono che preferenze autoindotte. Di solito, le preferenze autoindotte – ad esempio, la decisione di intraprendere una determinata carriera, di intrecciare una relazione affettiva stabile, di sottoscrivere un mutuo ventennale per acquistare un appartamento – implicano una pianificazione del futuro. Non solo: implicano anche una diversa concezione della massimizzazione. Rispetto alle preferenze

autoindotte io devo nutrire la convinzione che la loro soddisfazione sia intertemporale. Ad esempio, devo pensare che se ho sottoscritto un mutuo ventennale a rate mensili superiori (ma non troppo) a quanto pagherei per un affitto, la mia soddisfazione sarà massimizzata ad ogni pagamento mensile, anche prima dell'estinzione del mutuo, perché ogni rata pagata mi approssima a quel risultato, anche se a questa soddisfazione devo aggiungere quella relativa all'approssimarsi al risultato finale. Si noti bene che queste preferenze non implicano affatto una qualche etica del sacrificio (che pure deve alimentarsi, soprattutto se non esclusivamente, di preferenze autoindotte), ma è una semplice descrizione di una modalità fondamentale dell'azione che tutti noi adottiamo. Una modalità di massimizzazione diretta di preferenze attuali può invece condurci spesso a risultati indesiderati. Infatti, potrebbe essere che un agente abbia al momento attuale, poniamo, un intenso desiderio di vendetta verso un'amante che lo ha tradito. Si trova anche nelle condizioni di poterla esercitare senza probabilmente subirne le conseguenze. Sarebbe tuttavia irrazionale per lui realizzare ciò che è al momento il desiderio dominante, dal momento che non è certo che dal suo atto non discendano conseguenze negative delle quali potrebbe pentirsi in futuro. Inoltre, poiché ognuno sa che i desideri molto violenti tenderanno ad affievolirsi nel corso del tempo piuttosto rapidamente, l'agente sa già ora che le sue preferenze molto probabilmente tenderanno a cambiare. Facciamo, però, un altro esempio che risulta essere maggiormente convincente. Immaginiamo che tu in assenza di gravi malattie fisiche invalidanti sia profondamente depresso e che tu nutra desideri autodistruttivi. Immaginiamo anche che a uno stadio della tua depressione questi divengano dominanti. Ti comporteresti irrazionalmente se tentassi di soddisfarli, perché è ragionevole pensare che, se i tuoi desideri depressivi non saranno più dominanti in futuro, il saldo netto della rimanente porzione della tua vita potrà essere positivo. In realtà, per fare della soddisfazione dei tuoi desideri autodistruttivi una scelta ampiamente irrazionale sarebbe sufficiente che il tuo saldo netto di soddisfazione per quanto ti rimane da vivere sia leggermente positivo. Infatti, se tu massimizzi la soddisfazione dei tuoi desideri autodistruttivi, alla fine non ci sarebbe nessun saldo positivo. Salvo circostanze molto specifiche (dolorose malattie terminali e atti di sacrificio supererogatori), spesso è irrazionale escludere che non ci possa essere un maggior saldo netto da distribuire lungo l'arco temporale che non sopravanzi un intenso desiderio dominante da soddisfare al presente.

7. Le ragioni estese nel corso del tempo sono particolarmente evidenti in due circostanze generali. Quelle relative all'apprendimento e quelle relative alla

