

IL PROBLEMA DELL'IDENTITÀ PERSONALE NELLA FILOSOFIA ANALITICA CONTEMPORANEA

1. INTRODUZIONE

Questo XXI colloquio maceratese sull'interpretazione verte sul tema dell'individualità con particolare enfasi su quel peculiare tipo di individualità costituito dal vivente e soprattutto dalla persona umana. Nel suo intervento, il professor Alici si è dunque giustamente soffermato sul tema della persona, mettendo molto bene in rilievo alcune delle problematiche più importanti che su questo terreno entrano in gioco da un punto di vista filosofico. Lo ha fatto soprattutto nella prospettiva dell'ermeneutica filosofica continentale, in particolare in quella di Ricoeur, contrapponendola ad approcci fisicalisti nell'ambito della filosofia analitica. Il contributo del professor Alici offre già un quadro significativo della ricchezza della riflessione filosofica contemporanea sul tema della persona. In questo breve intervento cercherò di mostrare ancora qualche altro aspetto di questa riflessione, soffermandomi su come la filosofia analitica contemporanea ha affrontato e sta affrontando il tema dell'identità personale, sia da un punto di vista fisicalista, sia da punti di vista non strettamente riducibili al fisicalismo.

2. DUE DOMANDE SULL'IDENTITÀ PERSONALE

Il tema dell'identità personale in filosofia analitica si dipana da queste due fondamentali domande:

- (D1) Quali criteri ci permettono di "individuare" delle persone in un dato momento t (ossia di "contarle" come numericamente differenti)? Per esempio, che cosa ci permette di dire che ci sono, supponiamo, 30 *persone* in questa stanza?
- (D2) Quali criteri ci consentono di "identificare nel tempo" una stessa persona? Ossia, in base a quali criteri diciamo, per esempio, che la persona X esistente al momento t (per esempio, il governatore della Banca d'Italia il 26 marzo 1977) è la persona Y esistente al momento t' (per esempio, il presidente della Repubblica Italiana il 20 Marzo 2001)?

La prima domanda in effetti ci invita a chiederci che cosa significa essere una persona, in contrapposizione ad un essere inanimato o vivente, ma privo dei requisiti necessari per essere considerato "persona". La seconda domanda è quella che più da vicino caratterizza il tema dell'identità personale così come viene inteso dai filosofi analitici. La prima domanda è però presupposta nella seconda. Difficilmente possiamo rispondere alla seconda senza avere risposto alla prima, e, come si vedrà, la risposta data alla prima domanda condiziona la risposta data alla seconda domanda.

3. POSSIBILI RISPOSTE ALLA PRIMA DOMANDA

Nel movimento analitico, almeno a partire dal secondo dopoguerra, le risposte alla prima domanda sono prevalentemente di tipo anti-cartesiano, poiché negano una rigida demarcazione tra una *res cogitans* (la mente) distinta da una *res extensa* (il corpo), una cui parte (il cervello) media tra la prima e la seconda. All'interno di questa prospettiva anti-cartesiana, possiamo grosso modo distinguere questi due modi di rispondere alla prima domanda:

- (D1/AC1) Criterio dell'identità (mente-cervello). Per sapere quante persone (menti con un certo tipo di complessità) ci sono, basta sapere quanti sono i cervelli funzionanti di sufficiente complessità (cervelli di esseri umani).

(D2/AC2) Criterio della dipendenza (del mentale dal fisico). Per sapere quante persone (menti con un certo tipo di complessità) sono presenti in una stanza, basta sapere quanti "software mentali" sono implementati in entità fisiche presenti nella stanza, oppure quanti "flussi di coscienza privati" sono supportati da entità fisiche presenti nella stanza.