disassuefazione da sostanze psicotrope. Merita esaminare qualche esempio in proposito. Immaginiamo che io sia un dirigente di medio livello di una multinazionale con forti interessi in Cina. Le mie prospettive future di carriera sono legate al fatto che apprenda in un tempo ragionevolmente breve il cinese mandarino. Ho perciò al momento attuale una motivazione ad apprendere il cinese mandarino, non soltanto perché la soddisfazione che potrei trarne adesso, se io lo sapessi già, è soverchiante, ma anche perché è plausibile pensare che lo sarà, poniamo, tra sei mesi. La motivazione a soddisfare il mio desiderio è perciò svincolata dal presente, nel senso che trae la sua forza (e certamente la sua difficoltà) dal non essere limitata all'attualità. Tra sei mesi è perciò probabile che la motivazione ad apprendere il cinese mandarino avrà conservato la sua validità intrinseca, nel senso preciso che tra sei mesi io potrò ancora sostenere che avevo le ragioni di pensare sei mesi fa che dopo sei mesi la motivazione ad un apprendimento probabilmente faticoso avrebbe conservato la sua capacità motivazionale. Sembra che lo stesso si verifichi con i migliori programmi di disassuefazione da droghe. Il consiglio di mantenersi puliti per almeno ventiquattro ore e di pensare di mantenersi puliti nelle successive ventiquattro quando queste si presenteranno e così via, va esattamente nella medesima direzione, poiché confida sul fatto che la forza motivazionale si mantenga per lo meno identica distribuendosi nel corso del tempo.

In nessuno di questi due esempi la capacità di offrire valutazioni sui propri stati futuri è indipendente dalla volontà di soddisfare desideri presenti. Accettare la positività di uno stato ed assegnargli un valore positivo implica, alla luce della discussione precedente, che si ha un desiderio attuale per quello stato, poiché le motivazioni non possono non essere sorrette dai desideri. Il risultato è che anche le motivazioni maggiormente legate a quelli che si ritengono essere dei valori di alto profilo assumono la loro forza se noi li desideriamo e se pensiamo che soddisfarli sia meglio che non soddisfarli. Dal momento che la soddisfazione di questi valori di alto profilo è tipicamente intertemporale, tu devi essere in grado già ora di concepire che, desiderando nel momento attuale la tua soddisfazione relativamente a questi, la desidererai anche nel futuro in maniera analoga. Disporre la soddisfazione in senso intertemporale rimanda alla capacità di sapere come saranno le proprie preferenze nel futuro, ossia rimanda al ruolo dell'immaginazione nell'azione. Per dirla con le parole di Hare "Qui incontriamo nuovamente un'intima relazione concettuale fra gli stati cognitivi, affettivi, e conativi, ma questi ultimi hanno per oggetto degli stati di cose ipotetici, non reali".⁹ Questa dimensione ipote-

⁹ R. Hare, *Il pensiero morale* (1981), Bologna, Il Mulino, 1989, p. 134.

tica è della più grande importanza per renderci comprensibili le esperienze che gli altri fanno e quelle che potremmo fare noi. In altre parole, è parte integrante dell'esperienza umana che io in futuro possa avere le preferenze di un altro. "Supponiamo che io dica: 'Sì, so esattamente come ti senti, ma non mi importerebbe affatto se ora qualcuno facesse lo stesso a me', non dimostrerei forse che in realtà non sapevo, e nemmeno pensavo che ci si sentiva *in quella maniera?*'".¹⁰ In maniera del tutto analoga, se ho delle preferenze per situazioni future, in certa misura devo essere in grado di replicarle al presente, immaginando come mi sentirei se fossi capace di soddisfarle nel segmento temporale in cui mi ora le sto collocando.

Per questo motivo Hare può sostenere che la seguente proposizione: "(1) io attualmente preferisco con intensità I che, se fossi in quella situazione, avvenga X anziché il contrario" unita a "(2) se fossi in quella situazione preferirei con intensità I che X avvenisse anziché il contrario"¹¹ esprimono una verità concettuale. Per quanto non siano affatto proposizioni equivalenti, se uso nel senso usuale il verbo 'sapere' non posso sostenere la verità di (2) a meno che non sottoscriva la verità di (1). Da questa affermazione Hare ne trae un'altra di estremo interesse, ossia l'idea che il termine 'io' non sia per intero un termine descrittivo, bensì inglobi anche una dimensione prescrittiva. "Identificandomi realmente o ipoteticamente con un'altra persona, io mi identifico con le sue prescrizioni".¹² Detto in altre parole, quella che è stata identificata come 'funzione narrativa' indispensabile a pensare la propria esperienza etica come propria, presuppone la capacità di riferire questa esperienza a un nome, ossia a una descrizione definita. Questo non significa affatto dare per risolto il problema dell'identità personale, che comporta molte questioni di ordine differente (gnoseologiche e metafisiche), ma più semplicemente indica la capacità di riferirsi a un nome. Questo spiega anche l'efficacia della deterrenza. "Se considerare la persona che viene punita come me stesso, implica avere un'avversione contro il fatto che egli venga punito pari alla mia avversione futura, questo spiega perché io eviti di commettere il reato per il quale egli viene punito".¹³ Se io tengo conto delle mie attitudini cognitive, posso allora far assumere a quelli che sarebbero dei meri desideri futuri la forza di preferenze attuali.

Questa attitudine simpatetica verso preferenze altrui può essere illuminante per quanto riguarda le ragioni per cui io avrei dei buoni motivi per cu-

¹⁰ R. Hare, *Il pensiero morale*, cit., p. 134.

¹¹ R. Hare, *Il pensiero morale*, cit., p. 136.

¹² R. Hare, *Il pensiero morale*, cit., p. 137.

¹³ R. Hare, *Il pensiero morale*, cit., p. 136.

rarmi di preferenze che potrebbero essere mie. Hare sostiene che tutto questo non coinvolge il problema dell'identità personale. Io penso che occorra essere leggermente scettici sulla validità di questa affermazione, poiché credo sia inevitabile interrogarsi sulle ragioni che effettivamente io potrei avere per interessarmi delle mie preferenze future. L'unica risposta che io riesco a immaginare è una risposta basata su una soluzione specifica del problema dell'identità personale, e precisamente quella basata sulla continuità psicofisica nel corso del tempo. Del resto, è piuttosto facile ammettere che il sentimento di simpatia che io posso provare nei confronti di un altro e che mi permette di immaginare alcuni suoi ordini di preferenze, è un sentimento non originario, ma piuttosto derivato. Per quanto io possa essere simpatetico verso gli altri, lo sarò sempre di meno che nei confronti di me stesso. E se io sono simpatetico verso me stesso non lo sono tanto e soltanto verso l'io che adesso puntualmente sono, bensì piuttosto verso questo io in quanto si nutre delle aspirazioni e dei progetti che precisamente lo costituiscono in quanto è l'io che adesso è, ossia, perché riesco a immaginare una continuità di me stesso nel futuro come unità psicofisica. Non sto sostenendo che il fatto che il mio atto immaginativo sia efficace renda metafisicamente reale questa unità. Ci sono delle buone ragioni, anzi, per continuare a metterla in questione. Quanto sostengo, invece, è che la credenza sulla sussistenza relativa di tale unità è quella che rende ragionevole preoccuparsi delle nostre preferenze future ed anche è quanto fonda il nostro interesse simpatetico per le preferenze degli altri.

8. Mentre questa implicazione mi sembra del tutto corretta e necessaria alla prospettiva simpatetica che si è costretti ad assumere nella valutazione intertemporale delle preferenze, la medesima prospettiva rende anche dubbia la specifica prospettiva di Hare riguardo l'universalizzazione. Hare, infatti, ha ampiamente sostenuto che le nostre ragioni per agire moralmente sono chiaramente fondate perché rispondono a una determinata logica, quella dell'universalizzazione appunto. Se sei in grado di pensare che la tua azione debba essere compiuta da chiunque in circostanze simili alle tue, allora siamo nella sfera morale, altrimenti rimaniamo confinati alla sfera della prudenza personale, se non addirittura dell'egoismo. Tuttavia, il riferimento alla continuità psicofisica dell'agente che io sono non può palesemente essere universalizzabile. Sono io che immagino questa unità futura. Pretendere che non solo venga immaginata, ma anche assunta nei miei medesimi termini sarebbe un narcisismo che inclina verso la patologia. D'altra parte, l'imputazione di quella unità ad altri è operazione non meno dubbia. Tutti sappiamo che le