La prima risposta si sposa in modo naturale alla cosiddetta *teoria dell'identità*. Secondo questo punto di vista, una persona, in quanto entità dotata di proprietà mentali, si può in ultima analisi identificare con un cervello "funzionante". Infatti, secondo la teoria dell'identità, tutte le proprietà mentali di una certa persona sono identiche a proprietà fisiche del cervello della persona in questione, nel senso in cui la proprietà di essere acqua è identica alla proprietà di essere un composto di molecole di H_2O . Per esempio, la proprietà mentale di soffrire per un mal di denti della persona *X* è la proprietà del cervello di *X* di trovarsi in un certo stato elettrochimico. Data l'identità di proprietà mentali e cerebro-fisiche è legittimo, secondo questa prospettiva, accettare l'equazione mente = cervello. Questo punto di vista è stato proposto dapprima da Ullian Place nel 1956¹ ed ha poi goduto di notevole popolarità fino agli anni '70, soprattutto ad opera di filosofi quali Smart e Armstrong.

La seconda risposta si sposa in modo naturale ad almeno due prospettive, il *funzionalismo* e il *fisicalismo non riduttivista*.

In un tentativo di ristabilire l'autonomia della psicologia², filosofi quali Putnam e Lewis hanno proposto l'ipotesi funzionalista³, molto popolare soprattutto negli anni '70 e '80, secondo la quale le proprietà mentali sono "stati funzionali" che almeno in linea di principio sono "realizzabili" in un computer fatto di materia inorganica, tanto quanto lo sono in cervelli viventi fatti di materia biologica. Da questo punto di vista, il nostro avere proprietà mentali dipende dal fatto che nei nostri cervelli sono implementati dei "programmi" (nel senso informatico del termine) molto sofisticati. Questo "software mentale" permette che il nostro cervello funzioni in modo da avere proprietà quali credere che Parigi è la capitale della Francia, desiderare un gelato, soffrire per un mal di denti, e così via. Quando la branca dell'informatica chiamata "intelligenza artificiale" avrà compreso bene la natura di questi "programmi mentali", sarà possibile implementarli in computer inorganici, creando così delle intelligenze artificiali paragonabili agli esseri umani. Questo (con le necessarie semplificazioni) il punto di vista del funzionalismo; nella cui prospettiva, potrebbe non essere sufficiente contare il numero di cervelli umani funzionanti per sapere quante persone ci sono. Potrebbe esservi un computer con un programma di intelligenza artificiale sofisticato al punto da farci ammettere che vi è una persona non "localizzata" in un cervello biologico. Dal punto di vista funzionalista quindi vi sono tante persone quanti sono i programmi (implementati) capaci di generare stati funzionali classificabili come "mentali".

Il fisicalismo non riduzionista può sposarsi con il funzionalismo, ma è indipendente da esso. Il punto centrale (con le necessarie semplificazioni) è che le proprietà mentali non vanno identificate rigidamente con proprietà fisiche di un certo tipo. Per esempio, potrebbe darsi che un essere umano ed un pesce, che hanno entrambi una ferita, non hanno nei loro cervelli una stessa proprietà fisica identificabile con la proprietà mentale di provare dolore (poiché i due cervelli sono troppo diversi). Eppure entrambi, si può supporre, provano dolore. Quindi, il provare dolore, e in generale qualsiasi proprietà mentale, non è una proprietà fisica, anche se ha sempre bisogno di una base fisica di un qualche tipo perché si

¹ U. Place, "Is Consciousness a Brain Process?", in C. Borst (a cura di), *The Mind-Brain Identity Theory*, McMillan, Londra, 1970.

² La teoria dell'identità in effetti propone una riduzione del mentale al neurofisiologico e non sembra lasciare spazio ad uno studio autonomo dei fenomeni psicologici.

³ Cfr. H. Putnam, *Mind, Language and Reality, Philosophical Papers, Volume 2*, Cambridge University Press, Londra, 1975 (trad. it., *Mente, linguaggio e realtà*, Adelphi, Milano, 1987); D. Lewis, "Psychophysical and Theoretical Identifications", in *Australasian Journal of Philosophy*, 50, 1972, pp. 249-258.

manifesti. Se si scarta il funzionalismo, questa base fisica deve presumibilmente essere di natura organica. Questo punto di vista si è affermato a partire dagli anni '80 di fronte a difficoltà emerse nelle due prospettive considerate sopra⁴. Dal punto di vista del fisicalismo non riduzionista, possiamo dire che vi è una persona, quando vi è un certo fascio di proprietà mentali appropriatamente legate tra loro in modo tale da costituire una "unità", che chiameremo "flusso di coscienza unitario"⁵, la cui esistenza dipende dall'esistenza di una certa entità fisica (un certo cervello funzionante), ma che non è identificabile con tale entità fisica.