persone cambiano e non ci scandalizziamo affatto se in circostanze simili, ma posti in due segmenti temporali ragionevolmente distanti, operano scelte o avanzano ragioni per giustificare i propri atti che non sono simili. Questo non ci indirizza verso una eliminazione del pronome ‘io’ dal nostro lessico; anzi: ci sono molte buone ragioni per continuare a mantenerlo non soltanto dal punto di vista legale e giuspositivo della necessità dell’imputazione personale (che per altro comporta eccezioni nello stesso diritto, come è noto), ma anche per l’opportunità di mantenere tale ‘centro focale’ a fini di unità normativa. Certamente, siamo qui in presenza di una ambiguità. Da un lato, infatti, l’universalizzazione richiede l’eliminazione di tutte le variabili individuali che rendono sensato, narrativamente ed eticamente sensato, l’uso del pronome ‘io’, in senso specifico, ossia come *token*. Da un altro lato, secondo una interpretazione più stretta dell’universalizzazione, la motivazione di un’azione può essere universalizzabile a patto che in essa precisamente non compaiano riferimenti né a variabili né a costanti individuali. Questa versione dell’universalizzazione, che a me pare una versione particolarmente stringente di kantismo, ritiene che il riferimento all’io tanto come *token* quanto come *type*, renda l’atto di universalizzare la motivazione impuro. Infatti, io posso interpretare le preferenze di un soggetto A come accettabili da un soggetto B, senza che questo implichi in alcun modo che un terzo soggetto C debba farle proprie. Questo, in realtà non varrebbe nemmeno per B, se noi accettiamo l’idea che il pronome ‘io’ implichi una prescrizione, prescrizione in linea con una unità psicofisica nel corso del tempo. Infatti, né B né C sono continui con A; per quanto simili possano essere le loro preferenze, questo non significa che debbano adottare quelle di A per se stessi. Per quanto simili siano le loro motivazioni, questo non significa che debbano riconoscere un paradigma di generalità in A tanto forte da trasformarsi in universalità per ciascuno di loro. In definitiva, quello che l’interpretazione dell’io come termine descrittivo e prescrittivo richiede è una capacità di immedesimazione tale da poter immaginare che la soddisfazione delle preferenze di altri soggetti simili a lui per aspetti rilevanti, abbiano lo stesso peso della soddisfazione delle proprie preferenze. Per questo motivo Hare può scrivere. “Ora, i casi multilaterali presentano meno difficoltà di quanto sembrasse a prima vista. Infatti, anche in tali casi i conflitti interpersonali, per quanto complessi siano e per quante persone vi siano coinvolte, si ridurranno a conflitti intrapersonali, posto che si dia una completa conoscenza delle preferenze altrui”.¹⁴ In realtà, questa mossa è cruciale per l’idea che la logica interna del prescrittivismo conduca a una forma di utilitarismo. Ma se le cose stanno in questo mo-

¹⁴ R. Hare, *Il pensiero morale*, cit., p. 152.

do, a me pare che siamo precisamente ricondotti a quella forma particolarmente esigente di universalizzazione di cui si diceva sopra. Il fatto è che l'identità numerica dovrebbe essere esclusa da una forma rigorosa di utilitarismo, poiché la fonte di legittimazione per la soddisfazione di alcune preferenze anziché di altre si ridurrebbe a una forma di pregiudizio, quel pregiudizio che ci fa usare il pronome 'io' al tempo stesso come un indicatore di una descrizione definita, che può quindi essere sostituita da un nome proprio, e come qualcosa che può essere assunto in linea di principio da chiunque altro, senza che vengano alterate le condizioni di prescrittibilità di una motivazione specifica, condizione senza dubbio estremamente esigente. Se queste osservazioni sono pertinenti se ne dovrebbe ricavare che la derivazione dell'utilitarismo dal prescrittivismismo, che costituisce l'ambizione teorica principale di Hare, non è destinata ad essere soddisfatta, per lo meno in questi termini. Inoltre, possono essere avanzati dei dubbi sulla legge psicologica che sembra sottostare alla posizione prescrittivistica. Questa legge o intuizione psicologica suona più o meno in questi termini: "se io giudico che esista un ente psicofisicamente continuo con me stesso, allora sarà razionale che desideri che le preferenze di tale ente siano soddisfatte in misura proporzionale alla loro intensità". In altre parole, la mia preoccupazione sulla soddisfazione delle preferenze di tale ente è garantita dalla sua continuità con me stesso. Ma ammettiamo che io vada a dormire e che per qualche sconosciuto processo atomico il mio corpo e la mia mente si dissocino per essere ricomposti quando suona la sveglia, per avvertirmi che devo alzarmi e prepararmi per andare a lezione, in una unità che è leggermente diversa da quella che avevo quando sono andato a dormire, leggermente diversa ma indistinguibile agli effetti macroscopici. Avrei davvero acquisito qualche motivo per non preoccuparmi della soddisfazione delle preferenze di questa entità leggermente discontinua con quella che esisteva quando sono andato a coricarmi? Noi sappiamo che nella realtà le cellule del nostro corpo vengono sostituite progressivamente. Non pensiamo che questo comporti una maggiore o minore preoccupazione per la soddisfazione delle preferenze di tale entità che muta. Per quale motivo, allora, le cose dovrebbero cambiare se il ricambio cellulare dovesse avvenire simultaneamente? Quello che vorrei suggerire è che sembra ovvio che l'identità personale debba far riferimento a una qualche forma di continuità spazio-temporale, poiché la sola continuità materiale non ne è né condizione sufficiente, ma nemmeno condizione necessaria. Preoccuparsi della soddisfazione delle preferenze di un sé successivo, cosa che normalmente accade quando si ritiene di essere in presenza di una qualche forma di continuità dei sé, non è causato dalla mera continuità, ma piuttosto dal fatto che questa forma sia una somiglianza prossima a un qualche sé passato. Esistono

molto probabilmente delle pressioni evolutive che fanno sì che noi, esseri dotati di autocoscienza, reagiamo in maniera così pronta e sensibile alla somiglianza dei nostri sé successivi. Dovessimo comportarci altrimenti, probabilmente non sopravvivremmo a lungo. Che questo sia un investimento emotivo o razionale o più probabilmente un intreccio tra i due non penso sia necessario e importante stabilirlo ai nostri fini, ma è questo investimento a rendere ragione del fatto che ci preoccupiamo maggiormente dei sé che ci sono prossimi anziché di quelli che ci sembrano estranei. Questa era stata anche l'intuizione di Hume, quando aveva rilevato che i meccanismi della simpatia non si estendono più di tanto nel tempo e nello spazio. Hume non pensava tanto alla simpatia nei confronti di sé successivi in quel contesto, quanto alla simpatia che ci è difficile estendere oltre la cerchia di coloro che ci sono più prossimi. Nei termini della presente discussione, tuttavia, la cosa potrebbe essere formulata nei seguenti termini: ci può essere somiglianza anche senza continuità. La legge psicologica che sembrerebbe stare alla base della razionalità a preoccuparsi della soddisfazione delle preferenze di sé ritenuti successivi in virtù della continuità spazio-temporale deve perciò essere abbandonata perché non necessaria, dal momento che ci può essere somiglianza senza che ci sia continuità. Ma deve essere abbandonata anche per un altro motivo, ossia perché ci può essere continuità senza che ci sia somiglianza. Mi rendo conto che mentre la somiglianza senza continuità può essere facilmente accettata, per quanto riguarda il secondo caso le cose sembrano, almeno intuitivamente, maggiormente controverse. Eppure casi di questo genere sono tutt'altro che rari. Immagina che tu sia una persona assolutamente media, rispettosa della legge, leale verso i tuoi amici e verso la tua famiglia. Per uno di quegli strani casi della vita che mai avresti contemplato, un giorno assumi una droga che ti costringe da subito a un'elevata dipendenza, facendoti divenire un tossicodipendente la cui principale preoccupazione è procurarsi la prossima dose, in spregio alla legge e violando ogni regola di lealtà verso i tuoi amici e verso la tua famiglia. Ammettiamo pure che tu sia stato costretto a diventare un tossicodipendente e che quindi tu non sia responsabile di ciò che sei diventato. Dal momento che prima di essere un tossicodipendente non avresti mai voluto diventarlo, nel senso preciso che non avresti riconosciuto le preferenze di un tossicodipendente come tue, non si può certo dire che il tuo sé anteriore avrebbe dovuto in qualche modo preoccuparsi della soddisfazione delle preferenze di un sé che mai avrebbe voluto essere.