Per quanto la prospettiva anti-cartesiana è stata e continua a essere dominante in filosofia analitica, il punto di vista cartesiano, secondo il quale le proprietà mentali sono proprietà di una *res cogitans* distinta dal corpo, ha (o ha avuto nel recente passato) i suoi rappresentanti, talvolta personalità di spicco quali C. J. Ducasse, R. Chisholm o il premio Nobel J. C. Eccles⁶. Da questo punto di vista, una persona è un'entità capace di persistere nel tempo e non identificabile con il cervello o con qualsiasi altra entità fisica. Chisholm per la verità ha concesso (non molto cartesianamente) che tale entità possa anche essere di natura fisica (per esempio, una particella fisica elementare), purché indivisibile. Questa ipotesi sembra però più che altro una concessione allo "spirito del tempo", e in ogni caso non è rilevante per le questioni che considereremo nel seguito. In questa prospettiva cartesiana, alla prima domanda possiamo così rispondere:

(D1/C) Criterio della *res cogitans*. Quante persone vi sono dipende da quante menti "indivisibili" (di sufficiente complessità, in un senso non spaziale) vi sono. Il fatto che sappiamo che vi sia un cervello funzionante è solo un criterio epistemologico che ci porta ad assumere che ("collegata" a tale cervello) vi è anche una mente, ma quest'ultima è distinta dal cervello stesso.

4. POSSIBILI RISPOSTE ALLA SECONDA DOMANDA

Tralascieremo nel seguito (per semplicità) l'idea funzionalista della mente come software e quindi l'ipotesi di persone create artificialmente con i metodi dell'intelligenza artificiale. Pur limitandoci a persone dotate di (o "collegate a") cervelli organici, variano le risposte offerte alla seconda domanda, a seconda della risposta data alla prima domanda. Se si assume il criterio dell'identità (D1/AC1), emerge in modo naturale questa risposta:

(D2/AC2) Criterio della persistenza del cervello. Assumendo l'equazione mente = cervello, segue che il persistere nel tempo di una persona non è altro che il persistere nel tempo di un certo cervello funzionante. Questo è il punto di vista difeso per esempio da Thomas Nagel⁷. In questa prospettiva, il governatore della Banca d'Italia il 26 marzo 1977 è il presidente della Repubblica Italiana il 20 Marzo 2001, perché uno stesso cervello è durato nel tempo rivestendo, per così dire, le funzioni di governatore prima e di presidente poi. (Quindi, i meriti o le colpe ascrivibili al governatore sono ascrivibili al presidente).

Se invece si assume il criterio della dipendenza (D1/AC2), si tende ad approdare a questa risposta:

(D2/AC2) Criterio della relazione. Siccome il cervello non è che la "piattaforma", per quanto necessaria, di un flusso di

⁴ Sul fisicalismo non riduzionista si veda J. Kim, *La mente e il mondo fisico*, McGraw-Hill Italia, Milano, 2000.

⁵ L'idea della mente come fascio di proprietà mentali risale quanto meno a David Hume.

⁶ Cfr. C. J. Ducasse, *Nature, Mind and Death*, Open Court, La Salle, Illinois, 1951; K. Popper e J. C. Eccles, *The Self and Its Brain*, Springer, Berlino, 1981; R. M. Chisholm, *On Metaphysics*, University of Minnesota Press, Minneapolis, 1989.

⁷ Cfr. T. Nagel, *The View from Nowhere*, Oxford University Press, Oxford, 1986, Cap. III.

coscienza unitario, perché si possa parlare del persistere di una certa persona dal momento t al momento t' , è necessario che vi sia una certa relazione di continuità tra due flussi di coscienza unitari F (esistente al momento t) ed F' (esistente al momento t'); relazione data, per esempio, dal persistere di certe memorie, caratteristiche caratteriali, aspirazioni, ecc., insomma (semplificando un po') una "sufficiente somiglianza". Inoltre, è necessario che questa relazione sussista tra F ed F' più di quanto non sussista tra F ed altri flussi diversi da F' esistenti al momento t' . Questa prospettiva è difesa vigorosamente soprattutto da Derek Parfit⁸. Seguendo Parfit, chiameremo "R" la relazione in questione. Secondo questo modo di vedere, quando asseriamo che il governatore della Banca d'Italia il 26 marzo 1977 è il presidente della Repubblica Italiana il 20 Marzo 2001 non asseriamo un'identità (o almeno non dovremmo), ma semplicemente che un certo flusso unitario di coscienza che esisteva il 26 marzo 1977 ha la relazione R con un altro flusso unitario di coscienza che esisteva il 20 Marzo 2001 (e con nessun altro flusso esistente in quella data).

Se invece accettiamo il criterio della *res cogitans*, approdiamo a questo punto di vista:

- (D2/C) Teoria della persistenza della *res cogitans*. Il persistere nel tempo di una persona non è altro che il persistere nel tempo di una certa mente indivisibile. Da questo punto di vista, sapere che persiste un certo cervello funzionante, o che sussiste una relazione R, dà luogo al massimo a criteri epistemologici che ci possono guidare (fallibilmente) a supposizioni riguardanti il persistere o meno di una certa persona.

Chisholm ha difeso questa posizione, prendendo spunto dalle critiche del vescovo Butler (1692-1752) alla teoria di Locke dell'identità personale⁹.

5. CRITERI IN DISACCORDO.

Il dibattito su queste questioni ha messo in rilievo che le due prospettive anti-cartesiane, i criteri dell'identità e della persistenza da un lato e i criteri della dipendenza e della relazione dall'altro, possono entrare in conflitto tra loro, alla luce di dati empirici o di esperimenti mentali, contrariamente a quanto si potrebbe pensare ad un primo esame. Questo conflitto è interessante anche nella prospettiva cartesiana, poiché il cartesiano è interessato comunque alle risposte anti-cartesiane, anche se in chiave epistemologica e non ontologica.

Per quanto riguarda la prima domanda, ci si è interrogati su pazienti che hanno subito la bisezione cerebrale nel tentativo di curare forme molto gravi di epilessia. Tali pazienti sembrano presentare due flussi unitari di coscienza, distinti e non reciprocamente consapevoli, in uno stesso cervello, l'uno "legato" all'emisfero destro e l'altro all'emisfero sinistro¹⁰. Analogamente, in casi di cosiddetta "personalità multipla", tipicamente dovuti a gravi traumi infantili, sembra vi siano flussi di coscienza unitari distinti e separati all'interno di un unico cervello, che a turno "comandano" il corpo, certe volte con un flusso capace di "osservare telepaticamente" un altro flusso¹¹. In casi del genere, sembrerebbe che dal punto di vista del criterio dell'identità (D1/AC1) dovremmo dire che c'è una sola persona (un solo cervello), mentre dal punto di vista del criterio

⁸ Cfr. *Reasons and Persons*, Clarendon Press, Oxford, 1984. Il criterio relazionale di Parfit si può far risalire a John Locke.

⁹ Cfr. Chisholm, cit., Cap. 4.

¹⁰ Cfr. R. W. Sperry, "Brain Bisection and Mechanisms of Consciousness", in J. C. Eccles (a cura di), *Brain and Conscious Experience*, Springer, Berlino, 1966, pp. 298-313.

¹¹ Stephen E. Braude, *First person Plural*, Routledge, Londra, 1991.

della dipendenza (D1/AC2), dovremmo dire che ci sono due o più persone (diversi flussi di coscienza).