Ammettiamo, tuttavia, che il presupposto della continuità psicofisica sia valido. Anche in questo caso non ne conseguirebbe nient'altro che questo: che io ho una ragione in più per desiderare la soddisfazione delle preferenze espresse da qualche sé successivo di quante ne avrei rispetto a un sé che non

sia successivo. Si è detto anche che l'intensità delle preferenze di un sé successivo dovrebbe essere quanto entra nella considerazione relativa alla decisione di soddisfarle. Naturalmente, l'intensità delle preferenze è una delle caratteristiche principali che entrano in gioco quando decidiamo che valga o meno la pena di soddisfare qualche nostro desiderio, ma non è certo l'unica e nemmeno quella più rilevante in ogni circostanza. Noi valutiamo i nostri desideri anche in base ad altri criteri, che sono criteri che sinteticamente possono essere chiamati criteri di virtù e non hanno a che fare semplicemente con la massimizzazione. Piuttosto, tali criteri rappresentano un vincolo alla massimizzazione. Tale vincolo entra come contenuto dei nostri desideri. Mi riferisco ad oggetti quali la lealtà, il desiderio di non manipolare gli altri, di non trarre ingiusto profitto dalle nostre azioni. Tali contenuti hanno a che fare con la nostra descrizione percepita come maggiormente autentica e sono precedenti alla massimizzazione. Ma anche se immaginiamo un agente puramente massimizzante, tale agente avrebbe dovuto pur sempre decidere che la massimizzazione e non altro è ciò che conta come contenuto e oggetto dei propri desideri e tale decisione sarebbe precedente l'altra di massimizzare le proprie preferenze. Sarebbe, in altre parole, una modalità di descrivere l'agente stesso. Nella struttura per così dire trascendentale della propria descrizione è racchiusa una preoccupazione per la coerenza delle serie temporali che si devono riferire a operatori sufficientemente prossimi al sé attuale.

Un agente razionale porrà attenzione al fatto che per quanto è in suo potere, non vengano fatti sorgere dei desideri futuri la cui soddisfazione sarebbe in contrasto con quella che pensa essere la sua autorealizzazione attraverso il tempo. A meno che l'agente non sia animato da una narcisistica e irrealistica volontà di potenza, un agente razionale deve riconoscere che le sue azioni si svolgono e i propri desideri prendono forma in un mondo parzialmente deterministico, nel senso che la possibilità di agire e la soddisfacibilità delle nostre preferenze eccedono la capacità di controllo di chiunque a causa della indefinibilità delle relazioni cooperative coinvolte e dell'incompletezza dell'informazione. Questa notazione del tutto ragionevole può tuttavia generare dei notevoli problemi proprio nella prospettiva della massimizzazione vincolata, che comunque mi pare quella maggiormente adeguata a una concezione narrativa dell'agire. Infatti, poniamo il caso che io sappia che qualcuno agirà o qualcosa avrà luogo con un'alta probabilità nel futuro prossimo, in maniera tale da modificare profondamente il mio carattere e da farmi fare cose che io al momento attuale non farei. L'agente dovrebbe accondiscendere a questa possibilità per il semplice fatto che pensa di non essere in grado di opporvisi? Nella maggior parte dei casi noi immaginiamo che le cose non siano affatto così semplici e non bolliamo come irrazionale un comportamento