Per quanto riguarda la seconda domanda, seguendo Parfit (1984, p. 255) si potrebbe argomentare come segue (cit., Cap. 12): i pazienti che hanno subito la bisezione dimostrano che un flusso di coscienza si può dividere in due. Quindi, la completa divisione di un cervello in due parti è soltanto (per ora) una difficoltà tecnica. In linea di principio, un cervello funzionante *X* potrebbe essere diviso in due emisferi funzionanti *Y* e *Z*, trapiantati in due diversi corpi. In una tale situazione, non c'è più il persistere di un unico cervello, quindi dal punto di vista del criterio della persistenza del cervello, si potrebbe dire che una certa persona ha cessato di esistere. Ma dal punto di vista del criterio relazionale (D2/AC2) si può ipotizzare che la relazione *R* sussiste (in misura maggiore) tra il flusso di coscienza legato a *X* prima della divisione e quello legato a *Y*, oppure quello legato a *Z*. In entrambi i casi, si può parlare di persistenza di una stessa persona. Si possono anche immaginare altri esperimenti mentali che portano ad una analoga divergenza tra i due punti di vista anti-cartesiani. Per esempio, si può immaginare un graduale rimpiazzamento dei neuroni di un certo cervello *X* con altri neuroni (o con microchips) "funzionalmente equivalenti" (come nella classica storia della nave di Teseo). Oppure si può immaginare che con la "teletrasportazione" venga fatta su Marte una copia identica atomo per atomo di un cervello *X* che si trova sulla Terra. Alla fine della copiatura il cervello originale *X* viene distrutto e la copia *Y* rimane (Cfr. Parfit, cit., Cap. 10). In entrambi i casi, alla fine ci troveremo con un cervello *Y* completamente diverso da *X*. Ma presumibilmente il flusso di coscienza che era legato a *X* ha la relazione *R* con quello legato a *Y*.

Forse questi dati empirici ed esperimenti mentali possono indurre a preferire i criteri della dipendenza e della relazionale rispetto a quelli dell'identità e della persistenza del cervello. Però questa scelta ha conseguenze controintuitive. Per esempio, nel caso del cervello diviso in due emisferi trapiantati in due corpi diversi, stabilire se la persona "collegata" al cervello *X* precedente alla divisione "andrebbe con" *Y* oppure con *Z* sembra dipendere dalle convenzioni che decidiamo di adottare nel caratterizzare la relazione *R*, senza che vi sia una verità di fatto al riguardo. Ma se ciò è vero, sembrerebbe che le risposte a domande come "io sarò domani la persona *X* oppure la persona *Y*?", o addirittura "io esisterò domani?" si possano dare su base convenzionale. Tuttavia, il nostro punto di vista soggettivo sembra reclamare che le risposte a queste domande siano verità di fatto e non decisioni convenzionali. Su questa base Chisholm (cit., Cap. 4) argomenta che *dobbiamo* ammettere una mente indivisibile, perché il nostro schema concettuale *presuppone* che queste questioni sono fattuali e non convenzionali [rivendicando quindi la legittimità dei criteri cartesiani (D1/C) e (D2/C)]. Dal punto di vista di Parfit, ammesso che sia così, si dovrebbe semplicemente concludere che il nostro schema concettuale incorpora una teoria sbagliata e va quindi cambiato (per arrivare sostanzialmente allo schema concettuale del buddismo¹²). Secondo Parfit, un tale cambiamento è anche vantaggioso, perché permette di addomesticare la paura della morte e di favorire l'altruismo. Infatti, dal punto di vista di Parfit, quello che chiamiamo il sopravvivere di una certa persona, Mario, non è che un continuo "morire" di flussi di coscienza sostituito da nuovi flussi di coscienza in una catena in cui i primi membri della serie (legati a un bambino) possono essere molto diversi dagli ultimi (legati a un anziano) e magari più simili a membri di un'altra serie di flussi di coscienza, che chiamiamo il sopravvivere di un'altra persona, Paolo.

6. CONCLUSIONE

Non posso qui argomentare a favore dell'una o dell'altro prospettiva. Il mio scopo è stato semplicemente quello di offrire una piccola finestra sul dibattito riguardante l'identità personale nella filosofia analitica contemporanea.

¹² Cfr. Parfit, cit., Cap. 12, § 92.