anche se si svolge in condizioni molto avverse. Si pensi a quanto è accaduto nei regimi totalitari. La forza di penetrazione e persuasione di tali regimi non è dovuta solo alla loro forza repressiva, ma anche a un fenomeno psicologico che inducono in coloro che sono oppressi: tali regimi appaiono molto stabili. E tuttavia non mancano mai in essi fenomeni variegati e significativi di resistenza. Tale resistenza è motivata dalla convinzione che tali sistemi politici modificano oltre una determinata soglia critica il carattere di coloro che vi resistono. Questi comportamenti smentiscono la linea di minore resistenza, per così dire, nella soddisfazione di preferenze future che non si vorrebbero avere. Vi è un altro esempio che mi pare suggerire che il requisito della coerenza delle preferenze nel corso del tempo rappresenti senz'altro delle stigmate di razionalità impresse all'agente. Supponiamo che per un agente sia coerente avere un gruppo di preferenze *b* nel futuro non analoghe a un altro gruppo di preferenze *a* che ha già soddisfatto nel passato. In un qualche senso, inoltre, l'agente sa che queste preferenze costituiscono parte importante della propria autodescrizione. Tuttavia, decide di non soddisfarle. È sufficiente questo per sostenere che si sta comportando irrazionalmente? Non credo. Infatti, se tali preferenze sono considerate dall'agente come anti-sociali, potrebbe avere delle ragioni molto valide per non soddisfarle. È innegabile che esistano individui che possiedono tali tendenze ed decidono consapevolmente di non dar loro corso. E magari ognuno di noi si sarà potuto trovare in qualche situazione nella quale avrebbe potuto dar sfogo a un forte impulso anti-sociale senza subirne le conseguenze, eppure non lo ha fatto. Esistono poi anche gli atti di autopunizione che possono essere fatti rientrare in una fattispecie analoga. Tutti questi si presentano come atti non massimizzanti.

Penso che a questo punto sia possibile trarre qualche conclusione almeno provvisoria dalla nostra discussione. Abbiamo visto che la proposta influente di Hare è di assegnare al termine 'io' un significato tanto descrittivo quanto prescrittivo. Si è visto anche che è possibile formalizzare questa posizione esprimendola nella forma di una legge psicologica, la quale afferma che se qualcuno giudica che un sé sia un suo proprio successore psicofisico prossimo, allora è razionale avere come obiettivo la soddisfazione delle sue preferenze in ragione della loro intensità. Si è visto però che esistono ragioni per pensare che questa legge abbia una universalità esclusivamente presunta. Queste ragioni fanno perciò pensare che non sia vera. In effetti, questa legge psicologica, se forse non può essere attribuita ad Hare nella sua forma letterale, gli va ascritta in senso profondo, nella misura in cui realizza una congiunzione tra una posizione fattuale sull'identità personale – la continuità psicofisica dei sé – e una attitudine prescrittiva – quella relativa alla simpatia –. Ciò che si

dovrebbe concludere è che questo legame tra attualità e disposizione alla simpatia non è adeguatamente supportato né in una forma larga né in una forma stretta. Sembrano non esserci ragioni valide per soddisfare delle preferenze sulla sola base del fatto che apparterranno a un successore prossimo; né sembra che ve ne siano per scegliere tra diverse preferenze di un successore sulla sola base del fatto che si dovrebbe scegliere in base alla loro intensità. Appartenenza e massimizzazione possono essere criteri irrilevanti nel decidere la soddisfazione di alcuni desideri di fronte ad altri, proprio in virtù di caratteristiche considerate indispensabili alla propria autodescrizione.